

Volume 12 Issue 6

June 2021



ISSN 2156-5570(Online)

ISSN 2158-107X(Print)



Editorial Preface

From the Desk of Managing Editor...

It may be difficult to imagine that almost half a century ago we used computers far less sophisticated than current home desktop computers to put a man on the moon. In that 50 year span, the field of computer science has exploded.

Computer science has opened new avenues for thought and experimentation. What began as a way to simplify the calculation process has given birth to technology once only imagined by the human mind. The ability to communicate and share ideas even though collaborators are half a world away and exploration of not just the stars above but the internal workings of the human genome are some of the ways that this field has moved at an exponential pace.

At the International Journal of Advanced Computer Science and Applications it is our mission to provide an outlet for quality research. We want to promote universal access and opportunities for the international scientific community to share and disseminate scientific and technical information.

We believe in spreading knowledge of computer science and its applications to all classes of audiences. That is why we deliver up-to-date, authoritative coverage and offer open access of all our articles. Our archives have served as a place to provoke philosophical, theoretical, and empirical ideas from some of the finest minds in the field.

We utilize the talents and experience of editor and reviewers working at Universities and Institutions from around the world. We would like to express our gratitude to all authors, whose research results have been published in our journal, as well as our referees for their in-depth evaluations. Our high standards are maintained through a double blind review process.

We hope that this edition of IJACSA inspires and entices you to submit your own contributions in upcoming issues. Thank you for sharing wisdom.

Thank you for Sharing Wisdom!

Kohei Arai
Editor-in-Chief
IJACSA
Volume 12 Issue 6 June 2021
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)

Editorial Board

Editor-in-Chief

Dr. Kohei Arai - Saga University

Domains of Research: Technology Trends, Computer Vision, Decision Making, Information Retrieval, Networking, Simulation

Associate Editors

Alaa Sheta

Southern Connecticut State University

Domain of Research: Artificial Neural Networks, Computer Vision, Image Processing, Neural Networks, Neuro-Fuzzy Systems

Domenico Ciuonzo

University of Naples, Federico II, Italy

Domain of Research: Artificial Intelligence, Communication, Security, Big Data, Cloud Computing, Computer Networks, Internet of Things

Doroła Kaminska

Lodz University of Technology

Domain of Research: Artificial Intelligence, Virtual Reality

Elena Scutelnicu

"Dunarea de Jos" University of Galati

Domain of Research: e-Learning, e-Learning Tools, Simulation

In Soo Lee

Kyungpook National University

Domain of Research: Intelligent Systems, Artificial Neural Networks, Computational Intelligence, Neural Networks, Perception and Learning

Krassen Stefanov

Professor at Sofia University St. Kliment Ohridski

Domain of Research: e-Learning, Agents and Multi-agent Systems, Artificial Intelligence, e-Learning Tools, Educational Systems Design

Renato De Leone

Università di Camerino

Domain of Research: Mathematical Programming, Large-Scale Parallel Optimization, Transportation problems, Classification problems, Linear and Integer Programming

Xiao-Zhi Gao

University of Eastern Finland

Domain of Research: Artificial Intelligence, Genetic Algorithms

CONTENTS

Paper 1: Real-Time Driver's Focus of Attention Extraction and Prediction using Deep Learning

Authors: Pei-heng Hong, Yuehua Wang

PAGE 1 – 10

Paper 2: Predicting Strength Ratio of Laminated Composite Material with Evolutionary Artificial Neural Network

Authors: Huiyao Zhang, Atsushi Yokoyama

PAGE 11 – 18

Paper 3: UIP2SOP: A Unique IoT Network applying Single Sign-On and Message Queue Protocol

Authors: Lam Nguyen Tran Thanh, Nguyen Ngoc Phien, The Anh Nguyen, Hong Khanh Vo, Hoang Huong Luong, Tuan Dao Anh, Khoi Nguyen Huynh Tuan, Ha Xuan Son

PAGE 19 – 30

Paper 4: A Hybrid Encryption Solution to Improve Cloud Computing Security using Symmetric and Asymmetric Cryptography Algorithms

Authors: Hossein Abroshan

PAGE 31 – 37

Paper 5: Evil Twin Attack Mitigation Techniques in 802.11 Networks

Authors: Raja Muthalagu, Sachin Sanjay

PAGE 38 – 41

Paper 6: Solving Reactive Power Scheduling Problem using Multi-objective Crow Search Algorithm

Authors: Surrender Reddy Salkuti

PAGE 42 – 48

Paper 7: Controlling a Wheelchair using a Brain Computer Interface based on User Controlled Eye Blinks

Authors: Sebastián Poveda Zavala, Sang Guun Yoo, David Edmigio Valdivieso Tituana

PAGE 49 – 55

Paper 8: IoT based Smart Water Quality Prediction for Biofloc Aquaculture

Authors: Md. Mamunur Rashid, Al-Akhir Nayan, Sabrina Afrin Simi, Joyeta Saha, Md. Obaidur Rahman, Muhammad Golam Kibria

PAGE 56 – 62

Paper 9: A Novel High-order Linguistic Time Series Forecasting Model with the Growth of Declared Word-set

Authors: Nguyen Duy Hieu, Pham Dinh Phong

PAGE 63 – 71

Paper 10: Opinion Mining of Saudi Responses to COVID-19 Vaccines on Twitter

Authors: Fahad M. Alliheibi, Abdulfattah Omar, Nasser Al-Horais

PAGE 72 – 78

Paper 11: Indoor Localization and Navigation based on Deep Learning using a Monocular Visual System

Authors: Rodrigo Eduardo Arevalo Ancona, Leonel Germán Corona Ramírez, Oscar Octavio Gutiérrez Frías

PAGE 79 – 86

Paper 12: Firm Performance Prediction for Macroeconomic Diffusion Index using Machine Learning

Authors: Cu Nguyen Giap, Dao The Son, Dinh Thi Ha, Vu Quang Huy, Do Thi Thu Hien, Le Mai Trang

PAGE 87 – 99

Paper 13: A Review on SDR, Spectrum Sensing, and CR-based IoT in Cognitive Radio Networks

Authors: Nadia Kassri, Abdeslam Ennouaary, Slimane Bah, Hajar Baghdadi

PAGE 100 – 121

Paper 14: Artificial Intelligence based Recommendation System for Analyzing Social Business Reviews

Authors: Asma Alanazi, Marwan Alseid

PAGE 122 – 132

Paper 15: Comprehensive Analysis for Sensor-Based Hydraulic System Condition Monitoring

Authors: Ahmed Alenany, Ahmed M. Helmi, Basheer M. Nasef

PAGE 133 – 140

Paper 16: A New Communication Protocol for Drones Cooperative Network: 5G Site Survey Case Study

Authors: Youssef Shawky Othman, Mohamed Helmy Megahed, Mohamed Abo Rezka, Fathy Ahmed Elsayed Amer

PAGE 141 – 152

Paper 17: Real Time Face Expression Recognition along with Balanced FER2013 Dataset using CycleGAN

Authors: Fatma Mazen Ali Mazen, Ahmed Aly Nashat, Rania Ahmed Abdel Azeem Abul Seoud

PAGE 153 – 164

Paper 18: A Deep Learning Approach Combining CNN and Bi-LSTM with SVM Classifier for Arabic Sentiment Analysis

Authors: Omar Alharbi

PAGE 165 – 172

Paper 19: Emotional Intelligence Robotics to Motivate Interaction in E-Learning: An Algorithm

Authors: Dalia khairy, Salem Alkhalaf, M. F. Areed, Mohamed A. Amasha, Rania A. Abougalala

PAGE 173 – 183

Paper 20: Building Research Productivity Framework in Higher Education Institution

Authors: Ahmad Sanmorino, Ermatita, Samsuryadi, Dian Palupi Rini

PAGE 184 – 191

Paper 21: Method for Determination of Support Length of Daubechies Basis Function for Wavelet MRA based Moving Characteristic Estimation

Authors: Kohei Arai

PAGE 192 – 200

Paper 22: DeepfakeNet, an Efficient Deepfake Detection Method

Authors: Dafeng Gong, Yogan Jaya Kumar, Ong Sing Goh, Zi Ye, Wanle Chi

PAGE 201 – 207

Paper 23: A New Key Generation Technique based on Neural Networks for Lightweight Block Ciphers

Authors: Sohel Rana, M. Rubaiyat Hossain Mondal, A. H. M. Shahariar Parvez

PAGE 208 – 216

Paper 24: Fully Automated Ontology Increment`S user Guide Generation

Authors: Kaneeka Vidanage, Rosmayati Mohemad, Noor Maizura Mohamad Noor, Zuriana Abu Bakar

PAGE 217 – 227

Paper 25: Design of an Efficient RPL Objective Function for Internet of Things Applications

Authors: Sonia Kuwelkar, H.G. Virani

PAGE 228 – 235

Paper 26: Enhance Risks Management of Software Development Projects in Concurrent Multi-Projects Environment to Optimize Resources Allocation Decisions

Authors: Ibraheem M Alharbi, Adel A Alyoubi, Majid Altuwairiqi, Mahmoud Abd Ellatif

PAGE 236 – 246

Paper 27: Fog-based Remote in-Home Health Monitoring Framework

Authors: Fatma H. Elgendy, Amany M. Sarhan, Mahmoud A. M. Alshewimy

PAGE 247 – 254

Paper 28: What Drives Airbnb Customers' Satisfaction in Amsterdam? A Sentiment Analysis

Authors: Heyam Abdullah Bin Madhi, Muna M. Alhammah

PAGE 255 – 263

Paper 29: Motivational Factors Impacting the Use of Citizen Reporting Applications in Saudi Arabia: The Case of Balagh Application

Authors: Muna M. Alhammah, Layla Hajar, Sahar Alshathry, Mashael Alqasabi

PAGE 264 – 272

Paper 30: Securing Student Data Privacy using Modified Snake and Ladder Cryptographic Algorithm

Authors: Kamaladevi Kunkolienker, Vaishnavi Kamat

PAGE 273 – 279

Paper 31: A Method to Accommodate Backward Compatibility on the Learning Application-based Transliteration to the Balinese Script

Authors: Gede Indrawan, I Ketut Paramarta, I Gede Nurhayata, Sariyasa

PAGE 280 – 286

Paper 32: Validation: Conceptual versus Activity Diagram Approaches

Authors: Sabah Al-Fedaghi

PAGE 287 – 297

Paper 33: Identifying Small and Medium Enterprise Smart Entrepreneurship Training Framework Components using Thematic Analysis and Expert Review

Authors: Anis Nur Assila Rozmi, Puteri N.E. Nohuddin, Abdul Razak Abdul Hadi, Mohd Izhar A. Bakar

PAGE 298 – 309

Paper 34: A Framework for Protecting Teenagers from Cyber Crimes and Cyberbullying

Authors: Sultan Saud Alanazi, Adwan Alownie Alanazi

PAGE 310 – 315

Paper 35: Anonymity Feature in Android Mobile Apps for Interest Groups

Authors: Lee Sin Yi, Nurul A. Emran, Norharyati Harum

PAGE 316 – 323

Paper 36: A Method to Prevent SQL Injection Attack using an Improved Parameterized Stored Procedure

Authors: Kamsuriah Ahmad, Mayshara Karim

PAGE 324 – 332

Paper 37: From User Stories to UML Diagrams Driven by Ontological and Production Model

Authors: Samia Nasiri, Yassine Rhazali, Mohammed Lahmer, Amina Adadi

PAGE 333 – 340

Paper 38: Law Architecture for Regulatory-Compliant Public Enterprise Model: A Focus on Healthcare Reform in Egypt

Authors: Alsayed Abdelwahed Mohamed, Nashwa El-bendary, A. Abdo

PAGE 341 – 353

Paper 39: Microorganisms: Integrating Augmented Reality and Gamification in a Learning Tool

Authors: Ratna Zuarni Ramli, Nor Athirah Umairah Marobi, Noraidah Sahari Ashaari

PAGE 354 – 359

Paper 40: Emotional Evocative User Interface Design for Lifestyle Intervention in Non-communicable Diseases using Kansei

Authors: Noor Afiza Mat Razali, Normaizeerah Mohd Noor, Norulzahrah Mohd Zainudin, Nur Atiqah Malizan, Nor Asiakin Hasbullah, Khairul Khalil Ishak

PAGE 360 – 368

Paper 41: Exploring the Socio-economic Implications of Artificial Intelligence from Higher Education Student's Perspective

Authors: Sarah Areef, Lobna Amouri, Nahla El-Haggar, Aishah Moneer

PAGE 369 – 376

Paper 42: Feistel Network Assisted Dynamic Keying based SPN Lightweight Encryption for IoT Security

Authors: Krishna Priya Gurumanapalli, Nagendra Muthuluru

PAGE 377 – 392

Paper 43: Cloud-based Secure Healthcare Framework by using Enhanced Ciphertext Policy Attribute-Based Encryption Scheme

Authors: Siti Dhalila Mohd Satar, Mohamad Afendee Mohamed, Masnida Hussin, Zurina Mohd Hanapi, Siti Dhalila Mohd Satar

PAGE 393 – 399

Paper 44: Determining Optimal Number of K for e-Learning Groups Clustered using K-Medoid

Authors: S. Anthony Philomen Raj, Vidyaathulasiraman

PAGE 400 – 407

Paper 45: Abnormal Pulmonary Sounds Classification Algorithm using Convolutional Networks

Authors: Alva Mantari Alicia, Arancibia-Garcia Alexander, Chávez Frías William, Cieza-Terrones Michael, Herrera-Arana Víctor, Ramos-Cosi Sebastian

PAGE 408 – 415

Paper 46: Cost Effective Hybrid Fault Tolerant Scheduling Model for Cloud Computing Environment

Authors: Annabathula.Phani Sheetal, K. Ravindranath

PAGE 416 – 422

Paper 47: An Efficient Image Clustering Technique based on Fuzzy C-means and Cuckoo Search Algorithm

Authors: Lahbib KHRISSI, Nabil EL AKKAD, Hassan SATORI, Khalid SATORI

PAGE 423 – 432

Paper 48: Prediction of Cantilever Retaining Wall Stability using Optimal Kernel Function of Support Vector Machine

Authors: Rohaya Alias, Siti Jahara Matlan, Aniza Ibrahim

PAGE 433 – 438

Paper 49: Analyzing the Performance of Anomaly Detection Algorithms

Authors: Chiranjit Das, Akhtar Rasool, Aditya Dubey, Nilay Khare

PAGE 439 – 445

Paper 50: Current Perspective of Symbiotic Organisms Search Technique in Cloud Computing Environment: A Review

Authors: Ajoze Abdulraheem Zubair, Shukor Bin Abd Razak, Md. Asri Bin Ngadi, Aliyu Ahmed

PAGE 446 – 453

Paper 51: Sarcasm Detection in Tweets: A Feature-based Approach using Supervised Machine Learning Models

Authors: Arifur Rahaman, Ratnadip Kuri, Syful Islam, Md. Javed Hossain, Mohammed Humayun Kabir

PAGE 454 – 460

Paper 52: Evaluation of Agent-Network Environment Mapping on Open-AI Gym for Q-Routing Algorithm

Authors: Varshini Vidyadhar, R. Nagaraja

PAGE 461 – 469

Paper 53: Internet of Things (IoT) Based ECG System for Rural Health Care

Authors: Md. Obaidur Rahman, Mohammad Abul Kashem, Al-Akhir Nayan, Most. Fahmida Akter, Fazly Rabbi, Marzia Ahmed, Mohammad Asaduzzaman

PAGE 470 – 477

Paper 54: Design of a Plastic Shredding Machine to Obtain Small Plastic Waste

Authors: Witman Alvarado-Diaz, Jason Chicoma-Moreno, Brian Meneses-Claudio, Luis Nuñez-Tapia

PAGE 478 – 483

Paper 55: An HC-CSO Algorithm for Workflow Scheduling in Heterogeneous Cloud Computing System

Authors: Jai Bhagwan, Sanjeev Kumar

PAGE 484 – 492

Paper 56: Improving Imbalanced Data Classification in Auto Insurance by the Data Level Approaches

Authors: Mohamed Hanafy, Ruixing Ming

PAGE 493 – 499

Paper 57: A New Feature Filtering Approach by Integrating IG and T-Test Evaluation Metrics for Text Classification

Authors: Abubakar Adu, Mustafa Mat Deris, Noor Azah Samsudin, Aliyu Ahmed

PAGE 500 – 510

Paper 58: An Ontological Framework for Healthcare Web Applications Security

Authors: Mamdouh Alenezi

PAGE 511 – 516

Paper 59: A Proposal to Improve the Bit Plane Steganography based on the Complexity Calculation Technique

Authors: Cho Do Xuan

PAGE 517 – 522

Paper 60: Multiplicative Iterative Nonlinear Constrained Coupled Non-negative Matrix Factorization (MINC-CNMF) for Hyperspectral and Multispectral Image Fusion

Authors: Priya K, Rajkumar K K

PAGE 523 – 530

Paper 61: Optimal Operation of Smart Distribution Networks using Gravitational Search Algorithm

Authors: Surender Reddy Salkuti

PAGE 531 – 538

Paper 62: Analyzing the Performance of Stroke Prediction using ML Classification Algorithms

Authors: Gangavarapu Sailasya, Gorli L Aruna Kumari

PAGE 539 – 545

Paper 63: A Novel Threshold based Method for Vessel Intensity Detection and Extraction from Retinal Images

Authors: Farha Fatina Wahid, Sugandhi K, Raju G, Debabrata Swain, Biswaranjan Acharya, Manas Ranjan Pradhan

PAGE 546 – 554

Paper 64: Integration of Identity Governance and Management Framework within Universities for Privileged Users

Authors: Shadma Parveen, Sultan Ahmad, Mohammad Ahmar Khan

PAGE 555 – 562

Paper 65: Dorsal Hand Vein Extraction in Uncontrolled Environment

Authors: Nisha Charaya, Anil Kumar, Priti Singh

PAGE 563 – 568

Paper 66: An Ensemble GRU Approach for Wind Speed Forecasting with Data Augmentation

Authors: Anibal Flores, Hugo Tito-Chura, Victor Yana-Mamani

PAGE 569 – 574

Paper 67: Relative Merits of Data Mining Algorithms of Chronic Kidney Diseases

Authors: Harsha Herle, Padmaja K V

PAGE 575 – 583

Paper 68: Comparing the Balanced Accuracy of Deep Neural Network and Machine Learning for Predicting the Depressive Disorder of Multicultural Youth

Authors: Haewon Byeon

PAGE 584 – 588

Paper 69: Privacy Preserving Dynamic Provable Data Possession with Batch Update for Secure Cloud Storage

Authors: Smita Chaudhari, Gandharba Swain

PAGE 589 – 598

Paper 70: Classification Model Evaluation Metrics

Authors: Željko Đ. Vujović

PAGE 599 – 606

Paper 71: Cluster based Certificate Blocking Scheme using Improved False Accusation Algorithm

Authors: Chetan S Arage, K. V. V. Satyanarayana

PAGE 607 – 613

Paper 72: Augmented Reality based Adaptive and Collaborative Learning Methods for Improved Primary Education Towards Fourth Industrial Revolution (IR 4.0)

Authors: A F M Saifuddin Saif, Zainal Rasyid Mahayuddin, Azrulhizam Shapi'i

PAGE 614 – 623

Paper 73: Feasibility Study of a Small-Scale Grid-Connected PV Power Plants in Egypt; Case Study: New Valley Governorate

Authors: Mahmoud Saad, Hamdy M. Sultan, Ahmed Abdeltwab, Ahmed A. Zaki Diab

PAGE 624 – 631

Paper 74: Augmented Reality Adapted Book (AREmotion) Design as Emotional Expression Recognition Media for Children with Autistic Spectrum Disorders (ASD)

Authors: Tika Miningrum, Herman Tolle, Fitra A Bachtiar

PAGE 632 – 638

Paper 75: Towards Evaluating Adversarial Attacks Robustness in Wireless Communication

Authors: Asmaa FTAIMI, Tomader MAZRI

PAGE 639 – 646

Paper 76: An Approach for Optimal Feature Selection in Machine Learning using Global Sensitivity Analysis

Authors: G. Saranya, A. Pravin

PAGE 647 – 656

Paper 77: Energy Material Network Data Hubs

Authors: Robert R. White, Kristin Munch, Nicholas Wunder, Nalinrat Guba, Chitra Sivaraman, Kurt M. Van Allsburg, Huyen Dinh, Courtney Pailing

PAGE 657 – 667

Paper 78: A New Approach based on Fuzzy MADM for Enhancing Infrastructure Selection in the Case of VANET Network

Authors: Abdeslam Houari, Tomader Mazri

PAGE 668 – 675

Paper 79: A Modified Particle Swarm Optimization Approach for Latency of Wireless Sensor Networks

Authors: Jannat H. Elrefaei, Ahmed Yahya, Mouhamed K. Shaat, Ahmed H. Madian, Refaat M. Fikry

PAGE 676 – 685

Paper 80: Case Study of Self-Organization Processes in Information System Caching Components

Authors: Pavel Kurnikov, Nina Krapukhina

PAGE 686 – 692

Paper 81: Dynamic Management of Security Policies in PrivOrBAC

Authors: Jihane EL MOKHTARI, Anas ABOU EL KALAM, Siham BENHADDOU, Jean-Philippe LEROY

PAGE 693 – 701

Paper 82: Hyperspectral Image Classification using Convolutional Neural Networks

Authors: Shambulinga M, G. Sadashivappa

PAGE 702 – 708

Paper 83: Skeletonization of the Straight Leg Raise Movement using the Kinect SDK

Authors: Hustinawaty, T. Rumambi, M. Hermita

PAGE 709 – 713

Paper 84: MAGITS: A Mobile-based Information Sharing Framework for Integrating Intelligent Transport System in Agro-Goods e-Commerce in Developing Countries

Authors: Stivin Aloyce Nchimbi, Mussa Ally Dida, Gerrit K. Janssens, Janeth Marwa, Michael Kisangiri

PAGE 714 – 725

Paper 85: Software Project Estimation with Machine Learning

Authors: Noor Azura Zakaria, Amelia Ritahani Ismail, Afrujaan Yakath Ali, Nur Hidayah Mohd Khalid, Nadzurah Zainal Abidin

PAGE 726 – 734

Paper 86: Do Learning Styles Enhance the Academic Performance of University Students? A Case Study

Authors: Jorge Muñoz-Mederos, Elizabeth Acosta-Gonzaga, Elena Fabiola Ruiz-Ledesma, Aldo Ramírez-Arellano

PAGE 735 – 741

Paper 87: Application of Expert System with Smartphone-based Certainty Factor Method for Hypertension Risk Detection

Authors: Dodon Yendri, Derisma, Desta Yolanda, Sabrun Jamil

PAGE 742 – 749

Paper 88: Using API with Logistic Regression Model to Predict Hotel Reservation Cancellation by Detecting the Cancellation Factors

Authors: Sultan Almotiri, Nouf Alosaimi, Bayan Abdullah

PAGE 750 – 759

Paper 89: A Technique for Constrained Optimization of Cross-ply Laminates using a New Variant of Genetic Algorithm

Authors: Huiyao Zhang, Atsushi Yokoyama

PAGE 760 – 767

Paper 90: Kinematic Analysis for Trajectory Planning of Open-Source 4-DoF Robot Arm

Authors: Han Zhong Ting, Mohd Hairi Mohd Zaman, Mohd Faisal Ibrahim, Asraf Mohamed Moubark

PAGE 768 – 775

Paper 91: Fake News Detection in Arabic Tweets during the COVID-19 Pandemic

Authors: Ahmed Redha Mahlous, Ali Al-Laith

PAGE 776 – 785

Paper 92: Detecting COVID-19 Utilizing Probabilistic Graphical Models

Authors: Emad Alsuwat, Sabah Alzahrani, Hatim Alsuwat

PAGE 786 – 793

Paper 93: An Optimized Artificial Neural Network Model using Genetic Algorithm for Prediction of Traffic Emission Concentrations

Authors: Akibu Mahmoud Abdullah, Raja Sher Afgun Usmani, Thulasyammal Ramiah Pillai, Mohsen Marjani, Ibrahim Abaker Targio Hashem

PAGE 794 – 803

Paper 94: Discrete Time-Space Stochastic Mathematical Modelling for Quantitative Description of Software Imperfect Fault-Debugging with Change-Point

Authors: Mohd Taib Shatnawi, Omar Shatnawi

PAGE 804 – 811

Paper 95: The Role of Data Pre-processing Techniques in Improving Machine Learning Accuracy for Predicting Coronary Heart Disease

Authors: Osamah Sami, Yousef Elsheikh, Fadi Almasalha

PAGE 812 – 820

Paper 96: Numerical Investigation on System of Ordinary Differential Equations Absolute Time Inference with Mathematica®

Authors: Adeniji Adejimi, Surulere Samuel, Mkolesia Andrew, Shatalov Michael

PAGE 821 – 829

Paper 97: Maneuverable Autonomy of a Six-legged Walking Robot: Design and Implementation using Deep Neural Networks and Hexapod Locomotion

Authors: Hiep Xuan Huynh, Nghia Duong-Trung, Tran Nam Quoc Nguyen, Bao Hoai Le, Tam Hung Le

PAGE 830 – 839

Paper 98: Comprehensive Analysis of Augmented Reality Technology in Modern Healthcare System

Authors: Jinat Ara, Faria Benta Karim, Mohammed Saud A Alsubaie, Yeasin Arafat Bhuiyan, Muhammad Ismail Bhuiyan, Salma Begum Bhyan, Hanif Bhuiyan

PAGE 840 – 849

Paper 99: Identification of Abusive Behavior Towards Religious Beliefs and Practices on Social Media Platforms

Authors: Tanvir Ahammad, Md. Khabir Uddin, Tamanna Yesmin, Abdul Karim, Sajal Halder, Md. Mahmudul Hasan

PAGE 850 – 860

Paper 100: Fish Disease Detection System: A Case Study of Freshwater Fishes of Bangladesh

Authors: Juel Sikder, Kamrul Islam Sarek, Utpol Kanti Das

PAGE 861 – 865

Paper 101: Improving Data Services of Mobile Cloud Storage with Support for Large Data Objects using OpenStack Swift

Authors: Aslam B Nandyal, Mohammed Rafi, M Siddappa, Babu B. Sathish

PAGE 866 – 876

Paper 102: A Comparative Study of Stand-Alone and Hybrid CNN Models for COVID-19 Detection

Authors: Wedad Alawad, Banan Alburaidi, Asma Alzahrani, Fai Alflaj

PAGE 877 – 883

Paper 103: A Self-adaptive Algorithm for Solving Basis Pursuit Denoising Problem

Authors: Mengkai Zhu, Xu Zhang, Bing Xue, Hongchun Sun

PAGE 884 – 888

Paper 104: Wheelchair Control System based Eye Gaze

Authors: Samar Gamal Amer, Rabie A. Ramadan, Sanaa A. kamh, Marwa A. Elshahed

PAGE 889 – 894

Paper 105: Design and Implementation of a Most Secure Cryptographic Scheme for Lightweight Environment using Elliptic Curve and Trigonohash Technique

Authors: BhaskarPrakash Kosta, PasalaSanyasi Naidu

PAGE 895 – 912

Paper 106: Detecting Malware Infection on Infrastructure Hosted in IaaS Cloud using Cloud Visibility and Forensics

Authors: Lama Almadhour, A. A. Abd El-Aziz, Hedi Hamdi

PAGE 913 – 928

Paper 107: Reputation Measurement based on a Hybrid Sentiment Analysis Approach for Saudi Telecom Companies

Authors: Bayan Abdullah, Nouf Alosaimi, Sultan Almotiri

PAGE 929 – 937

Real-Time Driver's Focus of Attention Extraction and Prediction using Deep Learning

Pei-heng Hong¹, Yuehua Wang^{2*}

Department of Computer Science
Texas A&M University-Commerce
Commerce, Texas 75428 USA

*Corresponding Author

Abstract—Driving is one of the most common activities in our modern lives. Every day, millions drive to and from their schools or workplaces. Even though this activity seems simple and everyone knows how to drive on roads, it actually requires drivers' complete attention to keep their eyes on the road and surrounding cars for safe driving. However, most of the research focused on either keeping improving the configurations of active safety systems with high-cost components like Lidar, night vision cameras, and radar sensor array, or finding the optimal way of fusing and interpreting sensor information without considering the impact of drivers' continuous attention and focus. We notice that effective safety technologies and systems are greatly affected by drivers' attention and focus. In this paper, we design, implement and evaluate Dfaep, a deep learning network for automatically examining, estimating, and predicting driver's focus of attention in a real-time manner with dual low-cost dash cameras for driver-centric and car-centric views. Based on the raw stream data captured by the dash cameras during driving, we first detect the driver's face and eye and generate augmented face images to extract facial features and enable real-time head movement tracking. We then parse the driver's attention behaviors and gaze focus together with the road scene data captured by one front-facing dash camera faced towards the roads. Facial features, augmented face images, and gaze focus data are then inputted to our deep learning network for modeling drivers' driving and attention behaviors. Experiments are then conducted on the large dataset, DR(eye)VE, and our own dataset under realistic driving conditions. The findings of this study indicated that the distribution of driver's attention and focus is highly skewed. Results show that Dfaep can quickly detect and predict the driver's attention and focus, and the average accuracy of prediction is 99.38%. This will provide a basis and feasible solution with a computational learnt model for capturing and understanding driver's attention and focus to help avoid fatal collisions and eliminate the probability of potential unsafe driving behavior in the future.

Keywords—Driving; attention; interesting zones; deep neural network; deep learning; models

I. INTRODUCTION

While being a basic enabler of our modern society, road vehicle transportation has become a major source of societal concerns. In 2016 alone, over 330, 655 people [1]–[3] were killed in vehicle accidents in the world, which is far greater than the number of US Soldiers killed in action in Vietnam (58,220 fatalities). Comparing to 19 high-income countries [3] including Canada, Germany, France, Spain, United Kingdom, and other 14 countries, the United States has highest vehicle accident death rate, where about 90 people die each day

in vehicle accidents. The U.S. National Security Council (USNSC) reports that vehicle accidents cause estimated 40,200 fatalities, a 6% rise from 2015 (i.e., a 14% rise from 2014), which makes 2016 the deadliest year on American roads in nearly a decade. To prevent vehicle accidents, injuries, and deaths [4], we have enforced the use of seat belts, car seats and booster seats for children through at least age 8, restricted alcohol-impaired driving [5] and speeding, and suggested to avoid distracted driving (such as texting, talking on the phone, eating or doing something else that occupies the senses you need to drive). In reality, those are far from enough for safe driving.

A wide variety of approaches [6]–[8], systems [9]–[11], and self-driving cars [12]–[15] have been developed for vehicle surrounding environment perception, including object detection, tracking, localization, navigation. In the market, most of automotive companies like Tesla, Mercedes-benz, Audi, BMW, Rolls-Royce, GM, Ford, and Honda have offered diverse advanced driver-assistance systems (ADASs) with various high-cost components including Lidar, high resolution cameras, radar sensor array, sonar, GPS, odometry, and inertial measurement unit. ADASs are electronic systems that assist the human driver while driving or during parking and have significantly improved comfort and safety in nowadays driving. Waymo [13], [14] originated by Google's self-driving car, Navya automated bus driving system [16], Uber driverless car [17], and Apple car [15] are promising autonomous driving vehicle projects to promote fully personal self-driving vehicle and related technology development. Since there are still many technological difficulties and regulatory issues that are need to be addressed, it would be close to a decade before self-driving cars are ready to use safely in a large number [18].

The emergence of artificial intelligence and rapid development of computer vision technology provide more opportunities to improve driving safety without strongly relying on the high-cost components or external information surrounding vehicles. We observe that driver's attention and focus have significant impact on driving safety and can lead to vehicle accidents or even terrible disasters. It, therefore, is extremely important to investigate and understand driver's attention and focus taking into consideration the complexity and dynamics of driver's behaviors under realistic driving conditions. Unfortunately, this topic is under-investigated and the lack of realistic experimental system does not help.

Motivated by the fact that driver's attention and focus are always changing over time in an uncertain, dynamic, and

continuous environment, we propose to capture and examine driver attention and focus using two low-cost dash cameras in real-time. Instead of using expensive, commercial eye tracker/tracking devices, dual dash cameras are placed on the windshield inside the vehicle. One camera is facing front for car-centric view and the other is facing towards the driver for driver-centric view. Based on the raw streaming data from the dash cameras, we provide an approach to extract facial features with face and eye detection, augment face images, track and predicate gaze focus zones during driving. Our first contribution is to build our drive attention database by driving from/to our campus to/from Dallas, Texas. The underlying ideas is to travel different types of roads including campus road, rural road, suburban road, highways, and metropolitan roadways to collect enough driver-centric and car-centric data under various driving circumstances.

The second contribution is to extract the driver's facial features and generate augmented face images using open source toolkit, Openface [19], [20]. The main visual cues that we are analyzing are head and eye directions that are selected based on the insights that drivers who are paying attention on driving will have a tendency to look forward and keep eyes on roads. To enable to locate the driver's focus, the driver's view is grided into 11 grids within 5 gaze zones. The a feature vector can be formed with the facial features and gaze zone to label the data. The third contribution is the introduction of a data-driven deep learning network called Dfaep. This network uses the detected features and augmented face images as the input and output accurate the focus of driver's attention. If the focus is not within the main zones, a warning alert will be issued to bring it to the driver attention. We evaluate Dfaep on two datasets, DR(eye)VE [6] and our own dataset. On DR(eye)VE, we substantially improve the predictions of the driver's focus with accurate gaze data on the real road scene. On our dataset, we test the trained model and compare our Dfaep with other deep learning networks with the same set of images and extracted features in terms of accuracy, loss, and network complexity to identify the network that is best for real-time systems and applications. We outperform all other networks and show the results that Dfaep can achieve highest prediction accuracy, 99.38% with lowest validation loss, 0.018.

The paper is organized as follows: Section II provides a brief literature review about computer vision systems and artificial intelligence relate to driver attention behaviors; Section III details the design and structure of the proposed deep learning network Dfaep for modeling the driver attention behaviors and tracking focus; Section IV presents the datasets used in experiments. The numerical performance evaluation results and our finding are then reported; and Section V concludes the paper and describes our future research plan.

II. LITERATURE

In this section, we review the existing technologies and studies used to capture and detect driver's attention and focus, which can be classified based on eye/gaze tracking to capture humans attention and detection and prediction methods driven by sensing data.

A. Eye Tracking

Eye tracking data is widely held to be a good window into attention or non-attentive states for learners or people working on tasks [21], [22]. Eye tracking refers to the careful measurement of the movements of the eyes and/or body when the participants are either positively or negatively interacting with the learning environments in a time-varying manner. The measurement devices, hardware platforms, and systems are commonly known as eye trackers. Based on the hardware setup, we classify the existing eye tracking systems into four main types: tower-mounted eye-tracker [23]–[25], screen-based eye tracker [26], [27], head mounted/wearable eye tracker [28]–[30], and mobile eye tracker [31]–[33]. An detailed discussion on those four types of eye tracking devices and platforms can be found in our papers [34]–[36]. Eye trackers can be either well-established commercial devices/systems or low-cost, portable systems designed by educators or researchers. For commercial eye trackers, they are still expensive in the current market, but well developed and maintained by companies like Tobii, SMI, ISCAN, LC Technologies, EyeTech, Oculus, Ergoneers, SmartEye, and Mirametrix [36]. With the requirements of necessary purchasing, only authorized users can use the purchased hardware and software with warranties, documentations, reachable technical support. Comparing to the commercial eye trackers, open source eye tracking devices and systems [29], [30], [37]–[39] have unique abilities to support both and low-cost eye trackers, easily alter experiments to specific scenarios, quickly prototype ideas and enable eye tracking research and applications without major restrictions. Given that, in this study, we propose an active safe driving behavior detection system with two low-cost car dash cameras, ZEdge-Z3 in \$100.

B. Attention Detection and Prediction

There are existing studies have been conducted with the eye tracking or camera data to detect and predicate driver's attention. For instance, Tran et al. [40] proposed an assisted-driving testbed with a driving simulator in a laboratory to simulate the driving environment and control the simulator's behavior using a script. Ten distracted driving behaviors like drinking, operation the radio, talking, texting, reaching behind, and making-up were defined. Based on the definition, four deep learning models, CNN, VGG-16, AlexNet, and GoogleNet are trained and tested on the simulator for driving distraction detection by distinguishing the defined distracted behaviors. The results showed that the VGG-16 achieved fastest frequency (14 Hz) while the ResNet model yielded highest accuracy of 92% with highest complexity which needed to have longest time for data processing and model calculation.

In [41], a CNN(Convolution Neural Network) is proposed and trained to mimic the driver based on training data from human driver's driving, it builds a model which takes as an input raw data and map it directly to a decision, using minimum training data and minimum computation, the car learns to drive on roads with or without human-design features. The idea of end to end self-driving car were implemented by NVIDIA [42]. They trained CNN to map raw pixels from one single front camera directly to steering commands. The network architecture is shown in Fig. 2. that consists nine layers which includes five convolutional layers, one normalization

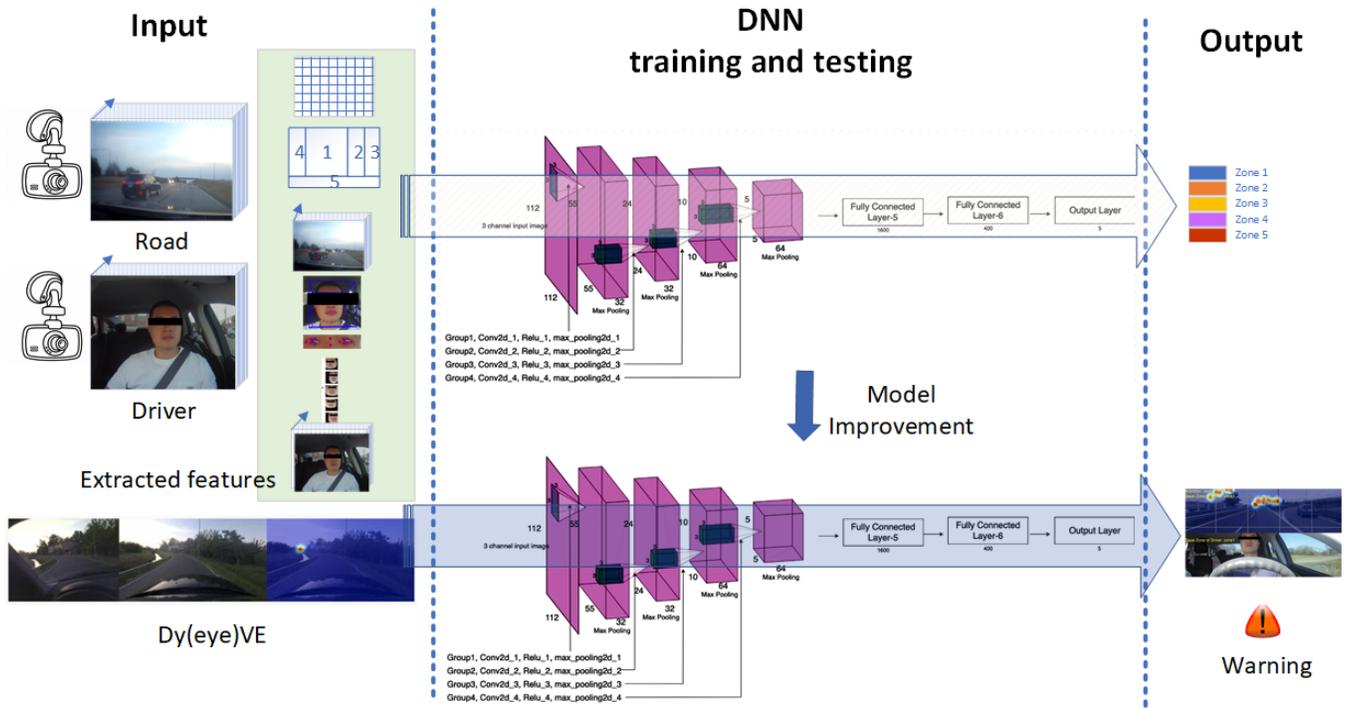
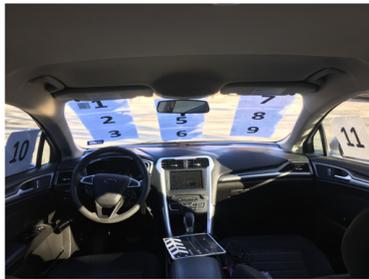


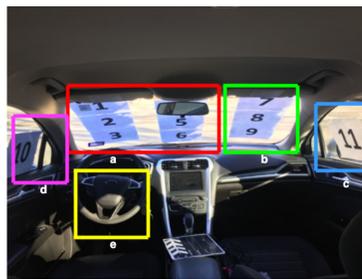
Fig. 1. Flowchart of Driver Attention System.



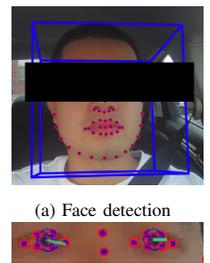
(a) 11 Grids of windshield



(b) Internal dual cameras



(c) 5 Main gaze zones



(a) Face detection
(b) Eye detection

Fig. 2. Internal Setting of our Experimental Vehicle.

Fig. 3. Driver Facial Detection.

layer and three fully connected layers, they used NVIDIA DevBox, Torch 7 and NVIDIA DRIVETM PX self-driving car computer for training and determining the steering angle. However, in their work, they did not consider about the throttle factor into training, they only focus on steering angle control, Therefore, based on their work, adding speed throttle factor and implementing it with RC car, can pose new challenges.

In this study, we compare the internal and external environment of the vehicle and propose a method that leverages deep learning network to predict the driver's attention in the environment inside and outside the vehicle. For the driver warning part, we have added scenes outside the vehicle as warning auxiliary factors to determine whether the driver is distracted by the driving of the vehicle.

III. PROPOSED SYSTEM AND METHODOLOGY

A. Overview of Our Driver Attention Database

The overview of this research is shown in Fig. 1. We use a camera to capture the raw images of the the driver's frontal view based on the corresponding windshield inside the vehicle, each zone for 1 minute. Then, the driver's facial features in raw images are extracted through the OpenFace [19], [20]. After that, each image of driver facial feature is cropped to 112×112 pixels for input DNN for training. We divide the data into training set, validation set and testing set. Here, our driver attention system will classify the driver's attention zone based on probability score of DNN model for each zone on the windshield. In DR(eye)VE database [6], we leverage saliency maps in the DR(eye)VE database to simulate our driving scene and label each frame of images, then we perform a synchronous comparison between the estimation zone of our driver attention detection classification system and attention zone of the DR(eye)VE database for evaluating the method

that we proposed. Finally, the synchronous comparison result is obtained, the system will give a warning to driver if the final label comparison result is not matched.

B. Detecting Facial Landmarks

The experimental environment with 11 grids within five gaze zones on the windshield of internal vehicle in this research is shown in Fig. 2. The reason we divided it into five gaze zones is to obtain more realistic driving scenarios. In this research we leverage a standard web camera to accomplish this task. A camera sensor is mounted in front of the driver. Since the size of the web camera is moderate (8 cm × 5 cm × 1.5 cm), it could be mounted in the vicinity of the windshield, and can continuously capture the driver's facial. Power to the web camera was supplied by the car using a universal serial bus (USB) line, connecting to the camera. The captured images by web camera is successively transmitted to the computer with graphics processing unit (GPU) for training via memory card of the web camera. The configuration of the web camera and the gaze zone on the windshield are shown in Fig. 2. We simulate driver's driving attention to different zone of windshield and their driving habits as if driving in real world, and save the videos recorded by the web camera, for each videos save the correspondent zone area performed by the driver. The dataset consists of 9,000 frames divided in five gaze zones, each of which is 1 minutes long, as shown in Fig. 4. The frames resolution size is 1920 × 1080 pixels.

C. Driver's Data Pre-processing

In our research, in order to improve the training efficiency and reduce unnecessary interference factors, our measures are to capture the driver's facial information and reduce the size of the original input image. In here, we leverage Openface to implement this goal. Openface was proposed by Tadas Baltrušaitis et al. [19], [20]. There are a lot of tools that can implement the face feature detection in images or videos. However, most of them did not provide the source code which makes it very difficult to reproduce experiments on different datasets. In Openface, it includes facial landmarks, head pose estimation, eye gaze estimation and the most importantly it is opens source, free and a tool that provides source code.

In our experiment, the facial feature of driver was captured based on Openface [19], [20]. Openface is a tool designed for computer vision and machine learning researchers, it leveraged Multi-task Convolutional Neural Network (MTCNN) facial feature tracker [43] to capture 68 face landmarks, eyelids, the iris, eyelids and the pupil are detected by a Constrained Local Neural Field (CLNF) landmark detector [44], and the head pose is extracted through 3D representation of facial landmarks of Convolutional Experts Constrained Local Model (CE-CLM) [45]. In the end, we obtained the appearance extraction face alignment images and cropped to 112×112 pixels for input to DFae as shown in Fig. 5.

D. Network Structure

In this research, we propose a deep learning network for driver's focus of attention extraction and prediction (DFae)(including six layers) as shown in Table I. The model used in this research is a simple stack of of four convolutional

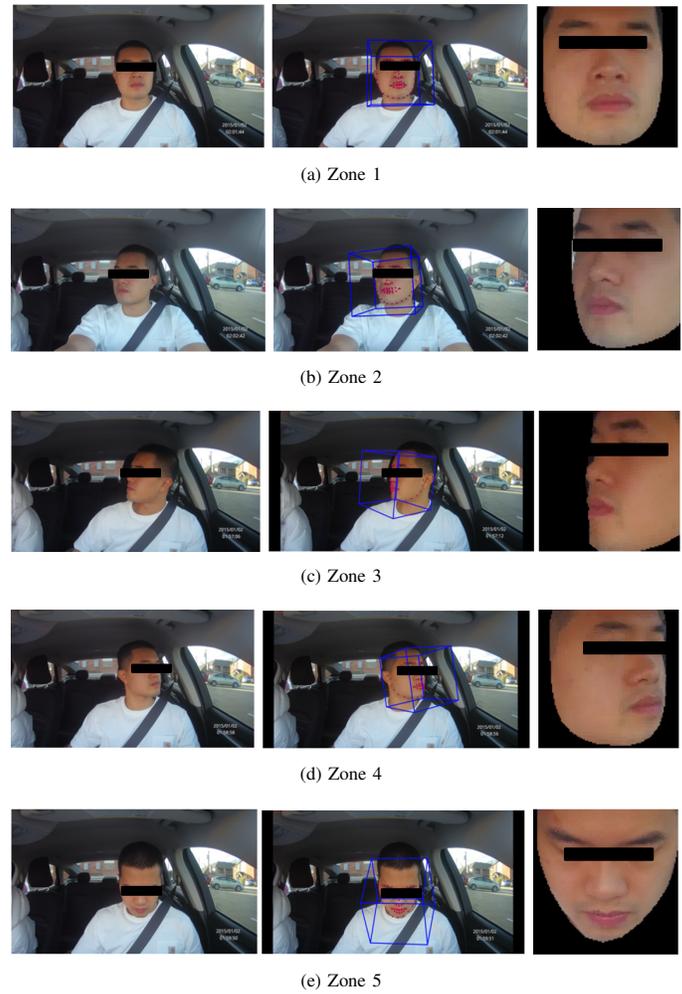


Fig. 4. Example Images of Driver's Face Landmark (left), after OpenFace Capture Driver's Attention (middle), and Driver's Facial Feature Image (right) while Looking at Distinct Zones of Fig. 2. Cases of Looking at (a) Zone 1; (b) Zone 2; (c) Zone 3; (d) Zone 4; (e) Zone 5.

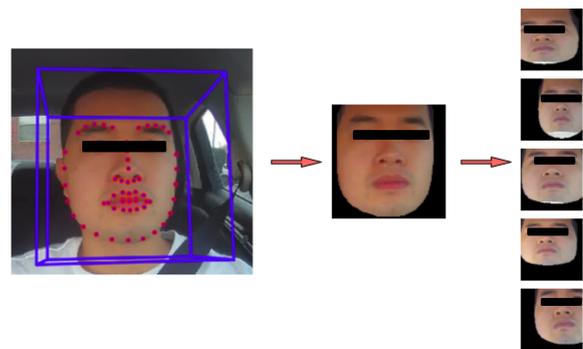


Fig. 5. The Example of Augmented Images.

layers (with ReLU activation [46]), four max pooling layers (with ReLU activation) and two fully connected layers (with ReLU and Softmax [47] activation). This model is modified from the LeNet [48]. We try to develop a DNN model with high accuracy and low parameter computation. Today we have far better models than this model, but this is the lightest one in terms of computation, and it is also the most suitable one for

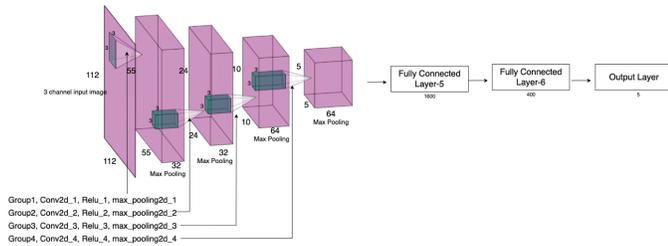


Fig. 6. Network Architecture used for Extracting Driver Gaze Features and Predicating Attention Focus.

TABLE I. STRUCTURE OF DFAEP MODEL

	Layers	# Filters	Size	Kernel	Stride	Padding
	Input layer	-	112 × 112 × 3	-	-	-
Group 1	1st Conv	32	110 × 110 × 32	3 × 3	1 × 1	0 × 0
	ReLU-1	-	110 × 110 × 32	-	-	-
	Pool-1	1	55 × 55 × 32	2 × 2	2 × 2	0 × 0
Group 2	2nd Conv	32	53 × 53 × 32	3 × 3	1 × 1	0 × 0
	ReLU-2	-	53 × 53 × 32	-	-	-
	Pool-2	1	26 × 26 × 32	2 × 2	2 × 2	0 × 0
Group 3	3rd Conv	64	24 × 24 × 64	3 × 3	1 × 1	0 × 0
	ReLU-3	-	24 × 24 × 64	-	-	-
	Pool-3	1	12 × 12 × 64	2 × 2	2 × 2	0 × 0
Group 4	4th Conv	64	10 × 10 × 64	3 × 3	1 × 1	0 × 0
	ReLU-4	-	10 × 10 × 64	-	-	-
	Pool-4	1	5 × 5 × 64	2 × 2	2 × 2	0 × 0
	1st FC	-	400 × 1	-	-	-
	ReLU-5	-	400 × 1	-	-	-
	Output layer	-	400 × 1	-	-	-
	2nd FC	-	400 × 1	-	-	-
	Softmax	-	400 × 1	-	-	-
	Output layer	-	400 × 1	-	-	-

our research. The performance comparison of different DNN models will be discussed in Section 4. In this model of DFAep, we executed the data augmentation in our training dataset. After extracting five features of last fully connected layer, the final attention zone of driver was estimated based on softmax function [47] of our DNN model.

The structure of our DFAep that is presented in Fig. 6 and will be illustrated in Table I. Our DFAep model is constructed of four convolutional layers (with ReLU activation [46]), 4 max pooling layers (with ReLU activation) and two fully connected layers. In the first convolutional layer, there are 32 filters of size 3 × 3 are applied to the input of 112 × 112 × 3. Here, 112 × 112 × 3 stands for width, height, and channel number, separately. Thus, a feature map of 112 × 112 × 32 is acquired. It can be computed according to the following standard: (output height = (input height - filter height + 2 × number of padding) / number of stride + 1) [49]. As shown in Table I, the height of input, height of filter, number of padding, and number of stride are 112, 3, 0, and 1 separately. From that the height of output of the feature map was obtained by the convolution is computed as 110 = ((112 - 3 + 2 × 0) / 1 + 1). If the value is positive, it can be valid output used in the next layer. If the value is negative, the output is 0.

E. Activation

The Rectified linear unit (ReLU) activation function [46].

$$ReLU = \max(0, x) \quad (1)$$

The ReLU function is not differentiable across the entire interval, but the non-differentiable part can be performed using Sub-gradient. Instead, ReLU is the most frequently used excitation function in recent years. Because of its following characteristics, including: solving the problem of gradient explosions, calculating quite quickly, and converging quickly, it will be analyzed in detail below.

For neural networks such as error back transfer operation (BN), gradient calculation considerations are most important when updating weights. Sigmoid function [50] is prone to vanishing gradient problems. When the input value approaches the saturation region (sigmoid function). When the excitation is performed outside [-4, +4], the first-order differential value approaches 0, and the problem of gradient disappearance occurs, which makes the backward transfer of the error calculation impossible, and the weight update cannot be performed effectively, and the neural network layer is deepened. The time is more obvious. Therefore, it is a difficulty encountered in deep neural network training, and the piecewise linear nature of ReLU can effectively overcome the problem of gradient disappearance. In our research, each convolutional layer is followed by one ReLU activation and one max pooling layer in Table I. The size of filter and stride within each max pooling layer is 2 × 2 and 1 × 1, respectively. As shown in Fig. 6, the size of each feature map is reduced by a pooling layer, ReLU-1 (110 × 110 × 32) is decreased to Pool-1 (55 × 55 × 32), ReLU-2 (55 × 55 × 32) to Pool-2 (24 × 24 × 32), ReLU-3 (24 × 24 × 32) to Pool-3 (10 × 10 × 64) and ReLU-4 (10 × 10 × 64) to Pool-4 (5 × 5 × 64).

After the four convolutional layers (with ReLU activation [46]), and 4 max pooling layers (with ReLU activation), we obtained the final feature map size of 5 × 5 × 64 pixels. And final feature map would be processed by the additional two fully connected layers. For each fully connected layers, the feature maps are separately of 400 × 1 and 5 × 1 as shown in Table I. In our driver attention classification system, the driver's attention zone is classified by . And the gaze zones of the vehicle are five as shown in Fig.6, hence after the softmax function' [47] in the last fully connected layer, we will determine which classification among these five is our final result according to the maximum probability calculated by the softmax function. In the last fully connected layer, the softmax activation is adopted as presented in Equation (1). Here, V_i is the output of the classifier's previous output unit; i is the category index; The total number of categories is C . S_i stands for the ratio of the index of the current element to the sum of the indices of all elements. (1) From the observation during the deep learning process that the accuracy of training set is very high, but the prediction accuracy is extremely low of validation set. This is due to the over-fitting issue. The over-fitting means that the learning is performed too thoroughly, and all the features of the training set have been learned, so the machine has learned too many local features, and too many fake features due to noise, which caused the model to fail.

In this research, we leverage data augmentation and dropout methods [51] for preventing the over-fitting problem. For data augmentation methods, the details of it will be discussed. In the dropout methods, we decide to dropout probability of 50 per cent after the the 1st fully connected layer. After the four convolutional layers (with ReLU activation),

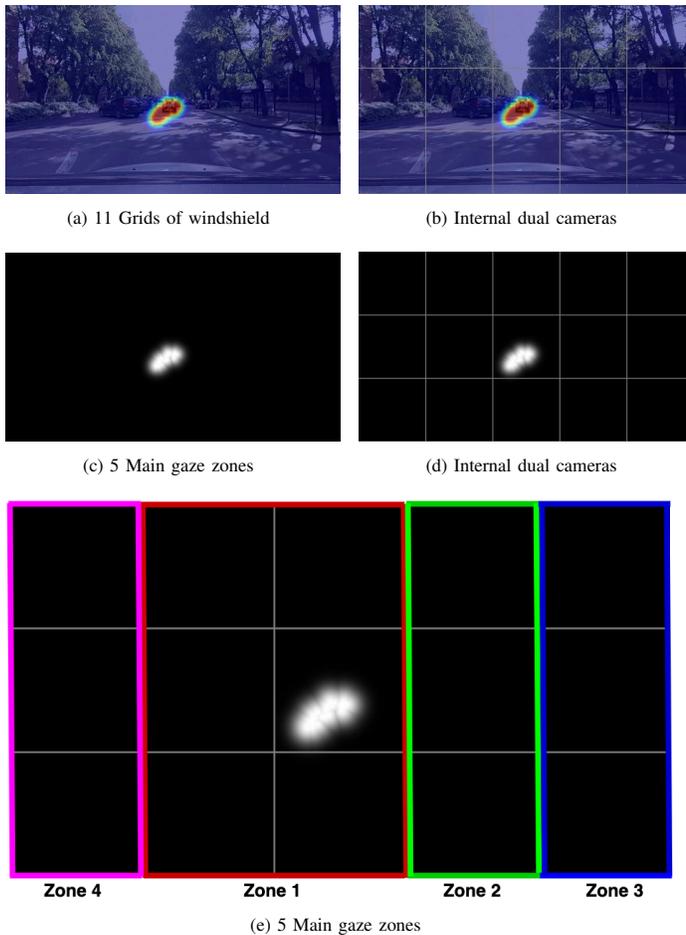


Fig. 7. Example Images of Driver's Face Landmark (left), after OpenFace Capture Driver's Attention (middle), and Driver's Facial Feature Image (right) while Looking at Distinct Zones of Fig. 2. Cases of Looking at (a) Zone 1; (b) Zone 2; (c) Zone 3; (d) Zone 4; (e) Zone 5.

and four max pooling layers (with ReLU activation), and 2nd fully connected layers (with ReLU and Softmax activation), the estimation zone of driver attention is obtained.

F. Model Improvement

We have used the DNN model to estimate the driver's gaze zone on the windshield of the vehicle. When the driver does not pay attention on the area where they should be focus, we should give a warning. We use the DR(eye)VE dataset to improve our model based on the attention zones of a real road driving scene. The DR(eye)VE dataset [6], it contains six hours of driving data, including 555,000 images, 74 video sequences of five minutes each. Videos were recorded in various environments, such as highway, downtown and countryside, sunny, rainy, cloudy, daytime and night. Because the DR(eye)VE project is currently the largest dataset of driving scenes with driving attention zones are available, we leverage its saliency maps dataset to compare to the estimation zone of driver to ensure whether or not to make the alarm during the driving.

Here we are using sequence videos in the DR (eye) VE dataset to simulate real road driving scene. In this work, we

divide at a rate of 5 frames per second to the videos into 1800 frames. In our research, the saliency maps are divided into the 15 grids. In our case, the driver's face landmark on the windshield is defined as four zones. As shown in the Fig. 7, the red block is defined as Zone 1, the reason for the larger proportion of this block is because the driver's concentration will usually focus on the front part, so our proportion in this area will be greater than other blocks. The green part is Zone 2, the blue part is Zone 3, and the purple part is Zone 4.

For performing a matching verification action with our driver's face concentration, we need to label each saliency map. For obtaining the label of our saliency maps, the filter of the size of 1×1 (stride number is 1) is used for the input of 1920×1080 pixels. We divide the 1920 by 5, then obtain a number 384 represents for each zone and then scan each pixel in each zone. The pixel of width from 1 to 384 is defined as Zone 4 (purple part of Fig. 7 (c)), 385 to 1152 is defined as Zone 1 (red part of Fig. 7 (c)), 1153 to 1536 is defined as Zone 2 (green part of Fig. 7 (c)), and 1537 to 1920 is defined as Zone 3 (blue part of Fig. 7 (c)). Since there are only pure black (0) and white (255) in the saliency map, thus the value of each pixel will be much unsophisticated. Finally, the attention zone whose final sum value is largest within 5 zones of Fig. 7 (c) is defined as the label of this saliency map.

G. Validation of Systems and Algorithms

With the road scene, we have labeled the attention area for each frame by calculating each pixel of frame. Then, we can match the prediction gaze zone of our current driver's facial image of Fig. 3 and the attention zone label of road scenarios of Fig. 7(c). After that, the driver prediction zone and road attention zone were determined whether or not identical based on label comparison. Finally, the label comparison result is obtained, the system will give a warning to driver if the final label comparison result is not matched.

IV. EXPERIMENTS AND RESULTS

A. Experiment Data

In our experiment, our dataset has been collected for the driver's attention classification system through a web camera as presented in Fig. 2. Our database was collected in an actual car instead of in a lab. There are many databases on driver concentration in previous research, such as Robust simultaneous modeling and tracking monitoring video (RS-DMV) dataset [52], The Chinese Face Database (CAS-PEAL) database [53] and Berkeley Deep Drive (BDD) [54], etc. However, these databases may not suitable for the research method we proposed this time due to the label of ground-truth attention zone is not defined. Without this key element, we cannot verify the method that purposed in this research. Therefore, we recorded our own database. In Fig. 4, 5 zones were split on the wind shield and designated to gaze at for this experiment. The size of raw image was $1920 \times 1080 \times 3$. When the experiment was starting, the driver has to perform normally, like they were actually driving in the ordinary day. After that, we cropped the size of images to 112×112 pixels of driver facial through OpenFace for Dfaep training and validation as presented in Fig. 3.

TABLE II. DESCRIPTION OF TRAINING, VALIDATION AND TESTING DATA IN OUR DATASET

Dataset	Amount
Training	7,000 (1,400 × 5) images for each gaze zone on the wind shield
Validation	2,010 (402 × 5) images for each gaze zone on the wind shield
Testing	1,802 images from the real driving scenarios

In our experiment, we performed three different DNN structures for training and validation. In here, we adopt AlexNet [55], VGG16 [56] and Dfaep model. The model structure is shown in Fig. 6. For preventing the over-fitting problem that we adopted some measures for data augmentation which are as follows. Shifting 2 pixels in vertically and horizontally and shear angle is 20 degrees in counter-clockwise direction as presented in Fig. 5. The raw images were used in validation process whereas the augmented images were used for training. In here, we used 7,000 augmented images in our dataset as training set, 2,010 images as validation set, and used completely independent 1800 images as testing set as shown in Table II. For the convolution neural network training, validation and testing, the desktop computer that we used is Intel i7-8700 CPU at 3.20 GHz and 16 GB memory. And the graphics card we used NVIDIA GeForce GTX 1070 (CUDA 10.0 and 8 GB memory). The backend was achieved by Keras (version 2.3.1) [57] with TensorFlow (version 2.1.0) [58]. And the algorithm was achieved by Visual Studio 2013, Dlib (version 19.19) and OpenCV (version 4.2.0) library.

B. Training of Dfaep Model

In this experiment, Adaptive Moment Estimation(Adam) [59] optimizer method was used in our Dfaep training. Adam is distinct to classical optimizer such as Stochastic gradient descent SGD [60], RMSprop [61] and Momentum [61], etc. Adam can be considered as a combination of RMSprop and Momentum, which uses the first and second moment estimates of the gradient to dynamically adjust the learning rate of each parameter. The main advantage of Adam is that after offset correction, the learning rate of each iteration has a certain range, making the parameters relatively stable as shown in Equations (2), (3) and (4).

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \end{aligned} \quad (2)$$

$$\begin{aligned} \hat{m}_t &= \frac{m_t}{1 - \beta_1^t} \\ \hat{v}_t &= \frac{v_t}{1 - \beta_2^t} \end{aligned} \quad (3)$$

$$W \leftarrow W - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \quad (4)$$

Where m_t and v_t are the weight to estimate of the first moment (the mean) and the second moment of the gradients respectively, g is the weight to gradient on current mini-batch, β_1 is used for decaying the running average of the gradient, β_2 is used for decaying the running average of the square of gradient, ε is the weight to prevent division from zero error. In our Dfaep training, the experiment was adopted for the

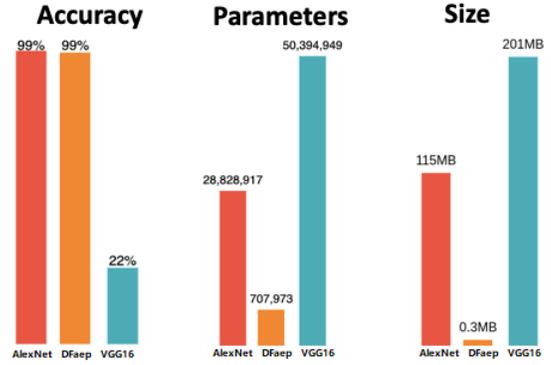


Fig. 8. The Example of Augmented Images.

TABLE III. COMPARISON OF PERFORMANCE IN THREE DIFFERENT CNNs

Network	#Parameters	#Layers	Model description	Accuracy
VGG16	50 M	16	13 conv + 3 fc layers	22 %
AlexNet	24 M	8	5 conv + 3 fc layers	99 %
Dfaep	0.7 M	6	4 conv + 2 fc layers	99 %

predefined setting. The β_1 , β_2 , and ε of Equations 2, 3 and 4 were set as 0.9, 0.999, and 10^{-8} , respectively.

In this research, we attempted AlexNet, VGG16 and Dfaep model for training and validation, respectively. As shown in Fig. 8 and Fig. 9, the visualization of the accuracy and loss during training of VGG16, AlexNet and Dfaep, respectively. The x-axis stands for epoch of training. The y-axis represents the loss and accuracy of training and validation, respectively. It can be observed from the results of training, the VGG16 model with loss 1.60952 and training accuracies 0.2046, the AlexNet model with loss 0.01979 and training accuracies 0.9934 and our Dfaep model with loss 0.01803 and training accuracies 0.9938.

It can be concluded from the observation that the accuracy of the VGG16 model is the lowest 0.2046 (20%), and the accuracy of AlexNet and our Dfaep model is very similar, their accuracy curves close to 1 (100%). However, compared to Dfaep and AlexNet, the depth of the Dfaep model structure is only 6 layers which is shallower than the 8 layers used by AlexNet. Moreover, the parameter amount of Dfaep is less than one-third of AlexNet and the storage size of Dfaep is only 0.3MB, this makes Dfaep's training efficiency and time consumption greatly lead AlexNet, and has achieved close to 100% accuracy as shown in Table III. From observation, because the rate of accuracy and amount of parameters can be changed according to the different construction of the Dfaep such as depth, dense and kernel size. Therefore, even the input data are the identical, the accuracy will not be proportional to the depth of the model.

C. Validation of Proposed Method

In the next experiment, we have made the comparison of the results of internal and external vehicle driving data as shown in Table III. We have matched the prediction results of driver's gaze zone and the salient point of DR(eye)VE

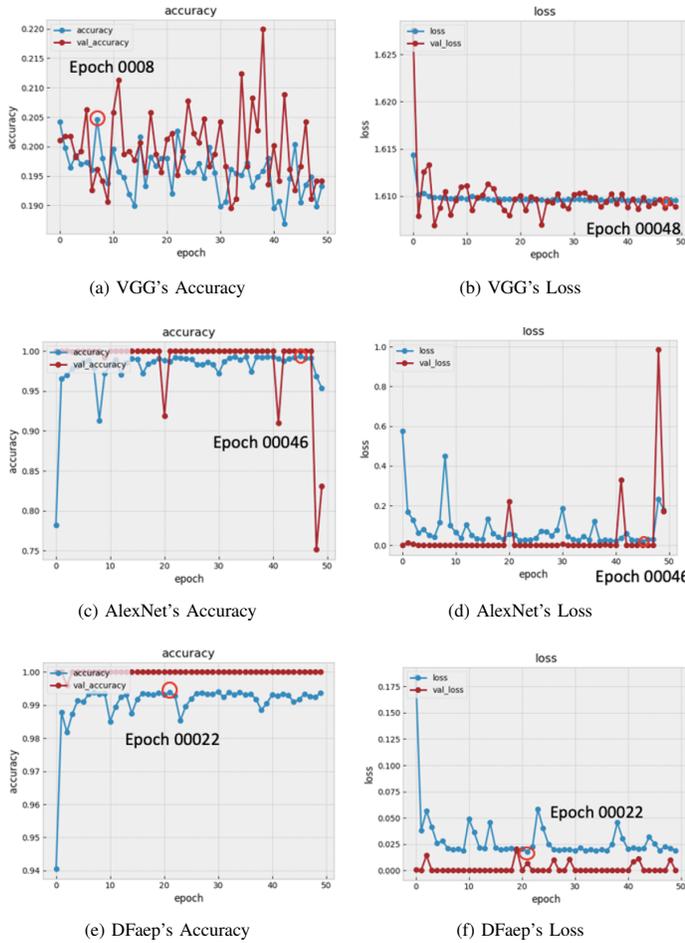


Fig. 9. The Diagram of Loss and Accuracy of Training and Validation According on the Number of Epoch with Three Models of (a-b) VGG; (b-c) AlexNet; and (e-f) DfAep.

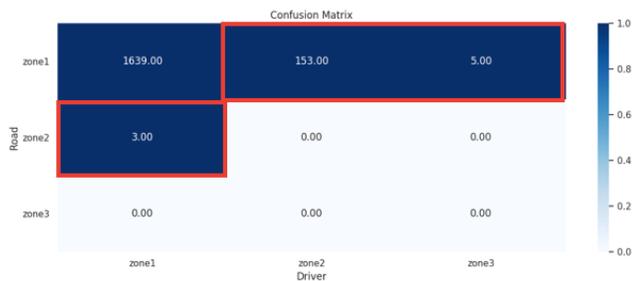


Fig. 10. The Confusion Matrix with the DR(eye)VE Dataset.

database through the confusion matrix (as shown in Fig. 10) for evaluating whether or not the results are consistently as shown in Table III. We compared the prediction results of driver's facial feature inside the vehicle with using external vehicle driving data. Table III presents the comparison results from the data internal and external the vehicle.

“Road” and “Driver” stand for the zone which driver should be attended to and estimation zone of driver, separately. By observing the results of the confusion matrix in Fig. 10, almost

TABLE IV. ESTIMATION ZONES AND ATTENTION AREAS OF DRIVER AND ROAD SCENE

Data	Zone 1	Zone 2	Zone 3	Total
Driver	1787	3	0	1800
Road Scene	1642	153	5	1800



Fig. 11. The Synchronous Comparison of Estimation Zones and Attention Areas.

all the attention zones have concentrated on zone 1. Although the direction of gaze regions is sparse, the method proposed in this research has demonstrated is doable. Later, we obtained the different matched results by confusion matrix for giving appropriate alert to driver as shown in red block of Table IV. These matched gaze zone are mostly concentrate in zone 1. Therefore, we found that it is comply to human driving behaviors in the normal daily driving. In Fig. 11 demonstrates the synchronous comparison of estimation zones and attention areas.

V. CONCLUSION

In this paper, we have proposed a deep learning network to map the driver's attention zone interior of the vehicle environment. For driver gaze zone mapping, driver's face landmark, head position, eye gaze images are captured by the dual low-cost dash cameras. We have performed facial feature tracker with Openface for extracting the images of driver's facial landmark. We have extracted the feature of input image and located the final gaze zone from our network based on the the final score of fully connected layer. The final score of fully connected layer is obtained based on all the extracted features calculated by softmax function to derive the final result. We have compared the losses and accuracies of our idea with three different DNN architecture. Based on the performance of these three models, our DfAep network not only has shown a high scale of accuracy but also performed low level of parameters and storage size. We have demonstrated that such a focus model can detect a large proportion of a driver's focus and is even fast with high acceptable accuracy for detecting distraction and sending a warning or an alert to the driver whenever it is needed. Additionally, we show evidence that such a deep learning network and its trained model provide a feasible way for understanding the driver's attention and focus and making it possible to tame the uncertainty and dynamics of driver's attention and corresponding behaviors.

Empowered by the learnt model trained by the DFaeP network, we have conducted multiple test driving. By comparing the estimation zone of driver and the normal attention area of the road scene, warning is issued when the driver has any abnormal driving behavior or distraction during driving (see Fig. 11). From the results, we can see that the method proposed in this study is feasible and effectiveness. The propose network and trained model can be used to offer potential reduction of driving distraction and help drivers be more focus on roads during driving in such a way that road and driving safety can be significantly improved. In the future, we need to further split the zones and improve the accuracy. In our study, we notice that the distribution of driver's attention and focus are highly skewed. To estimate driver attention and focus, the zones defined in this work are needed to be further splitted or refined dynamically by road scene. Moreover, it is noteworthy to introduce road sense semantic segmentation and object detection into our network to estimate driver's interests for better prediction performance.

REFERENCES

- [1] Fortum, "2016 was the deadliest year on american roads in nearly a decade?." Available at:<http://fortune.com/2017/02/15/traffic-deadliest-year/>, 2017.
- [2] Icebike, "Real time traffic accident statistics." Available at:<http://www.icebike.org/real-time-traffic-accident-statistics/>, 2017.
- [3] C. for Disease Control and prevention, "Vital signs: Motor vehicle injury prevention — united states and 19 comparison countries." Available at:<https://www.textrequest.com/blog/how-much-time-people-spend-mobile-phones-2017/>, 2016.
- [4] J. Briggs, "How in-dash night-vision systems work." Available at:<http://electronics.howstuffworks.com/gadgets/automotive/in-dash-night-vision-system.htm>, 2016.
- [5] D. A. Owens and M. Sivak, "Differentiation of visibility and alcohol as contributors to twilight road fatalities," *Human Factors*, vol. 38, no. 4, pp. 680–689, 1996.
- [6] A. Palazzi, D. Abati, F. Solera, R. Cucchiara, et al., "Predicting the driver's focus of attention: the dr (eye) ve project," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 7, pp. 1720–1733, 2018.
- [7] J. Xu, J. Min, and J. Hu, "Real-time eye tracking for the assessment of driver fatigue," *Healthcare technology letters*, vol. 5, no. 2, pp. 54–58, 2018.
- [8] Y. Tian, K. Pei, S. Jana, and B. Ray, "Deeptest: Automated testing of deep-neural-network-driven autonomous cars," in *Proceedings of the 40th international conference on software engineering*, pp. 303–314, 2018.
- [9] K. A. Brookhuis, D. De Waard, and W. H. Janssen, "Behavioural impacts of advanced driver assistance systems—an overview," *European Journal of Transport and Infrastructure Research*, vol. 1, no. 3, 2001.
- [10] A. Shaout, D. Colella, and S. Awad, "Advanced driver assistance systems—past, present and future," in *2011 Seventh International Computer Engineering Conference (ICENCO'2011)*, pp. 72–82, IEEE, 2011.
- [11] J. Piao and M. McDonald, "Advanced driver assistance systems from autonomous to cooperative approach," *Transport reviews*, vol. 28, no. 5, pp. 659–684, 2008.
- [12] A. Taeihagh and H. S. M. Lim, "Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks," *Transport reviews*, vol. 39, no. 1, pp. 103–128, 2019.
- [13] Waymo, "We are building the world's most experienced drivertm." Available at: <https://waymo.com/>, 2021.
- [14] P. LeBeau, "Waymo starts commercial ride-share service," URL: <https://www.cnbc.com/2018/12/05/waymo-starts-commercial-ride-share-service.html>, 2018.
- [15] AppleInsider, "Apple car." Available at: <https://appleinsider.com/inside/apple-car>, 2021.
- [16] N. Tech, "Navya shuttle brochure." Available at: https://navya.tech/wp-content/uploads/documents/Brochure_Shuttle_EN.pdf, 2021.
- [17] S. Shetty, "Uber's self-driving cars are a key to its path to profitability," 2020.
- [18] P. LeBeau, "Relax, experts say it's at least a decade before you can buy a self-driving vehicle," 2019.
- [19] T. Baltrušaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency, "Openface 2.0: Facial behavior analysis toolkit," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 59–66, IEEE, 2018.
- [20] T. Baltrušaitis, P. Robinson, and L.-P. Morency, "Openface: an open source facial behavior analysis toolkit," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–10, IEEE, 2016.
- [21] M. Carrasco, "Visual attention: The past 25 years," *Vision research*, vol. 51, no. 13, pp. 1484–1525, 2011.
- [22] J. L. Orquin and S. M. Loose, "Attention and choice: A review on eye movements in decision making," *Acta psychologica*, vol. 144, no. 1, pp. 190–206, 2013.
- [23] T. E. Halverson, "An" active vision" computational model of visual search for human-computer interaction. PhD thesis, University of Oregon, 2008.
- [24] S. E. Gaither, K. Pauker, and S. P. Johnson, "Biracial and monoracial infant own-race face perception: An eye tracking study," *Developmental science*, vol. 15, no. 6, pp. 775–782, 2012.
- [25] S. Hutt, C. Mills, S. White, P. J. Donnelly, and S. K. D'Mello, "The eyes have it: Gaze-based detection of mind wandering during learning with an intelligent tutoring system.," *International Educational Data Mining Society*, 2016.
- [26] A. J. Hornof and A. Cavender, "Eyedraw: enabling children with severe motor impairments to draw with their eyes," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 161–170, 2005.
- [27] M. Chau and M. Betke, "Real time eye tracking and blink detection with usb cameras," tech. rep., Boston University Computer Science Department, 2005.
- [28] T. Toyama, T. Kieninger, F. Shafait, and A. Dengel, "Gaze guided object recognition using a head-mounted eye tracker," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 91–98, 2012.
- [29] M. Kassner, W. Patera, and A. Bulling, "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction," in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication*, pp. 1151–1160, 2014.
- [30] D. Li, J. Babcock, and D. J. Parkhurst, "openeyes: a low-cost head-mounted eye-tracking solution," in *Proceedings of the 2006 symposium on Eye tracking research & applications*, pp. 95–100, 2006.
- [31] E. Miluzzo, T. Wang, and A. T. Campbell, "Eyephone: activating mobile phones with your eyes," in *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*, pp. 15–20, 2010.
- [32] X. Zhang, H. Kulkarni, and M. R. Morris, "Smartphone-based gaze gesture communication for people with motor disabilities," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2878–2889, 2017.
- [33] E. Wood and A. Bulling, "Eyetable: Model-based gaze estimation on unmodified tablet computers," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 207–210, 2014.
- [34] Y. Wang, S. Lu, and D. Harter, "Eye tracking and learning analytics for promoting proactive teaching and learning in classroom a survey," in *2020 The 4th International Conference on Education and E-Learning*, pp. 156–160, 2020.
- [35] Y. Wang, S. Lu, and D. Harter, "Towards a collaborative and intelligent framework for pervasive and proactive learning," *Submit to International Journal of Engineering Education (IJEE)*, 2021.
- [36] Y. Wang, S. Lu, and D. Harter, "Multi-sensor systems and infrastructure for capturing student attention and understanding engagement in learning: A review," *Submit to IEEE Sensors Journal*, 2021.

- [37] T. Santini, W. Fuhl, D. Geisler, and E. Kasneci, "Eyerectoo: Open-source software for real-time pervasive head-mounted eye tracking.," in *VISIGRAPP (6: VISAPP)*, pp. 96–101, 2017.
- [38] E. S. Kim, A. Naples, G. V. Gearty, Q. Wang, S. Wallace, C. Wall, J. Kowitz, L. Friedlaender, B. Reichow, F. Volkmar, *et al.*, "Development of an untethered, mobile, low-cost head-mounted eye tracker," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 247–250, 2014.
- [39] L. Yang, K. Dong, A. J. Dmitruk, J. Brighton, and Y. Zhao, "A dual-cameras-based driver gaze mapping system with an application on non-driving activities monitoring," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4318–4327, 2019.
- [40] D. Tran, H. M. Do, W. Sheng, H. Bai, and G. Chowdhary, "Real-time detection of distracted driving based on deep learning," *IET Intelligent Transport Systems*, vol. 12, no. 10, pp. 1210–1219, 2018.
- [41] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, *et al.*, "End to end learning for self-driving cars," *arXiv preprint arXiv:1604.07316*, 2016.
- [42] NVIDIA, "Designed from the ground up for the largest hpc and ai workloads." Available at: <https://www.nvidia.com/en-us/>, 2017.
- [43] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.
- [44] T. Baltrusaitis, P. Robinson, and L.-P. Morency, "Constrained local neural fields for robust facial landmark detection in the wild," in *Proceedings of the IEEE international conference on computer vision workshops*, pp. 354–361, 2013.
- [45] A. Zadeh, Y. Chong Lim, T. Baltrusaitis, and L.-P. Morency, "Convolutional experts constrained local model for 3d facial landmark detection," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 2519–2528, 2017.
- [46] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *ICML*, 2010.
- [47] T. Mikolov, S. Kombrink, L. Burget, J. Černocký, and S. Khudanpur, "Extensions of recurrent neural network language model," in *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 5528–5531, IEEE, 2011.
- [48] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [49] A. Karpathy, "Stanford university cs231n: Convolutional neural networks for visual recognition," *url: http://cs231n.stanford.edu/syllabus.html*, 2018.
- [50] J. Han and C. Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning," in *International Workshop on Artificial Neural Networks*, pp. 195–201, Springer, 1995.
- [51] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for lvcsr using rectified linear units and dropout," in *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 8609–8613, IEEE, 2013.
- [52] J. Nuevo, L. M. Bergasa, and P. Jiménez, "Rsmat: Robust simultaneous modeling and tracking," *Pattern Recognition Letters*, vol. 31, no. 16, pp. 2455–2463, 2010.
- [53] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The cas-peal large-scale chinese face database and baseline evaluations," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 38, no. 1, pp. 149–161, 2007.
- [54] "Berkeley DeepDrive." Available at: <https://bdd-data.berkeley.edu/>.
- [55] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.
- [56] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [57] "Keras api reference-utilities-backend utilities." Available at: https://keras.io/api/utills/backend_utills/.
- [58] "Tensorflow," "tensorflow/tensorflow," github." Available at: <https://github.com/tensorflow/tensorflow>.
- [59] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [60] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, pp. 177–186, Springer.
- [61] "Keras optimizers.." Available at: <https://keras.io/optimizers/>.

Predicting Strength Ratio of Laminated Composite Material with Evolutionary Artificial Neural Network

Huiyao Zhang¹, Atsushi Yokoyama²
Department of Fiber Science and Engineering
Kyoto Institute of Technology
Kyoto, JAPAN

Abstract—In this paper, an alternative methodology to obtain the strength ratio for the laminated composite material is presented. Traditionally, classical lamination theory and related failure criteria are used to calculate the numerical value of strength ratio of laminated composite material under in-plane and out-of-plane loading from a knowledge of the material properties and its layup. In this study, to calculate the strength ratio, an alternative approach is proposed by using an artificial neural network, in which the genetic algorithm is proposed to optimize the search process at four different levels: the architecture, parameters, connections of the neural network, and active functions. The results of the present method are compared to those obtained via classical lamination theory and failure criteria. The results show that an artificial neural network is a feasible method to calculate the strength ratio concerning in-plane loading instead of classical lamination and associated failure theory.

Keywords—Classical lamination theory; genetic algorithm; artificial neural network; optimization

I. INTRODUCTION

Fiber-reinforced composite materials have gained increasing attention due to their superior mechanical performance in stiffness, strength, and specific gravity of fibers over conventional materials. Laminated composite material takes advantage of fiber-reinforced composite material, and finds wide application in a variety of applications, which include electronic packaging, sports equipment, homebuilding, medical prosthetic devices, high-performance military structures, etc. The mechanical properties of composite laminated are determined by stacking sequence, ply thickness, fiber orientation, and material for each ply. Strength ratio [1], [2], [3], [4], [5], [6], [7], [8] is a critical index to predict the performance of a laminated composite material. There are two approaches for solving this problem: analytical methods, such as classical lamination theory (CLT); data-driven methods, such as artificial neural networks (ANN).

The analytical approach involves a two-step procedure to obtain strength ratio: first, develop the stress and strain relationship among in-plane loading using classical lamination theory based on a knowledge of the composite laminate properties of the individual layers and the laminate geometry; then calculate the strength ratio according to associated failure criteria, such as Tsai-Wu failure criterion, based on the above-obtained stress and strain relationship. However, the use of CLT needs intensive computation since it involves massive matrix multiplication and integration operation.

The other approach to this problem is using an artificial neural network, which is a data-driven method, instead of

an analytical method. ANN, heavily inspired by biology and psychology, is a reliable tool instead of a complicated mathematical model, which can accelerate the calculation process and reduce the computation cost. It has been widely used to solve various practical engineering problems in applications [9], [10], such as pattern recognition, nonlinear regression, data mining, clustering, prediction, etc. Evolutionary artificial neural networks are a subclass of artificial neural networks, in which evolutionary algorithms are introduced to design the topology of an ANN. For an artificial neural network, the number of layers, the connection between neurons, the activation functions used in every neuron, etc. are critical components to its performance. The design of an ANN can be treated as an optimization procedure of discrete variables, which can be solved by a genetic algorithm (GA). It is claimed that the combinations of artificial neural networks and evolutionary algorithm [11] can significantly improve the performance of intelligent systems than that rely on ANNs or evolutionary algorithms alone.

GA, inspired by Darwin's principle of survival of the fittest, is widely adopted to obtain the global optimal for discrete optimization problems. The techniques used in this algorithm, such as selection, crossover, mutation, are derived from natural selection, and individuals with better fitness get more chances to breed. Therefore, GA can be integrated into the design of ANN, in which encoding the information of an artificial neural network into a chromosome [12], [13].

The rest of this paper is organized as the following: Section II introduces the CLT and the failure criteria, which is used to check whether the composite material fails or not in the present study; Section III covers the design of an artificial neural network for a function approximation; Section IV reviews the use of the genetic algorithm in the design of neural network architecture, and the techniques of parameters optimization during the training process; Section V presents the result of the numerical experiments in different cases; in the conclusion part, we present and discuss the experiment results.

II. CLASSICAL LAMINATION THEORY AND FAILURE CRITERIA

A. Classical Lamination Theory

Classical lamination theory derives from three simplifying assumptions in laminated composite material: the laminate consist of plies bonded together through the thickness, the thickness of each ply is small, and it is consists of homogeneous, orthotropic material; the entire laminated composite

is only under in-plane loading; the Normal cross-section of the laminate is vertical to the deflected middle surface. Fig. 1 shows the coordinate system used for an angle lamina. The axis in the 1-2 coordinate system is called the local axis or the material axis, and the axis in the x-y coordinate system is called the global axis.

A few cases of laminates, such as symmetric laminates, cross-ply laminates, play an important role in the application of laminated composite material. A laminate is called an angle ply laminate if it has plies of the same material and thickness and is only oriented at $+\theta$ and $-\theta$ directions. A model of an angle ply laminate is as shown in Fig. 2.

1) *Stress and Strain in a Lamina:* For a single lamina under in-plane loading whose thickness is relatively small, suppose the upper and lower surfaces of the lamina are free from external loading. According to Hooke's law, the three-dimensional stress-strain equations can be reduced to two-dimensional stress-strain equations in the composite material. The stress-strain relation in local axis 1-2 is

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{12} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{bmatrix}, \quad (1)$$

where Q_{ij} is the stiffness of a lamina. And they are related to engineering elastic constants as follows:

$$\begin{aligned} Q_{11} &= \frac{E_1}{1-\nu_{12}\nu_{21}}, \\ Q_{22} &= \frac{E_2}{1-\nu_{12}\nu_{21}}, \\ Q_{66} &= G_{12}, \\ Q_{12} &= \frac{\nu_{21}E_2}{1-\nu_{12}\nu_{21}}, \end{aligned} \quad (2)$$

where $E_1, E_2, \nu_{12}, G_{12}$ are four independent engineering elastic constants, which are defined as follows: E_1 is the longitudinal Young's modulus, E_2 is the transverse Young's modulus, ν_{12} is the major Poisson's ratio, and G_{12} is the in-plane shear modulus.

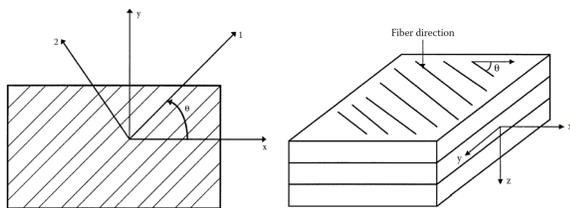


Fig. 1. The Left Diagram Shows the Local and Global Axis of an Angle Lamina, which is from a Laminate as Shown in the Right Diagram.

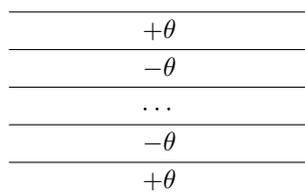


Fig. 2. Model for Angle Ply Laminate.

Stress-strain relation in the global x-y axis is

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} & \bar{Q}_{16} \\ \bar{Q}_{12} & \bar{Q}_{22} & \bar{Q}_{26} \\ \bar{Q}_{16} & \bar{Q}_{26} & \bar{Q}_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix}, \quad (3)$$

where

$$\begin{aligned} \bar{Q}_{11} &= Q_{11}\cos^4\theta + Q_{22}\sin^4\theta + 2(Q_{12} + 2Q_{66})\sin^2\theta\cos^2\theta, \\ \bar{Q}_{12} &= (Q_{11} + Q_{22} - 4Q_{66})\sin^2\theta\cos^2\theta + Q_{12}(\cos^4\theta + \sin^4\theta), \\ \bar{Q}_{22} &= Q_{11}\sin^4\theta + Q_{22}\cos^4\theta + 2(Q_{12} + 2Q_{66})\sin^2\theta\cos^2\theta, \\ \bar{Q}_{16} &= (Q_{11} - Q_{12} - 2Q_{66})\cos^3\theta\sin\theta - (Q_{22} - Q_{12} - 2Q_{66})\sin^3\theta\cos\theta, \\ \bar{Q}_{26} &= (Q_{11} - Q_{12} - 2Q_{66})\cos\theta\sin^3\theta - (Q_{22} - Q_{12} - 2Q_{66})\cos^3\theta\sin\theta, \\ \bar{Q}_{66} &= (Q_{11} + Q_{22} - 2Q_{12} - 2Q_{66})\sin^2\theta\cos^2\theta + Q_{66}(\sin^4\theta + \cos^4\theta). \end{aligned} \quad (4)$$

2) *Stress and Strain in a Laminate:* For forces acting on laminates, such as in plate and shell structures, the relationship between applied forces and displacement can be given by

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{16} \\ A_{12} & A_{22} & A_{26} \\ A_{16} & A_{26} & A_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^0 \\ \varepsilon_y^0 \\ \gamma_{xy}^0 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{16} \\ B_{11} & B_{12} & B_{16} \\ B_{16} & B_{26} & B_{66} \end{bmatrix} \begin{bmatrix} k_x \\ k_y \\ k_{xy} \end{bmatrix}, \quad (5)$$

where N_x, N_y refers to the normal force per unit length; N_{xy} means shear force per unit length; ε^0 and k_{xy} denotes mid plane strains and curvature of a laminate in x-y coordinates. The mid-plane strain and curvature is given by

$$\begin{aligned} A_{ij} &= \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k - h_{k-1}) i = 1, 2, 6, j = 1, 2, 6, \\ B_{ij} &= \frac{1}{2} \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k^2 - h_{k-1}^2) i = 1, 2, 6, j = 1, 2, 6, \\ D_{ij} &= \frac{1}{3} \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k^3 - h_{k-1}^3) i = 1, 2, 6, j = 1, 2, 6. \end{aligned} \quad (6)$$

The [A], [B], and [D] matrices are called the extensional, coupling, and bending stiffness matrices, respectively. The extensional stiffness matrix [A] relates the resultant in-plane forces to the in-plane strains, and the bending stiffness matrix [D] couples the resultant bending moments to the plane curvatures. The coupling stiffness matrix [B] relates the force and moment terms to the midplane strains and curvatures.

B. Failure Criteria for a Lamina

Failure criteria for composite materials are more difficult to predict due to structural and material complexity. The failure process of composite materials can be regarded from microscopic and macroscopic points of view. The most popular criteria about the failure of an angle lamina are from the macroscopic point of view, which are according to the tensile, compressive, and shear strengths. As shown in Fig. 3, there are two types of failure criteria [14], [15], [16], [17], [18], [19], [20], [21] according to failure surfaces. The first failure surface is a rectangle that includes the maximum stress failure criterion [22], and maximum strain failure criterion. The second failure surface is ellipsoidal that includes Tsai-Wu [23], [24], Chamis, Hoffman, and Hill criteria. In the present study, the two most

reliable failure criteria are adopted, Maximum stress and Tsai-wu. Both these failure criteria are based on the stress in the local axis instead of principal normal stress and maximum shear stresses, in which four normal strength parameters and one shear stress are involved. The five strength parameters are:

$(\sigma_1^T)_{ult}$ = ultimate longitudinal tensile strength(in direction 1),

$(\sigma_1^C)_{ult}$ = ultimate longitudinal compressive strength,

$(\sigma_2^T)_{ult}$ = ultimate transverse tensile strength,

$(\sigma_2^C)_{ult}$ = ultimate transverse compressive strength, and

$(\tau_{12})_{ult}$ = and ultimate in-plane shear strength.

1) *Maximum stress (MS) failure criterion:* Maximum stress failure criteria are consist of the normal stress theory and the shear stress theory. The stress applied to a lamina can be resolved into the normal stress and shear stress in the local axis. The lamina fails if either of the normal stress or shear stress in the local axis of a lamina is equal or exceeds the corresponding ultimate strengths of the unidirectional lamina. That is,

$$\begin{aligned} \sigma_1 &\geq (\sigma_1^T)_{ult} & \text{or} & & \sigma_1 &\leq -(\sigma_1^C)_{ult}, \\ \sigma_2 &\geq (\sigma_2^T)_{ult} & \text{or} & & \sigma_2 &\leq -(\sigma_2^C)_{ult}, \\ \tau_{12} &\geq (\tau_{12})_{ult} & \text{or} & & \tau_{12} &\leq -(\tau_{12})_{ult}, \end{aligned} \quad (7)$$

where σ_1 and σ_2 are the normal stresses in the local axis 1 and 2; τ_{12} is the shear stress in the symmetry plane 1-2.

2) *Tsai-Wu failure criterion:* The Tsai-Wu criterion is one of the most reliable static failure criteria derived from the von Mises yield criterion. A lamina is considered to fail if

$$\begin{aligned} H_{11}\sigma_1 + H_{22}\sigma_2 + H_6\tau_{12} + H_{11}\sigma_1^2 + H_{22}\sigma_2^2 \\ + H_{66}\tau_{12}^2 + 2H_{12}\sigma_1\sigma_2 < 1 \end{aligned} \quad (8)$$

is violated, where

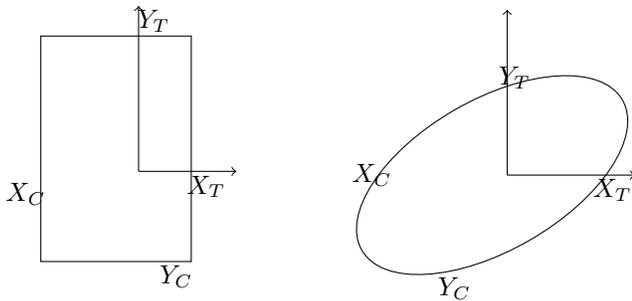


Fig. 3. Schematic Failure Surfaces for Maximum Stress and Quadratic Failure Criteria.

$$\begin{aligned} H_1 &= \frac{1}{(\sigma_1^T)_{ult}} - \frac{1}{(\sigma_1^C)_{ult}}, \\ H_{11} &= \frac{1}{(\sigma_1^T)_{ult}(\sigma_1^C)_{ult}}, \\ H_2 &= \frac{1}{(\sigma_2^T)_{ult}} - \frac{1}{(\sigma_2^C)_{ult}}, \\ H_{22} &= \frac{1}{(\sigma_2^T)_{ult}(\sigma_2^C)_{ult}}, \\ H_{66} &= \frac{1}{(\tau_{12})_{ult}^2}, \\ H_{12} &= -\frac{1}{2} \sqrt{\frac{1}{(\sigma_1^T)_{ult}(\sigma_1^C)_{ult}(\sigma_2^T)_{ult}(\sigma_2^C)_{ult}}}. \end{aligned} \quad (9)$$

H_i is the strength tensor of the second-order; H_{ij} is the strength tensor of the fourth-order. σ_1 is the applied normal stress in direction 1; σ_2 is the applied normal stress in direction 2; τ_{12} is the applied in-plane shear stress.

3) *Strength ratio:* The safety factor, or yield stress, is how much extra load beyond is intended a composite laminate will take. The strength ratio(SR) is defined as

$$SR = \frac{\text{Maximum Load Which Can Be Applied}}{\text{Load Applied}}. \quad (10)$$

III. EVOLUTIONARY ARTIFICIAL NEURAL NETWORK

A. General Neural Network

In this paper, the feedforward ANN is adopted in the current study, since it is straightforward to code. For function approximation through an artificial neural network, Cybenko demonstrated that a two-layer perceptron can form an arbitrarily close approximation to any continuous nonlinear mapping [25]. Therefore, a two-layer feedforward ANN is proposed in the present study. Fig. 4 shows a general framework for a two-layer ANN, in which the number of nodes in the hidden layer and the connection with inputs, are critical in the design of an ANN. The nodes in the hidden layer are treated as feature extractors or detectors. Therefore, nodes within this layer should partially be connected with the inputs of an ANN, since the unnecessary connections would increase the model's complicity, which will reduce an ANN's performance. The number of nodes in the hidden layer should be less than the number of inputs since the nodes in the hidden layer are features. For the nodes in the last layer, every node should be fully connected with nodes in the previous layer, the relationship between the outputs and features should be direct. The rest, which affects the performance of an artificial neural network, are the activation function, and ANNs training method. In the following section, the i th node in the input layer is denoted as i_i , and the j th node in the hidden layered denoted as h_j , respectively.

B. Activation Function

The activation function is one of the critical parts of an ANN. Liu [12] et al. claims that the performance of neural

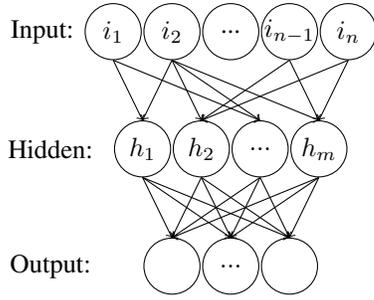


Fig. 4. Network Diagram for the Two-layer Neural Network. The Input, Hidden, and Output Variables are Represented by Nodes, and the Weight Parameters are Represented by Links between the Nodes. Arrows Denote the Direction of Information Flow Through the Network During Forward Propagation.

networks with different activation functions is different, even if they have the same architecture. A generalized activation function can be written as

$$y_i = f_i\left(\sum_{j=1}^n w_{ij}x_j - \theta\right) \quad (11)$$

where y_i is the output of the node i , x_j is the j th input to the node, and w_{ij} is the connection weight between adjacent nodes i and j . Table I display the most widely adopted activation functions in the design of an ANN, which is used in the current study.

C. Weights Learning

The weight training in an ANN is to minimize the error function, such as the most widely used mean square error function, which calculates the difference between the desired and the prediction output values averaged overall examples. Gradient descent algorithm is widely adopted to reduce the value of an error function, which has been successfully applied in many practical areas. However, this class of algorithms is plagued by the possible existence of local minima or “flat spots” and “the curse of dimensionality”. One method to overcome this problem is to adopt a genetic algorithm.

IV. METHODOLOGY

For an angle ply laminate, its strength ratio can be computed based on Tsai-Wu failure theory or maximum stress theory given the laminate’s lay-up, material properties, in-plane loading, etc. To model this function, we propose an ANN framework as shown in Fig. 6, which derives from the previous two-layer model. There are sixteen inputs of this ANN, which are in-plane loading N_x , N_y , and N_{xy} ; design parameters of a

laminate, two fiber orientation θ_1 and θ_2 , ply thickness t , total number of plies N ; five engineering constants of composite materials, E_1 , E_2 , G_{12} , and ν_{12} ; five strength parameters of a unidirectional lamina. Two outputs are strength ratio according to MS theory and strength ratio according to Tsai-Wu theory.

The work involved in the evolution process of ANN consists of three parts: search space, which includes the topology of an ANN, activation function, etc.; search strategy, which details how to explore the search space; performance estimation strategy refers to the measurement of the performance of an artificial neural network.

A. Search Space

We propose a general neural network framework as shown in Fig. 4. The search space is parameterized by four parts: (1) the number of nodes m (possibly unbounded) in the hidden layer, to further narrow down the search space, the assumption is m less than n ; (2) the type of operation every node executes, e.g., sigmoid, linear, Gaussian; (3) the connection relationship between the hidden nodes and inputs (4) the weight value in the connection if a connection exists.

Therefore, the evolution process in an evolutionary artificial neural network can be divided into four different levels: topology, learning rules, active functions, and connection weights. For the evolution of the topology, the aim is to find an optimal ANN architecture for a specific problem. The architecture

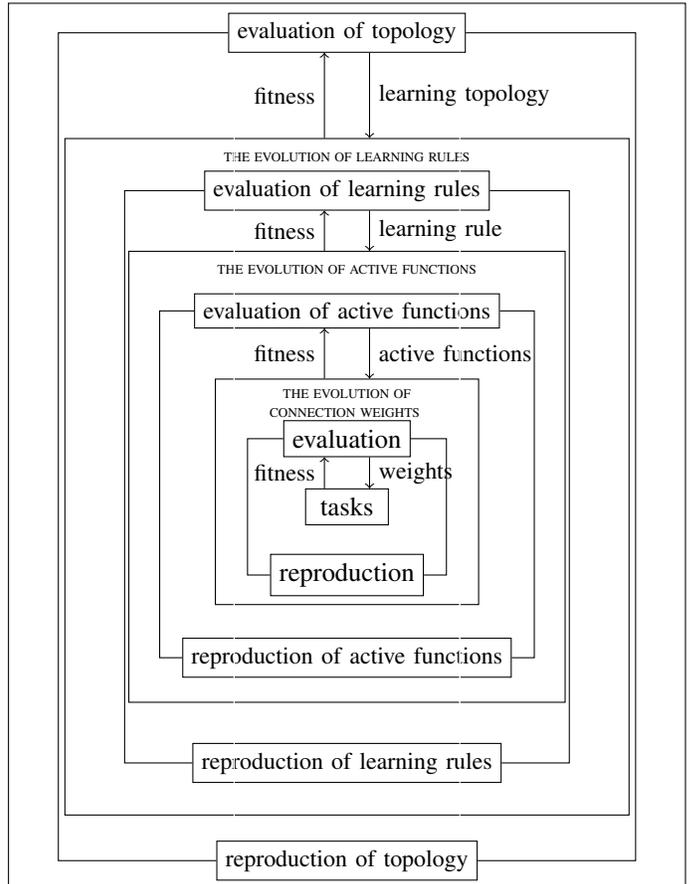


Fig. 5. A General Framework for Evolutionary Neural Network, in which Fitness Refers to the Corresponding Value of Objective Function.

TABLE I. EXAMPLES OF WIDELY USED ACTIVATION FUNCTIONS IN THE DESIGN OF AN ARTIFICIAL NEURAL NETWORK

Type	Description	Formula	Range	Encoding
Linear	The output is proportional to the input	$f(x) = cx$	$(-\infty, +\infty)$	00
Sigmoid	A family of S-shaped functions	$f(x) = \frac{1}{1+e^{-cx}}$	$(0, 1)$	01
ReLU	A piece-wise function	$f(x) = \max\{0, x\}$	$(0, +\infty)$	10
Softplus	A family of S-shaped functions	$f(x) = \ln(1 + e^x)$	$(0, +\infty)$	11

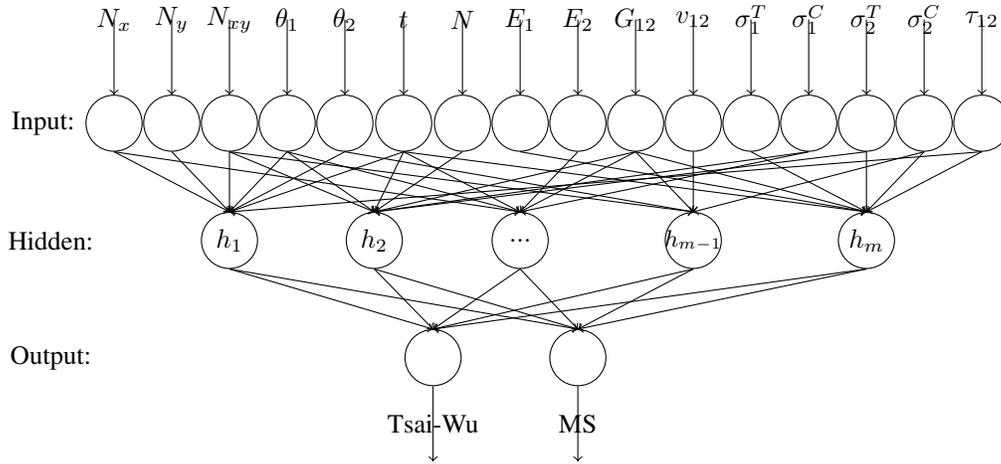


Fig. 6. Diagram for Modeling the Target Function of Strength Ratio Calculating for an Angle Ply Laminate.

TABLE II. THE BINARY REPRESENTATION OF PARENT 1, PARENT 2, AND CHILD CORRESPONDING TO FIG. 7(A), (B) AND (C), WITH i_1, i_2, \dots, i_{16} DENOTE SIXTEEN INPUTS AND h_1, h_2, \dots, h_{12} REFER TO NODES IN THE HIDDEN LAYER. 1 REPRESENTS AN EDGE FROM THE INPUT NODE TO THE HIDDEN NODE, AND 0 REPRESENTS NO EDGE FROM THE INPUT NODES TO THE HIDDEN NODE.

Hidden	Nodes	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9	i_{10}	i_{11}	i_{12}	i_{13}	i_{14}	i_{15}	i_{16}	f	f	
P1	h_1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	
	h_2	0	1	1	1	0	0	0	1	0	0	1	1	0	0	0	0	1	1	
	h_3	1	0	0	1	0	1	1	0	1	1	0	0	1	0	0	0	0	0	
	h_4	0	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	0	1	
	h_5	0	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	0	1
P2	h_1	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	0	
	h_2	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	h_3	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	1	1
	h_4	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	h_5	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	1
	h_6	0	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	0	1
	h_7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
	h_8	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0
	h_9	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1
	h_{10}	0	1	1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	1	1
	h_{11}	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	1
	h_{12}	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1
Child	h_1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	
	h_2	0	1	1	1	0	0	0	1	0	0	1	1	0	0	0	0	1	1	
	h_1	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	1	0	
	h_2	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	h_3	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1
	h_4	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
h_5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	
h_6	0	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	0	1	

of a neural network determines the information processing capability in an application, which is the foundation of the ANN. Two critical issues are involved in the search process of an ANN architecture: the representation and the search operators. Fig. 5 summarizes these four levels of evolution in an ANN.

B. Search Strategy

It is necessary to define related operations during the GA process, which includes the representation of an artificial neural network, the fitness function that determines how good a solution is, and the search operators, such as selection, mutation, and crossover.

For the representation of an ANN, encode the h_i node as an eighteen digits binary string. The initial sixteen digits in the string correspond to the connections between i_i and h_i , with '1' implying there exists a connection between them, with '0' implying no connection exists. The last two digits in the string refer to an activation function, such as "01" which means a sigmoid function. Table II are examples of the binary representation of ANNs whose architectures are as shown in Fig. 7.

For the objective function, treat the multiplicative inverse of the mean squared error, which is the difference between the target and actual output averaged overall examples, as the fitness function.

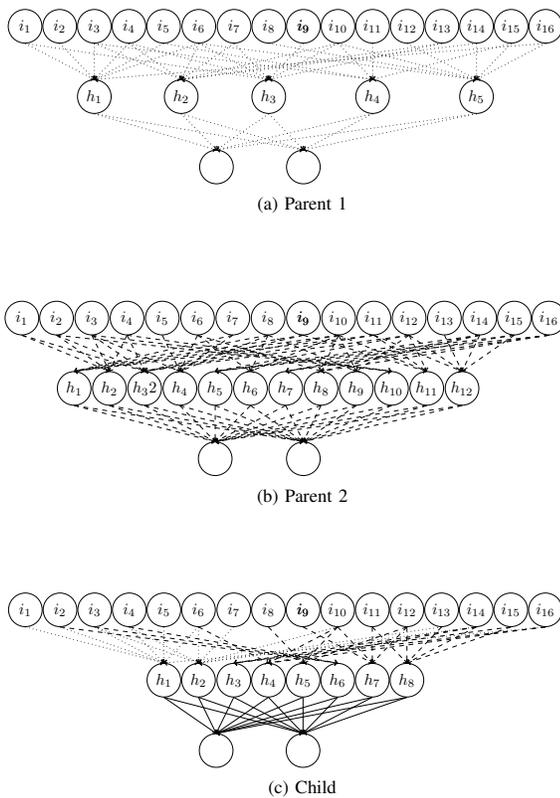


Fig. 7. Examples of three ANNs, with (a) and (b) as parent ANNs, and (c) as the child of (a) and (b). child c inherits the connection relationship part from parent 1 denoted by the darker dashed lines, and the rest from parent 2 denoted by the gray dashed line.

The crossover between individuals results in exploiting the area between the given two-parent solutions. In the present study, we search the local area by combining the genes of half number of nodes from both parents. Fig. 7 illustrates the crossover operator: Fig. 7(c) is the child of Fig. 7(a) and Fig. 7(b), the connection relationship of hidden nodes with inputs are from both parents, and the corresponding activation functions are also from both parents. In the binary representation Table II, it is shown that the first two rows of the child are the same as the first two rows of parent P_1 , and the last six rows of the child are the same as the first six rows of parent P_2 .

C. Performance Estimation Strategy

The simplest approach to this problem is to perform standard training and validation of the architecture on a dataset. However, this method is inefficient and computationally intensive. Therefore, much recent research [26] focuses on developing strategy reducing the cost of performance estimation. In this work, during the GA process, we adopt the following straightforward and efficient method to estimate the performance of an ANN: first, train a neural network one hundred times on the training dataset; second, do the validation test; measure the neural network's performance according to its fitness of objective function on the test dataset.

TABLE III. COMPARISON OF THE CARBON/EPOXY, GRAPHITE/EPOXY, AND GLASS/EPOXY PROPERTIES

Property	Symbol	Unit	Carbon/Epoxy	Graphite/Epoxy	Glass/Epoxy
Longitudinal elastic modulus	E_1	GPa	116.6	181	38.6
Transverse elastic modulus	E_2	GPa	7.67	10.3	8.27
Major Poisson's ratio	ν_{12}		0.27	0.28	0.26
Shear modulus	G_{12}	GPa	4.17	7.17	4.14
Ultimate longitudinal tensile strength	$(\sigma_1)_{ult}$	MP	2062	1500	1062
Ultimate longitudinal compressive strength	$(\sigma_1)_{ult}$	MP	1701	1500	610
Ultimate transverse tensile strength	$(\sigma_2)_{ult}$	MPa	70	40	31
Ultimate transverse compressive strength	$(\sigma_2)_{ult}$	MPa	240	246	118
Ultimate in-plane shear strength	$(\tau_{12})_{ult}$	MPa	105	68	72
Density	ρ	g/cm^3	1.605	1.590	1.903
Cost			8	2.5	1

TABLE IV. EXAMPLES OF THE TRAINING DATA

Load	Input			Output	
	Laminate Structure	Material Property	Failure Property	MS	Tsai-Wu
-70,-10,-40,	90,-90,4,1.27,	38.6,8.27,0.26,4.14,	1062.0,610.0,31,118,72,	0.0102,	0.0086
-10,10,0,	-86,86,80,1.27,	181.0,10.3,0.28,7.17,	1500.0,1500.0,40,246,68,	0.4026,	2.5120
-70,-50,80,	-38,38,4,1.27,	116.6,7.67,0.27,4.173,	2062.0,1701.0,70,240,105,	0.0080,	0.0325
-70,80,-40,	90,-90,48,1.27,	38.6,8.27,0.26,4.14,	1062.0,610.0,31,118,72,	0.0218,	0.1028
-20,-30,0,	-86,86,60,1.27,	181.0,10.3,0.28,7.17,	1500.0,1500.0,40,246,68,	0.6481,	0.9512
0,-40,0,	74,-74,168,1.27,	181.0,10.3,0.28,7.17,	1500.0,1500.0,40,246,68,	1.3110,	3.9619

V. EXPERIMENT

In the previous section, we present the details of our strategies for designing an ANN. In this section, we explain the details of the preparation of the training dataset and validation dataset.

A. Dataset Preparation

For composite material, it is impossible to obtain massive training data from the practical scenario. Therefore, we use classical lamination theory and failure theory, which follows a two-step procedure: first, evaluate the stress and strain according to classic lamination theory; second, substitute them into the corresponding equation to get the strength ratio. Repeat this procedure to yield 14000 points uniformly distributed over the domain space, and define the domain of the corresponding inputs as follows: the range of in-plane loading varies from 0 to 120; the range of fiber orientation θ is from -90 to 90; ply thickness t is 1.27mm, the number of plies ranges N is from 4 to 120. Three different composite material is used in this experiment, as shown in Table III. Table IV shows part of the training data, which are randomly selected from the generated training dataset. To speeds up the learning and accelerate convergence, the input attributes of the dataset are rescaled to between 0 and 1.0 by a linear function.

B. ANN Training and Validation

The ANN training procedure is carried out by optimizing the multinomial logistic regression objective using mini-batch gradient descent [27] with momentum. The batch size is set to 1000, momentum to 0.9. the learning rate is set to 10^{-2} . The ratio of the training dataset and validation dataset is 70/30, with 70% of the entire data for training and 30% for validation.

C. Genetic Algorithm

The genetic algorithm involves the evolution of an artificial neural network's topology, activation function, etc. in the optimizing process. The corresponding parameters are as the following. The population is 10, the percentage of parents in

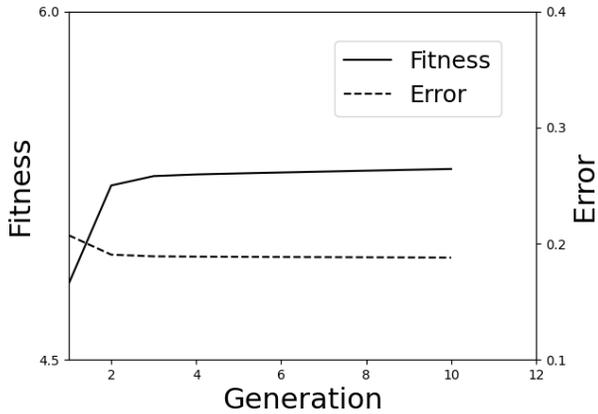


Fig. 8. Fitness and Averaged Sum-of-squares Errors of the Pre-trained Artificial Neural Network as Generations Proceed.

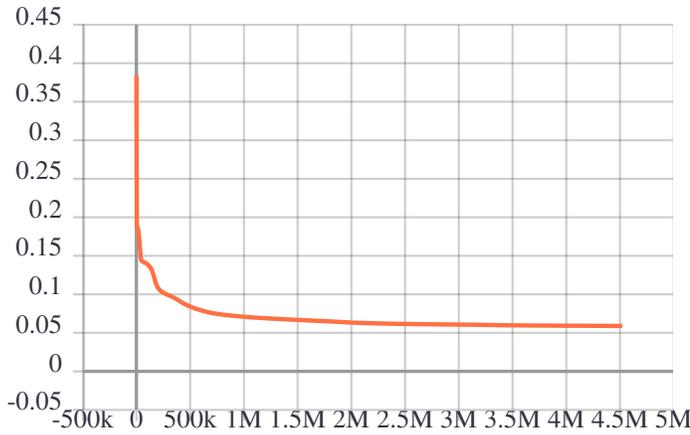


Fig. 9. The illustration of the Behaviour of Fitness on the Training Dataset During the Training Session.

the population is 40%; the strategy of selecting parents is rank-based; the mutation rate of the offspring is 0.3.

VI. RESULT AND DISCUSSION

In this work, we propose to use an artificial neural network as an alternative way to compute the strength ratio of composite material instead of a two-step procedure, based on classical lamination and failure theory. Fig. 8 shows the changes of the fitness and error during the evolution procedure. The fitness is obtained through the performance estimation technique of an artificial neural network. As shown in this figure, fitness grows during the initial stage; then, it slowly converges as generation proceeds. It implies genetic algorithm can find a better artificial neural network with the evolution of the number of neurons in the hidden layer, connection relationship, activation functions, and connection weights.

Fig. 9 shows the rest training of the artificial neural network obtained from the GA, which is a pre-trained ANN. Continue to train it with a standard gradient-based descent algorithm until the error converges. The target neural network converges rapidly at first, and further training doesn't reduce the error

TABLE V. ANN PREDICTIONS OF THE TSAI-WU AND MS STRENGTH RATIO WITH THE NUMERICAL RESULTS OBTAINED BY CLT.

Load	Input			Output			
	Laminate Structure	Material Property	Failure Property	CLT MS Tsai-Wu		ANN MS Tsai-Wu	
-10,40,20	26,-26,168,1.27	116,6,7,67,0,27,4,17	2062,0,1701,0,70,240,105	0.342	0.476	0.351	0.492
20,-70,-30	10,-10,196,1.27	181,0,10,3,0,28,7,17	1500,0,1500,0,40,246,68	0.653	0.489	0.612	0.445
60,-20,0	82,-82,128,1.27	181,0,10,3,0,28,7,17	1500,0,1500,0,40,246,68	1.663	0.112	1.673	0.189

efficiently. Then, this artificial neural network is used to predict the strength ratio of laminated composite material.

To present the evaluation result of the ANN straightforwardly, several experiment results from the validation dataset are displayed in Table V, which are randomly selected. Comparing the strength ratio outputs based on CLT and ANN from Table V, it is shown that the calculation of strength ratio can be achieved using a two-layer neural network, without the intensive computation of matrix multiplication.

VII. CONCLUSION

In this paper, an evolutionary artificial neural network model was developed to predict the strength ratio of laminated composite material under in-plane loading. We review the use of genetic algorithms and artificial neural networks as an alternative approach for calculating the strength ratio of an angle ply laminate under in-plane loading. Traditionally, it is obtained through CLT and corresponding failure criteria, such as Maximum Stress theory and Tsai-Wu failure theory.

The main contribution of this work is as follows: 1) propose a two-layer diagram model for designing a sophisticated neural network in simulating the calculation of strength ratio, and use a genetic algorithm to explore the search space; 2) suggest an efficient method to compute the strength ratio instead of adopting the two-step procedure based on classical lamination theory and related failure criteria. Compared with experimentally obtained data, it is demonstrated that ANN is an efficient and simple tool to compute the strength ratio, instead of the complex analytical mathematical model. Our findings underline the practical applicability of ANN on the analysis of composite material.

There are more improvements we can make over the search strategy and application in the area of laminated composite material. The future work is to develop a more sophisticated ANN, which not only can predict the properties for angle ply laminate, but also the other type of laminated composite material.

ACKNOWLEDGMENT

The work has partly been supported by China Scholarship Council(CSC) under grant no. 201806630112

REFERENCES

- [1] A. Todoroki and R. T. Haftka, "Stacking sequence optimization by a genetic algorithm with a new recessive gene like repair strategy," *Composites Part B: Engineering*, vol. 29, no. 3, pp. 277–285, 1998.
- [2] B. Liu, R. T. Haftka, M. A. Akgün, and A. Todoroki, "Permutation genetic algorithm for stacking sequence design of composite laminates," *Computer methods in applied mechanics and engineering*, vol. 186, no. 2-4, pp. 357–372, 2000.

- [3] K. Sivakumar, N. Iyengar, and K. Deb, "Optimum design of laminated composite plates with cutouts using a genetic algorithm," *Composite Structures*, vol. 42, no. 3, pp. 265–279, 1998.
- [4] M. Walker and R. E. Smith, "A technique for the multiobjective optimisation of laminated composite structures using genetic algorithms and finite element analysis," *Composite structures*, vol. 62, no. 1, pp. 123–128, 2003.
- [5] C.-C. Lin and Y.-J. Lee, "Stacking sequence optimization of laminated composite structures using genetic algorithm with local improvement," *Composite structures*, vol. 63, no. 3-4, pp. 339–345, 2004.
- [6] J.-H. Kang and C.-G. Kim, "Minimum-weight design of compressively loaded composite plates and stiffened panels for postbuckling strength by genetic algorithm," *Composite structures*, vol. 69, no. 2, pp. 239–246, 2005.
- [7] M. Murugan, S. Suresh, R. Ganguli, and V. Mani, "Target vector optimization of composite box beam using real-coded genetic algorithm: a decomposition approach," *Structural and Multidisciplinary Optimization*, vol. 33, no. 2, pp. 131–146, 2007.
- [8] M. Akbulut and F. O. Sonmez, "Optimum design of composite laminates for minimum thickness," *Computers & Structures*, vol. 86, no. 21-22, pp. 1974–1982, 2008.
- [9] S. Yan, X. Zou, M. Ilkhani, and A. Jones, "An efficient multiscale surrogate modelling framework for composite materials considering progressive damage based on artificial neural networks," *Composites Part B: Engineering*, vol. 194, p. 108014, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1359836820303279>
- [10] N. Mentges, B. Dashtbozorg, and S. Mirkhalaf, "A micromechanics-based artificial neural networks model for elastic properties of short fiber composites," *Composites Part B: Engineering*, vol. 213, p. 108736, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1359836821001281>
- [11] F. Lobo, C. F. Lima, and Z. Michalewicz, *Parameter setting in evolutionary algorithms*. Springer Science & Business Media, 2007, vol. 54.
- [12] Y. Liu and X. Yao, "Evolutionary design of artificial neural networks with different nodes," in *Proceedings of IEEE international conference on evolutionary computation*. IEEE, 1996, pp. 670–675.
- [13] S. Rodzin, O. Rodzina, and L. Rodzina, "Neuroevolution: problems, algorithms, and experiments," in *2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)*. IEEE, 2016, pp. 1–4.
- [14] T. N. Massard, "Computer sizing of composite laminates for strength," *Journal of reinforced plastics and composites*, vol. 3, no. 4, pp. 300–345, 1984.
- [15] J. Reddy and A. Pandey, "A first-ply failure analysis of composite laminates," *Computers & Structures*, vol. 25, no. 3, pp. 371–393, 1987.
- [16] C. Fang and G. S. Springer, "Design of composite laminates by a monte carlo method," *Journal of composite materials*, vol. 27, no. 7, pp. 721–753, 1993.
- [17] A. Soeiro, C. C. António, and A. T. Marques, "Multilevel optimization of laminated composite structures," *Structural optimization*, vol. 7, no. 1-2, pp. 55–60, 1994.
- [18] J. L. Pelletier and S. S. Vel, "Multi-objective optimization of fiber reinforced composite laminates for strength, stiffness and minimal mass," *Computers & structures*, vol. 84, no. 29-30, pp. 2065–2080, 2006.
- [19] P. Jadhav and P. R. Mantena, "Parametric optimization of grid-stiffened composite panels for maximizing their performance under transverse loading," *Composite structures*, vol. 77, no. 3, pp. 353–363, 2007.
- [20] S. Omkar, R. Khandelwal, S. Yathindra, G. N. Naik, and S. Gopalakrishnan, "Artificial immune system for multi-objective design optimization of composite structures," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 8, pp. 1416–1429, 2008.
- [21] A. Choudhury, S. Mondal, and S. Sarkar, "Failure analysis of laminated composite plate under hygro-thermo mechanical load and optimisation," *International Journal of Applied Mechanics and Engineering*, vol. 24, no. 3, pp. 509–526, 2019.
- [22] R. Watkins and A. Morris, "A multicriteria objective function optimization scheme for laminated composites for use in multilevel structural optimization schemes," *Computer Methods in Applied Mechanics and Engineering*, vol. 60, no. 2, pp. 233–251, 1987.
- [23] P. Martin, "Optimum design of anisotropic sandwich panels with thin faces," *Engineering optimization*, vol. 11, no. 1-2, pp. 3–12, 1987.
- [24] C. M. Soares, V. F. Correia, H. Mateus, and J. Herskovits, "A discrete model for the optimal design of thin composite plate-shell type structures using a two-level approach," *Composite structures*, vol. 30, no. 2, pp. 147–157, 1995.
- [25] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [26] B. Baker, O. Gupta, R. Raskar, and N. Naik, "Accelerating neural architecture search using performance prediction," *arXiv preprint arXiv:1705.10823*, 2017.
- [27] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.

UIP2SOP: A Unique IoT Network applying Single Sign-On and Message Queue Protocol

Lam Nguyen Tran Thanh¹, Nguyen Ngoc Phien^{2*}, The Anh Nguyen³, Hong Khanh Vo⁴,
Hoang Huong Luong⁵, Tuan Dao Anh⁶, Khoi Nguyen Huynh Tuan⁷, Ha Xuan Son⁸
VNPT Information Technology Company, Ho Chi Minh city, Vietnam¹
Center for Applied Information Technology, Ton Duc Thang University,
Ho Chi Minh City, Viet Nam²
Faculty of Information Technology, Ton Duc Thang University,
Ho Chi Minh City, Vietnam²
Department of Computer Science, Faculty of Electrical Engineering and Computer Science,
VSB-Technical University of Ostrava, Ostrava, Czech Republic²
FPT University, Can Tho City, Viet Nam^{3,4,5,6,7}
University of Insubria, Varese, Italy⁸
Corresponding Author: Ton Duc Thang University, Ho Chi Minh City, Vietnam*

Abstract—Internet of Things (IoT), currently, plays an importance role in our life, also, this is one of the most rapidly developing technology trends. However, the present structure has some limitation - one of these is the communication via client-server model - the users, devices, and applications using IoT services where all the connection/requirement is managed at IoT service providers. On the one hand, the IoT service providers (e.g., individual, organization) have different method to manage their devices, services, and users. Thus, the unique standard (i.e., communication method among the service providers and between client server) is still the challenge for the developers. On the other hand, Message Queuing Telemetry Protocol (MQTT) that is one of the most popular protocols in IoT deployments, has significant security and privacy issues by itself (e.g., authentication, authorization, as well as privacy problem). Therefore, this paper proposes UIP2SOP - an unique IoT network by using Single Sign-On (SSO) and message queue to improve the MQTT protocol's security problem. Besides, this model allows the organizations to provide the IoT services to connect into a single network but does not change the architecture of organization at all. The evaluation section proves the effectiveness of our proposed model. In particular, we consider the number of concurrent users publishing messages simultaneously in the two scenarios i) internal communication and ii) external communication. In addition, we evaluate recovery ability of system when occurred broken connection. Finally, to engage further reproducibility and improvement, we share a complete code solution is publicized on the GitHub repository.

Keywords—Internet of Things (IoT); MQTT; OAuth; Single Sign-On; Kafka; message queue

I. INTRODUCTION

The Internet of Thing (IoT) services/applications grown and play a vital role in our life such services/applications have applied in most fields such as smart cities, healthcare, supply chains, industry, and agriculture. In fact, by 2025, the whole world have approximately 75.44 billion IoT-connected devices [1], [2]. However, these devices is own by the different individuals or organizations, its characteristic (e.g., capacity, communication, computation ability) is totally different. Therefore, the ability to connect IoT service providers as well as

the security issues are still an open problem and need more considerable from the developers.

The most of IoT systems, currently, are a centrally designed according to a client-server architecture [3], [4], the individual/organization requires all devices and users to authenticate exchange information through one or more of the its servers. This architecture is suitable when the number of devices is limited - one advantage can easily aware is this model can easily setup and deploy in the real environment. However, when the system is extended with a millions of users/devices join the IoT network, these may create a billion transaction among them in the short time, we should consider the latency or even the deadlock issues.

For the IoT devices, the current architecture have the limited network connectivity, power, and processing capabilities [5], [6], so there is a specific requirement for separate machine-to-machine (M2M) protocols, unlike traditional communication protocols. The five most prominent protocols used for IoT devices (i.e., communicate among them and communicate with the upper layers) are Hypertext Transfer Protocol (HTTP), Constrained Application Protocol (CoAP), Extensible Messaging and Presence Protocol (XMPP), Advanced Message Queuing Protocol (AMQP), and Message Queuing Telemetry Protocol (MQTT) [7], [8]. For the communication requirements in limited networks (constrained networks), MQTT and CoAP are more reasonable to be used [9]. Besides, we found that the MQTT protocol has faster creation time, and transmission time of the packet is twice as fast as the CoAP protocol [10]. Furthermore, for developers of low bandwidth and memory devices, MQTT is the most preferred protocol [11]. Therefore, this paper applies the MQTT protocol to develop UIP2SOP platform.

Nevertheless, the current MQTT model has some drawback especially in security and privacy issues, namely data confidentiality, availability, integrity, and privacy [12]. This model only provides identity, authentication, and authorization for the security mechanism [13] but it is very simple. Lundgren et al. [14] indicate that a simple ruby script is used to subscribe

to topic # of any random MQTT server that is public on the Internet, and the obtained result reveals the data, including the device's GPS location, without any authentication, which is a severe security risk. In the MQTT protocol, the Client ID is unique, and MQTT Broker uses this Client ID to identify the client and its status. From this feature, an attacker can take advantage of performing a denial-of-service (DoS) attack. This is considered a risk in terms of the availability of the MQTT protocol. As the studies in [12], [15] show, if the attacker subscribes to topics with the client ID, the victim encounters denial of service status, and all information sent to the victim is forwarded to the attacker.

Regarding the authentication mechanism, MQTT supports authentication by username and password pairs, but the authentication mechanism is optional and not encrypted. The MQTT client authenticates itself by sending the username and password plaintext in the CONNECT package. Attacker attacks are made quickly by blocking packets [10]. Besides, the MQTT protocol allows any user to subscribe to a broadcast topic without any authentication, and anyone who has it can easily subscribe to any MQTT server available on the Internet [13]. According to a survey of article [16] about Shodan - the world's first IoT search engine for Internet-connected devices shows that there are 67,000 MQTT servers on the public Internet, with most of them without authentication. MQTT does support an authorization mechanism to access a specific topic based on the access list (ACL). This access list must be predefined in the MQTT broker config file¹ and we must restart the MQTT broker service to apply a new access-list configuration. This is inconvenient and difficult to expand, especially for systems with billions of devices, and these devices can only have the right to act for a specified period on specific topics. The problem of access control and authorization is a significant challenge [17].

The security problems in the MQTT protocol are also in the internet connection systems aspect. In particular, due to user behavior, the MQTT's security flaws are generally vulnerable to attack by the malicious users. Users often ignore this aspect, especially privacy, until the loss of critical data [18]. Statistics from [19] show that a significant proportion of the users are not fully aware of where their pieces of information are shared. Hence, with IoT systems having a huge number of users and devices (e.g., a millions), it is quite challenging to manage all user's behavior, especially when users among IoT systems exchange information with each other.

To address the risk of security and availability of MQTT, UIP2SOP manages users, things and channels (topic). This model allows the users to own, use, and exchange messages through ensuring precisely on which channels they are sharing information. To improve the Authentication and Authorization protocol of the MQTT protocol, we propose a combination of MQTT and OAuth protocol by adding a centralized authentication management system (Single Sign-On system). Finally, we use Kafka to build a message queue system that connects discrete IoT systems. To engage further reproducibility or improvement, we share the completely code solution which is publicized on the our Github².

The rest of the paper is organized as follows. We provide the background and the related work in the next two sections. Section 4 introduces architecture of UIP2SOP and its prototype system in Section 5. In Section 6, we discuss our evaluation in the different scenarios. Finally, we conclude the key points paper and discuss further work directions.

II. BACKGROUND

A. MQTT Protocol

MQTT (Message Queue Telemetry Transport) is a messaging protocol in a publish-subscribe model, using low bandwidth and high reliability. MQTT architecture consists of two main components: Broker and Client. MQTT Broker is the central server. It is the intersection point of all the connections coming from the client. The broker's main task is to receive messages from all clients and then forward them to a specific address. Clients are divided into two groups: publisher and subscriber. The publisher is the client that publishes messages on a specific topic. Subscribers are clients that subscribe to one or more topics to receive messages going to these topics.

B. OAuth Protocol and Single Sign-On

OAuth is an authentication mechanism that enables third-party applications to be authorized by the user to access the user's resources located on another application. OAuth version 2, an upgrade of OAuth version 1, is an authentication protocol that allows applications to share a portion of resources without authenticating via username and password as the traditional way. Thereby limiting the hassle of having to enter the username, password in too many places or register too many accounts for many applications that they cannot remember.

According to the OAuth document³, there are four basic concepts, namely, Resource owners, Resource server, Clients, and Authorization server.

- **Resource owners:** are users who have the ability to grant access, the owner of the resource that the application wants to get.
- **Resource server:** a place to store resources, capable of handling access requests to protected resources.
- **Clients:** are third-party applications that want to access the resource shared by the resource owner, and before accessing, the application needs to receive the user's authorization.
- **Authorization server:** authenticates, checks the information the user sent from there, grants access to the application by generating access tokens. Sometimes the same authorization server is the resource server.

A token is a random code generated by the Authorization server when a request comes from the client. There are two types of tokens, the access token, and the refresh token. The access token is a code used to authenticate access, allowing third-party applications to access user data. This token is sent by the client as a parameter during the request when it is necessary to access the resource in the Resource server. The

¹<https://mosquitto.org/man/mosquitto-conf-5.html>

²https://github.com/thanhlam2110/uiip2sop_platform

³<https://oauth.net/2/>

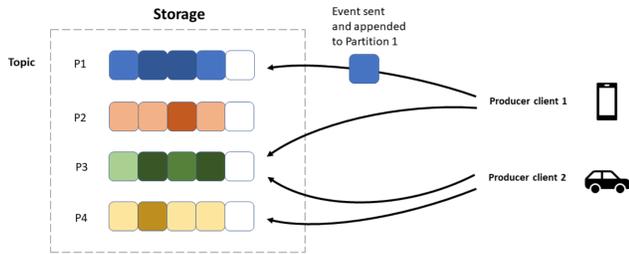


Fig. 1. The Activity Sample of Kafka [21]

access token has a reasonable time (30 minutes, 1 hour). When it expired, the client had to request the Authorization Server to get the new access token. The Authorization server also generates the refresh token simultaneously as the access token but with a different function. The refresh token is used to get the new access token when it expires, so the validity period is more extended than the access token.

Single Sign-On (SSO) is a mechanism that allows users to access multiple applications with just one authentication. SSO simplifies administration by managing user information on a single system instead of multiple separate authentication systems. It makes it easier to manage users when they join or leave an organization [20]. SSO supports many authentication methods such as OAuth, OpenID, SAML, and so on.

C. Kafka

Kafka is a distributed messaging system. Kafka is capable of transmitting a large number of messages in real-time. In case the kernel has not received the message, this message is still stored on the message queue and on the disk to ensure safety. Fig. 1 show a sample of Kafka [21].

Kafka includes the four components: producer, consumer, topic, and partition. Kafka producer is a client to publish messages to topics. Data is sent to the partition of the topic stored on the broker. Kafka consumers are clients that subscribe and receive messages from the topic. Group names identify consumers, whereas many consumers can subscribe to the same topic. Data is transmitted in Kafka by topic. Once it is necessary to transmit data for various applications, it can create many different topics. Partition is the data storage on a topic. Each topic can have one or more partitions. For each partition, the data is stored permanently and assigned an ID called offset. Besides, Kafka servers are also called a broker, and the zookeeper is a service to manage the brokers.

III. RELATED WORK

A. OAuth and MQTT

Paul Fremantle et al. [17] used OAuth to enable access control in the MQTT protocol. The paper results show that IoT clients can fully use OAuth token to authenticate with an MQTT broker. The article demonstrates how to deploy the Web Authorization Tool to create the access token and then embed it in the MQTT client. However, the article does not cover the control of communication channels, so when the properly authenticated MQTT client is able to subscribe to any topic on the MQTT broker, this creates the risk of data

disclosure. The paper presents the combined implementation of OAuth and MQTT for internal communication between MQTT broker and MQTT client in the same organization, but not the possibility of applying for inter-organization communication. Therefore, in our article, we implement a strict management mechanism for users, devices, and communication channels. We also present the mechanism of combining MQTT and OAuth protocols to authenticate users when communicating among organizations.

Benjamin Aziz et al. [22] investigated OAuth to manage the registration of users and IoT devices. These papers also introduce the concept of Personal Cloud Middleware (PCM) to perform internal communication between the device and a third-party application on behalf of the user. PCM is an MQTT broker that isolates and operates on a Docker or operating system. Each user has their PCM, and this can help limit data loss. However, Benjamin Aziz et al. also said that they do not have a mechanism for revoking PCM when users are no longer using IoT services, nor have they clearly stated the mechanism to ultimately connect PCMs to form a network for users of various organizations to communicate with each other.

B. Kafka and MQTT

A.S. Rozik et al. [21] found that the MQTT broker does not provide any buffering mechanism and cannot be extended. When large amounts of data come from a variety of sources, both of these features are essential. In the Sense Egypt IoT platform, A.S. Rozik et al. have used Kafka as an intermediary system to transport messages between the MQTT broker and the rest of the IoT system, which improves the overall performance of the system as well as provides easy scalability.

Moreover, in the previous studies [23], the authors presented Kafka Message Queue and MQTT broker's combined possibilities in Intelligent Transportation System. The deployment model demonstrates the ability to apply to bridge MQTT with Kafka for low latency and handle messages generated by millions of vehicles. They used MQTT Source Connector to move messages from MQTT topic to Kafka Topic and MQTT Sink connector to move messages from Kafka topic to MQTT topic as shown in Fig. 2.

Lam et al. [24] presented an architecture that combines MQTT broker and kafka message queue to connect different IoT service providers. This architecture allows individual service providers to communicate with each other easily without changing the existing architecture too much. In addition, Lam et al. [25] also evaluates power consumption, transfer speed, communication reliability, and security when using a combination of MQTT broker and kafka message queue. With Kafka's capabilities, we don't need to trade off transmission speed and reliability for power consumption (this is related to QoS-0 and QoS-2 levels).

In the implementation of the IoT Platform, we also adopt and extend this technique by building APIs that allow users to map their topics.

IV. UIP2SOP PLATFORM

The UIP2SOP Platform is a set of APIs combined with system architecture such as Single Sign-On system, Kafka

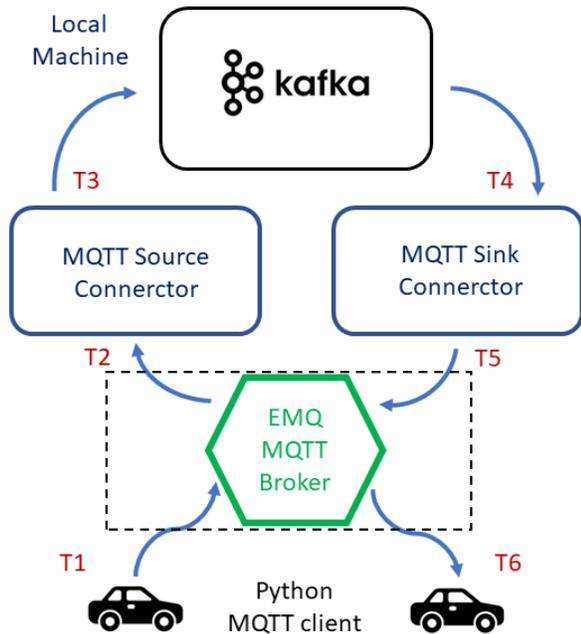


Fig. 2. Communication between MQTT Borker and Kafka [23].

Message Queue, and MQTT broker that provide the following capabilities:

- Authenticate information verifies the user info by granting OAuth access token and refresh token for the user.
- Management information exploits the tree model and the unlimited number of user levels created.
- Management of physical devices/ applications and communication channels participates in an IoT system.
- Sending and receiving messages locally defines a user, a thing, and a specific communication channel.
- Sending and receiving messages globally combines two various organizations through the Kafka message queue.

A. System Architecture

Fig. 3 presents an architectural proposal model of the IoT framework. Let's consider Organization A and Organization B are two separate organizations in the system. Each organization is a set of MQTT brokers that are interconnected to form a cluster MQTT broker. These MQTT brokers play two roles as follows: i) **Internal MQTT broker** is responsible for transporting messages communicating between users, IoT devices; ii) **External MQTT broker** is in charge of transporting messages communicating between two organizations. The UIP2SOP Platform system includes the Single Sign-On server and the Single Sign-On service's database cluster to perform the following tasks: i) authenticate user according to OAuth protocol; ii) manage user registration information, channels (public and local), and things. Finally, the UIP2SOP Platform

system uses the Kafka Message Queue to transport messages between two organizations' external brokers.

B. UIP2SOP Architecture

To meet the goals set out by the UIP2SOP Framework, we provide several definitions of the components involved in the system and the interaction of these components.

1) Users::

Users participate in an organization and use IoT services provided by that organization. They have two types (corresponding to the parameter field "usertype" in the database): a representation user and a normal user.

Each organization has only one representation user that created when an organization registers information of the organization with the UIP2SOP Platform. Representation users are not allow to send or receive messages through the MQTT broker or Kafka message queue. A representation user only creates an organization's public channel and the normal users use this channel to communicate with other organizations in the IoT network. Also, they manage the normal users of the organization.

The purpose in creating a representation user concept is to efficiently manage (e.g., send or receive) messages as well as the organization's (e.g., join or leave) the IoT network. All normal users have to send and receive public messages on the public channel that created and not allowed to create a public channel. Besides, the UIP2SOP Platform manager efficiently manages the entry and exit of an organization's IoT network via the organization's status (i.e., the `userstatus` parameter field in the database). When an organization leaves the IoT network or may be attacked, the UIP2SOP Platform administrator switches the `userstatus` from `ACTIVE` to `DISABLE`, which results in isolating the entire organization from the IoT network, and all normal users and devices within the organization cannot communicate with other organizations but can still communicate internally within the same organization.

Similarly, after the problem is resolved or wants to re-join the IoT network, through a representation user, the organization can request the manager of the UIP2SOP Platform to change his or her state to `ACTIVE`. By constructing a user hierarchy model tree with the child's `user_parent_id` value equal to the parent user's username, our UIP2SOP Platform allows the creation and management of multiple users' levels and is not restricted depending on the characteristics of the organization. This tree-modeled hierarchy of users makes the UIP2SOP platform more suitable for companies, especially when it comes to authorize a specific user. The user hierarchical management model is shown in Fig. 4.

A normal user registers to use the IoT services of a particular organization. They can create things, channels and assign things to channels to manage which things allowed to send and receive messages on a predefined channels. Each user has a unique `user_id` value, conforming to the UUID⁴ standard created by the API and managed by the UIP2SOP Platform (user is not aware of this value). When publishing or subscribing, the user must pass his access token obtained

⁴<https://tools.ietf.org/html/rfc4122>

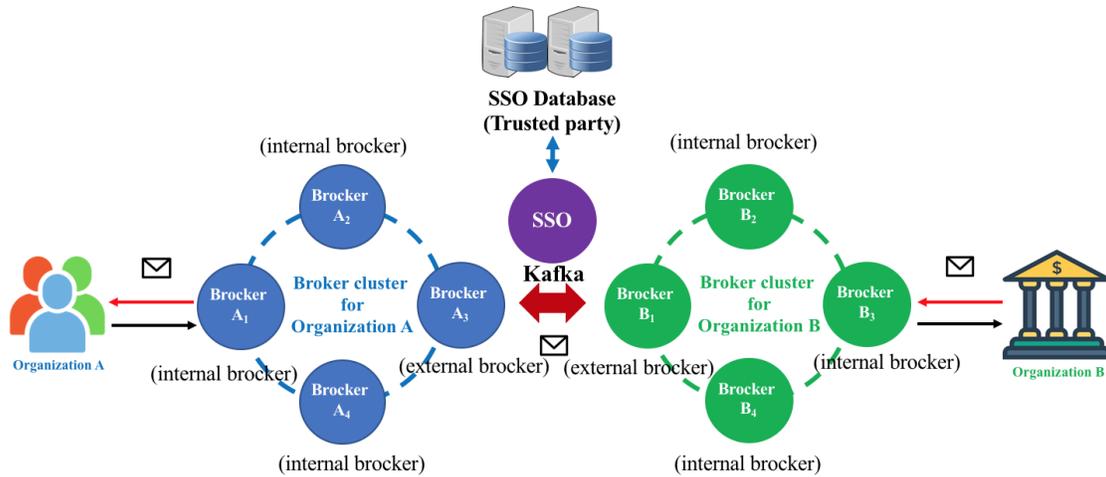


Fig. 3. UIP2SOP Platform.

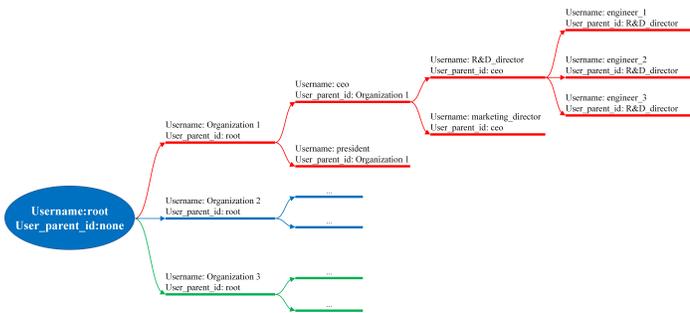


Fig. 4. Management user as a Model Tree.

from the Single Sign-On server, which contains the `user_id` information and is used as the MQTT protocol's `clientID` value. In this way, we have enhanced the authentication and authorization mechanisms for the user and minimize the risk of a denial-of-service attack when a hacker subscribes to a topic with the user's `clientID` affects the availability of the UIP2SOP Platform as we mentioned in Section I.

2) Things:

The things represent the physical devices info or the applications. To create the things, the owners need to call the API provided by the UIP2SOP Platform and pass in his/her valid OAuth token. If the user owns a valid token, the API generates the device's management information. The device's info includes two values `thing_id` and `thing_key`, corresponding to the device just created and return to the user. These two values are unique, created according to the UUID standard and used respectively with the username and password values used in the MQTT protocol. In practice, these two values are embedded in the physical device or application. Only the things that own the valid pair of `thing_id` and `thing_key`, can communicate with the MQTT broker since the other ones are not able to publish or subscribe directly to the MQTT broker. Instead, the things must go through the API layer. This layer API validates `thing_id` and `thing_key` sent by the things. Similarly, the user can create device management info to reduce the risk of denial-of-service attacks.

3) Channels and Assign Things to Channels:

In the UIP2SOP Platform, we propose, channels are the logical concept that governs topics where users and things publish and subscribe to messages. There are two types of channels: public channel and local channel.

The local channel is an MQTT topic that is created and managed by a normal user. Each normal user can create one or more local channels. Similarly, when creating a thing, the user who wants to create a channel must call the UIP2SOP Platform's API and pass in his valid OAuth token. If the user owns a valid token, the API generates the management information of the channel. This information includes value `channel_id` corresponding to the channel just created and return to the user. The `channel_id` value is unique, and according to the UUID standard, `channel_id` are returned to the user who created this channel. Each user only has his channel's informations and does not know the channel information of other users. Besides, the user has to assign things to this channels by calling the API and pass in `thing_id`, `channel_id` and his valid OAuth token. The purpose of this process is to only allow a thing with a valid `thing_id` and `thing_key` to publish and subscribe to messages on a predefined channel. From there, this help to avoid the client can subscribe to any topic. This enormously increases the authorization mechanism, which is a flexible way that the original MQTT protocol did not support. The assign things to channels mechanism also enhance security because, through our API, only things are mapped to the channel are allowed to publish and subscribe to messages on this channel.

The public channel is a Kafka topic are created by the representation user. Each organization has a unique public channel. All of the normal users of the organization have to communicate with another organization through the public channel.

4) Publish and Subscribe Message Locally: : After creating the things, the users (owners) have enough information including `channel_id`, `thing_id` and `thing_key` generated by the API layer of UIP2SOP Platform Proposal and returned to the user. Users embed three values `thing_id`, `thing_key` and their refresh token into the things (physical device or application). The process of embedding the above

information into a thing is out of this article's scope, so we are not able to elaborate on it here. The things with the necessary information embedded performs the API that provided by the UIP2SOP Platform Proposal to publish the message within the organization.

```
1 {
2   "token": "access token of publisher",
3   "thingid": "1960ff8f-ec57-4d81-b6d2-cb17e83016ad",
4   "thingkey": "3415e387-1b3b-49aa-8113-40799005f3bc",
5   "chanelid": "b7cd662c-7d15-4d6b-a8f0-d451e2f368ea",
6   "message": "Message local communication"
7 }
```

The information of token, thing_id, thing_key and channel_id are validated by the API services. If all information is correct, the message is sent to the MQTT broker; otherwise, the message is discarded. Similarly, for the Subscribe process, things also send information of token, thing_id, thing_key, channel_id to connect to the MQTT Broker through the API. If the information is not valid, the API are not allow the things to connect to the MQTT Broker.

5) *Publish and Subscribe message publicly*: : The process of publishing and subscribing to the message is described in Figure 5.

To support the communication among the users in the different organizations, our platform applies a public channel (or Kafka topic). In particular, we assign a local topic (MQTT topic) with a public channel (Kafka topic). For instance, organization A has a public channel with channel ID "95ce1a32-2136-417e-85b4-46b432f1c9ad". A user of organization A, called "user-a", wants to send a message to any user of another organization, called "user-b", he must go through this public channel and perform two steps as follows:

- Step 1: "user-a" creates a dedicated local channel to send messages to the public, assuming it is called "send-public-a". This channel is created via the API that creates a local channel, as shown in Section 4.2.3. In fact, the API uses channel_id, but for brevity, we cover the channel's name.
- Step 2: "user-a" uses the UIP2SOP Platform's API to assign the local channel just created above to the public channel. This process is equivalent create the MQTT Source Connector. After the mapping complete, "user-a" publishes the message to the "send-public-a" channel. Finally, the message is automatically routed to the public channel. The body structure to call API are as follows:

```
1 {
2   "name": "mqtt-source-for-user-a",
3   "config": {
4     "connector.class": "io.confluent.
5       connect.mqtt.MqttSourceConnector",
6     "tasks.max": "1",
7     "name": "mqtt-source-for-user-a",
```

```
  "mqtt.server.uri": "tcp://13.212.194.
    253:1883",
  "mqtt.topics": "send-public-a",
  "kafka.topic": "95ce1a32-2136-417e-85
    b4-46b432f1c9ad",
  "mqtt.clean.session.enabled": "true",
  "mqtt.connect.timeout.seconds": "30",
  "mqtt.keepalive.interval.seconds": "3
    0",
  "confluent.topic.bootstrap.servers":
    "localhost:9092",
  "confluent.topic.replication.factor":
    "1",
  "mqtt.qos": "0".
}
```

The UIP2SOP Platform builds the Kafka consumer service, which receive the message sent by "user-a", check the destination address (defined in body of public message), then forward it to the public channel of the "user-b". At that time, the message is on the public channel (Kafka topic) of "user-b". Therefore, to receive the message, the "user-b" must previously create a local channel (MQTT topic), called "receive-public-b" and map it to the public channel of "user-b". This is equivalent create an MQTT sink connector. The "user-b" uses an API provided by UIP2SOP Platform to create MQTT sink connector. The body structure creates MQTT sink connector is shown follow:

```
1 {
2   "name": "receive-public-b",
3   "config": {
4     "connector.class": "io.confluent
5       .connect.mqtt.
6       MqttSinkConnector",
7     "tasks.max": "1",
8     "mqtt.server.uri": "tcp://172.31
9       .46.150:1883",
10    "topics": "receive-public-b",
11    "mqtt.qos": "2",
12    "key.converter": "org.apache.
13      kafka.connect.storage.
14      StringConverter",
15    "value.converter": "org.apache.
16      kafka.connect.storage.
17      StringConverter",
18    "confluent.topic.bootstrap.servers
19      ": "localhost:9092",
20    "confluent.topic.replication.
21      factor": "1".
22  }
23 }
```

The body structure when using API public publishing is as follows:

```
1 {
2   "token": "access token of publisher"
3   ,
4   "source": "user-a",
```

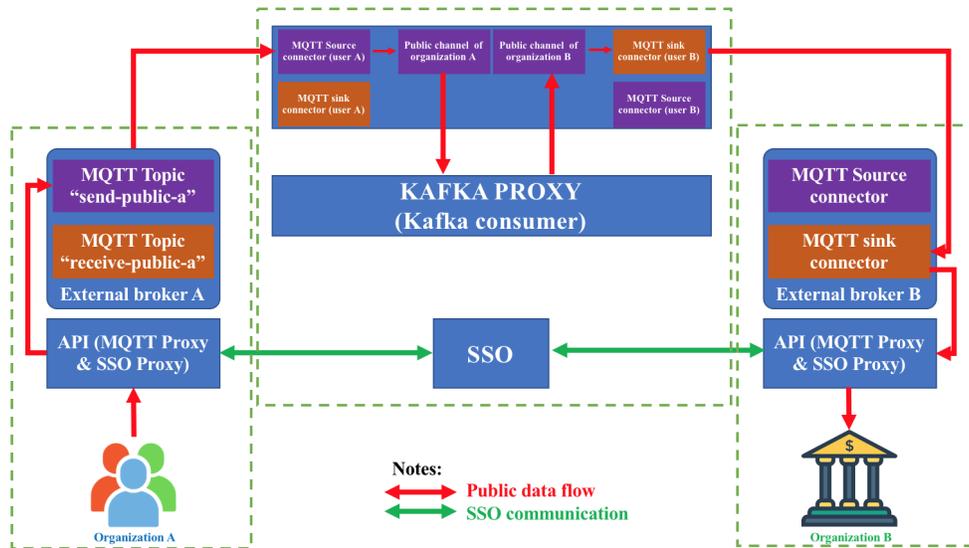


Fig. 5. Process Publish a Message to the Public.

```

4  ``destination": ``user-b",
5  ``message": ``Hello user-b"
6  }

```

V. IMPLEMENTATION

To demonstrate the practical implementation of the proposal platform, we have built a prototype of the system. We do this to answer the following research questions:

- Can UIP2SOP Platform be realizable?
- Are there any specific problems with the implementation?
- How does this performance compare to an existing system?

A. Database

As explained in Section V, the IoT Platform allows to manage the information of users, things, channels and implement assign things to channel. In practical implementation, we use MongoDB as a database management system belonging to NoSQL. To realize the model tree (select, update, delete), we have used the Aggregation technique provided by MongoDB⁵.

In the prototype system implementation we have seven data collections with the following roles:

Collection of users: store information of normal user and representation user. This collection of users are configured for the Single Sign-On system to read the information and authenticate with the parameters `username` and `password`. Our API layer (SSO proxy) also checks parameter `user_status`, if `user_status` is `ACTIVE` then API return Oauth access token and refresh token for user. We have designed the API that does not allow the users to change his or her `user_status`,

only his or her parent user can do this. Thanks to the user management design as tree model, the parent user can quickly change the status of its entire child user's status, which helps to quickly isolate all users when something goes wrong. The parameter `user_parent_id` is used to indicate the parent of the user. In our design the child user has information `user_parent_id` equal the `username` of its parent user.

Collection of things: store information of physical devices or applications managed by a normal user. The API layer (MQTT proxy) checks the parameter `thing_status`, if `thing_status` is `ACTIVE` then the things is allowed to connect to the MQTT broker. User can change the status of things which is managed by the himself. We have also designed an API that allows to change the status of the user resulting in changing the status of all of the user's things.

Collection of local channel: store local channel information (MQTT topic) managed by the normal user. API layer (MQTT proxy) checks the parameter `channel_status`, if `channel_status` is active then the thing is allowed to publish messages to the MQTT broker on this channel. User can change the status of a channel managed by the himself. We have also designed an API that allows to change the status of the user resulting in changing the status of all of the user's channels.

Collection of `things_map_channel`: stores information of thing that has been assigned to channels by a normal user to the channel. API layer (MQTT proxy) checks the parameter `map_status`, if `map_status` is active then the thing is allowed to publish the message to the MQTT broker on this channel. A user can change the status of `things_map_channel`, which is managed by himself. We have also designed an API that allows to change the status of the user resulting in changing the status of all of the user's `things_map_channel`.

Collection `MQTT_source_connector`: stores information MQTT source connector created by normal user. API layer (MQTT proxy) checks parameter `mqtt_source_status`,

⁵<https://docs.mongodb.com/manual/reference/operator/aggregation/graph-lookup/>

if `mqtt_source_status` is active then the message can be transferred from MQTT topic to Kafka topic. User can change the status of `MQTT_source_connector` managed by himself. We have also designed an API that allows to change the status of the user resulting in changing the status of all of the user's `MQTT_source_connector`.

Collection `MQTT_sink_connector`: stores the MQTT sink connector information generated by a normal user. API layer (MQTT proxy) check parameter `mqtt_sink_status`, if `mqtt_sink_status` is active then user will receive message from MQTT topic sent by Kafka topic. User can change the status of `MQTT_sink_connector` managed by himself. We have also designed an API that allows to change the status of the user resulting in changing the status of all of the user's `MQTT_sink_connector`.

B. Single Sign-On

In the prototype system, we use the open source CAS Apereo⁶ to provide the Single Sign-On service. CAS Apereo supports many protocols for implementing Single Sign-On services such as OAuth, SAML, OpenID, etc. The protocol that UIP2SOP Platform used to communicate with the Single Sign-On server is OAuth. In our implementation, the clients are not allowed interact directly with the CAS server but instead we provide the API for request OAuth token. For example, the API request token has the following body:

```
1 {
2   ``clientid``: ``exampleOauthClient``,
3   ``clientsecret``: ``
4     exampleOauthClientSecret``,
5   ``username``: ``thanhlam``,
6   ``password``: ``12345678``.
}
```

Fig. 6 describes an example of a the User with `DISABLE` status unable to request an `Access token`.

C. Mosquitto MQTT Broker

Mosquitto is an open source to implement MQTT broker that allows to transmit and receive data according to MQTT protocol. Mosquitto is also part of the Eclipse Foundation, the project iot.eclipse.org⁷. Mosquitto is very light and has the advantages of fast data transfer and processing speed, high stability.

D. Prototype System Deployment Model

We deploy prototype system as shown in Fig. 7.

The Prototype system we deployed on Amazon EC2 infrastructure consists of seven servers as shown in Table I.

Please add the following required packages to your document preamble:

In Table I, Organization 1 consists of two servers. The first server deploys the MQTT Broker 1a service for local communication. Second server, deploying MQTT Broker 1b

service for public communication. In addition, we deploy MQTT Proxy API service and SSO Proxy to provide APIs for users and devices to communicate with MQTT brokers and SSO through APIs in the first server. The implementation of organization 2 is similar to Organization 1.

Management Central is a single server deploying four services: Kafka, Kafka Proxy, Single Sign-On and database. Single Sign-On service creates access token and refresh token for users according to OAuth protocol. The Kafka service acts as the message queue to transport messages between Organization 1 and Organization 2. The Kafka Proxy service acts as the Kafka consumer to receive messages from Organization 1's public channel and forward it to the Organization's public channel 2. Database service to store information of users, things, channels.

VI. EVALUATION

A. Environment

After completing the deployment of the prototype system on the Amazon EC2 infrastructure, we conduct performance test scenarios of the IoT Framework. We check the number of concurrent users that can publish messages simultaneously for two cases of internal communication and external communication. We measure the time it takes to create a public channels. The test tool we use is the Apache Jmeter⁸.

Apache Jmeter is an open source, written 100% in Java, usable for performance testing on both static resources, dynamic resources and Web applications. It can be used to simulate a large number of virtual users, large requests on a server or a group of servers, a network or an object to test for load capacity or analyze the response time. Apache Jmeter provides the ability to test different applications, servers and protocols such as: Web-HTTP, HTTPS, SOAP / REST Webservice, FTP, LDAP, etc. Since our UIP2SOP Platform Protocol provides API as REST, Jmeter is a perfect fit for system load testing. Firstly, Jmeter makes requests and sends them to the server according to the predefined method, in this case it is REST. Then, it receives responses from the server, collects them and displays information in the report. Jmeter has many report parameters, but when performing system load test, we are mainly interested in two parameters: throughput and error, where the former is the number of requests processed by the server in a second and the latter is the percentage of the number of failed requests over the total number of requests. For our UIP2SOP Platform, the system crashed because the API layer is overload and can't handle request lead to crash API service.

B. Local Communication Test Cases

For the local communication test scenario, we compare the case where the user publishes the message through the MQTT Proxy API, i.e., the authentication process with Single Sign-On, checked by the MQTT Proxy and SSO Proxy, then passed to the local MQTT broker and finally received by the MQTT subscriber with case the user publishes the message directly to the MQTT broker. For two cases in the internal communication test scenario, we use Jmeter to call the API publish message locally then record the number of concurrent users and the

⁶<https://apereo.github.io/cas/6.3.x/index.html>

⁷<https://iot.eclipse.org/>

⁸<https://jmeter.apache.org/>

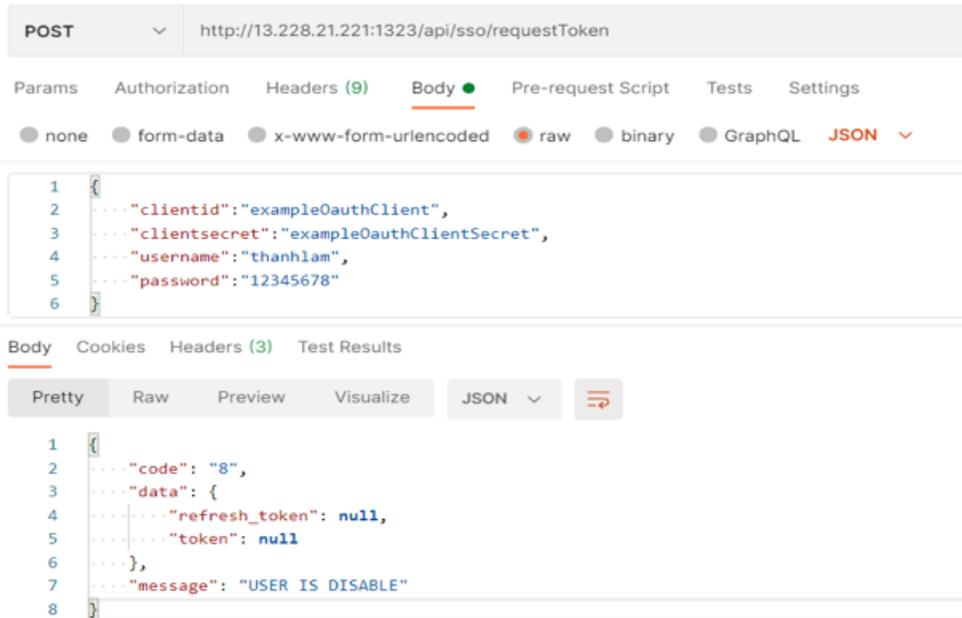


Fig. 6. User with DISABLE status is not Able to Request an Access Token.

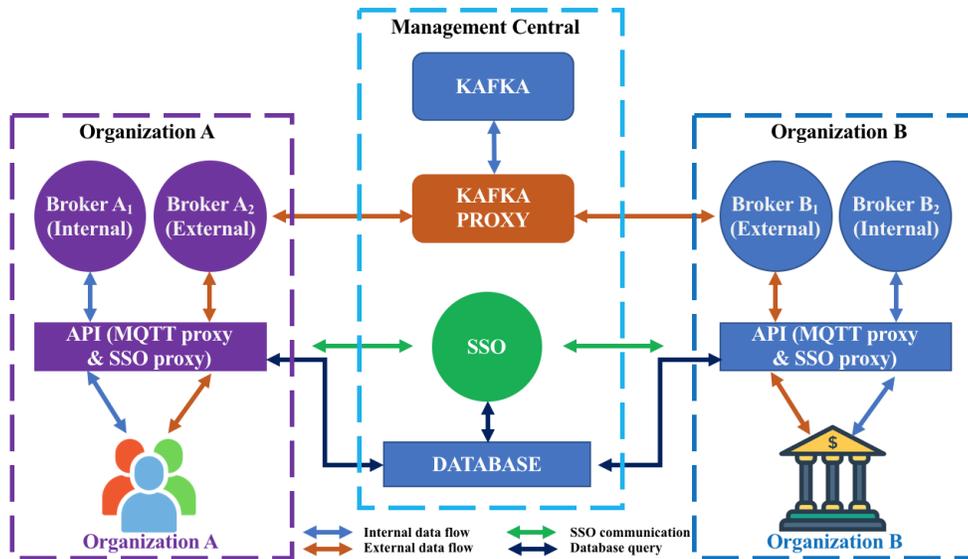


Fig. 7. Prototype System Deployment Model.

TABLE I. LIST OF INFRASTRUCTURE INFO IN THE PROTOTYPE

Zone	Server	Role	Server configuration
2*Organization A	Broker 1a	Deploy service MQTT broker for local communication Deploy service API (MQTT Proxy & SSO Proxy)	CPU 1 RAM 1 GB
	Broker 1b	Deploy service MQTT broker to serve public communication	CPU 1 RAM 1 GB
2*Organization B	Broker 2a	Deploy service MQTT broker for local communication Deploy service API (MQTT Proxy & SSO Proxy)	CPU 1 RAM 1 GB
	Broker 2b	Deploy service MQTT broker for public communication	CPU 1 RAM 1 GB
Management Central	Management Central	Deploy Single Sign-On service Deploy the Kafka service Deploy Kafka Proxy service Deploy the Database service	CPU 2 RAM 8 GB

number of requests processed per second. For MQTT Broker we use Mosquitto. The test model for internal communication is shown in the Fig. 8 and the test results are shown in Table II (via UIP2SOP Platform) and Table III (via MQTT Broker).

Through the results obtained, we found that with a server configuration of 1GB RAM and 1 CPU, the UIP2SOP Platform provides the ability to connect 450 concurrent users with a processing speed of 23.9 requests/s without error. Correspondingly, when publishing messages directly to the MQTT broker, Mosquitto can provide the ability to connect 550 concurrent users with a processing speed of 49.6 requests/s. This is reasonable and the result is completely acceptable because our UIP2SOP Platform has added the inspection and authentication mechanisms in Section IV.

C. Public Communication Test Cases

For the test scenario for public communication between organizations, we use Jmeter to call the API publish message publicly then record the number of concurrent users and the number of requests processed per second. Test model of communication between two organizations is shown in Fig. 7 and detail process is shown in Fig. 5. Test results are shown in Table IV.

According to aforementioned results, we found that with a server configuration of 1GB RAM and 1 CPU, the UIP2SOP Platform provides the ability to connect 170 concurrent users with a processing speed of 33 requests/s without error. To the best of our knowledge, we do not find any similar implementation models, so in this section we only record the measured parameters. This test result is reasonable and this result is completely acceptable because when sending messages between two organizations, the UIP2SOP Platform must perform more processing and authentication steps as we have presented in Section IV. The processing speed and the number of concurrent users can be improved by deploying the UIP2SOP Platform on a higher configuration server. In fact, APIs of the UIP2SOP Platform such as MQTT Proxy, SSO Proxy, SSO service and Kafka services should also be deployed as cluster to enhance processing capacity and high availability.

D. Broken Connection Test Cases

In addition, we also have taken test scenario with broken connection between publisher and subscriber. We compare number of receive messages in case with and without using UIP2SOP platform when occurred broken connection. The test model is shown in the Fig. 9.

The test result show that, in case without using UIP2SOP platform, subscriber only receive one message - the newest message that publisher send when occurred broken connection. This is the retain function of MQTT protocol. When we enable retain flag, MQTT broker is ability to keep only newest message that publish by publisher. This message is received by subscriber after it reconnects to MQTT broker⁹. However, when using UIP2SOP platform, subscriber can receive all message that published by publisher. This is ability of kafka message queue, therefore the system is guaranteed lost data.

⁹<http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html>

E. Future Work

To develop a larger scenario and increase the number of devices/users authorized quickly, other security issues such as security, privacy, availability for objects are still the challenges. For the security aspect, further works will be deployed in different scenarios like healthcare environment [26], [27], [28], cash on delivery [29], [30]. For the privacy aspect, we will exploit attribute-based access control (ABAC) [31], [32] to manage the authorization process of the IoT Platform via the dynamic policy approach [33], [34], [35]. Finally, we will apply the blockchain benefit to improve the availability issues [36], [37], [38].

VII. CONCLUSION

The UIP2SOP Platform provides significant improvements in the security capabilities of the MQTT protocol by combining OAuth protocol (Single Sign-On), model of management user, things and channels. The procedure of assigning things to channels, strict management of local and public communication channels helps to minimize careless behavior of users when sharing data. The Kafka message queue system not only easily connects among the organizations providing IoT services to the IoT network without changing too much of the available IoT system architecture of each organization but also reduce lost data when occurred broken connection.

REFERENCES

- [1] T. Alam, "A reliable communication framework and its use in internet of things (iot)," *CSEIT1835111—Received*, vol. 10, pp. 450–456, 2018.
- [2] V. Morfino and S. Rampone, "Towards near-real-time intrusion detection for iot devices using supervised learning and apache spark," *Electronics*, vol. 9, no. 3, p. 444, 2020.
- [3] H. F. Atlam, A. Alenezi, M. O. Alassafi, and G. Wills, "Blockchain with internet of things: Benefits, challenges, and future directions," *International Journal of Intelligent Systems and Applications*, vol. 10, no. 6, pp. 40–48, 2018.
- [4] O. Novo, "Blockchain meets iot: An architecture for scalable access management in iot," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1184–1195, 2018.
- [5] V. Karagiannis, P. Chatzimisios, F. Vazquez-Gallego, and J. Alonso-Zarate, "A survey on application layer protocols for the internet of things," *Transaction on IoT and Cloud computing*, vol. 3, no. 1, pp. 11–17, 2015.
- [6] D. Weissman and A. Jayasumana, "Integrating iot monitoring for security operation center," in *2020 Global Internet of Things Summit (GIoTS)*. IEEE, 2020, pp. 1–6.
- [7] A. Niruntasukrat, C. Issariyapat, P. Pongpaibool, K. Meesublak, P. Aiumsupucgul, and A. Panya, "Authorization mechanism for mqtt-based internet of things," in *2016 IEEE International Conference on Communications Workshops (ICC)*. IEEE, 2016, pp. 290–295.
- [8] B. Mishra and A. Kertesz, "The use of mqtt in m2m and iot systems: A survey," *IEEE Access*, vol. 8, pp. 201 071–201 086, 2020.
- [9] S. P. Jaikar and K. R. Iyer, "A survey of messaging protocols for iot systems," *International Journal of Advanced in Management, Technology and Engineering Sciences*, vol. 8, no. II, pp. 510–514, 2018.
- [10] B. H. Çorak, F. Y. Okay, M. Güzel, Ş. Murt, and S. Ozdemir, "Comparative analysis of iot communication protocols," in *2018 International symposium on networks, computers and communications (ISNCC)*. IEEE, 2018, pp. 1–6.
- [11] G. C. Hillar, *MQTT Essentials-A lightweight IoT protocol*. Packt Publishing Ltd, 2017.
- [12] J. J. Anthraper and J. Kotak, "Security, privacy and forensic concern of mqtt protocol," in *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM)*, Amity University Rajasthan, Jaipur-India, 2019.

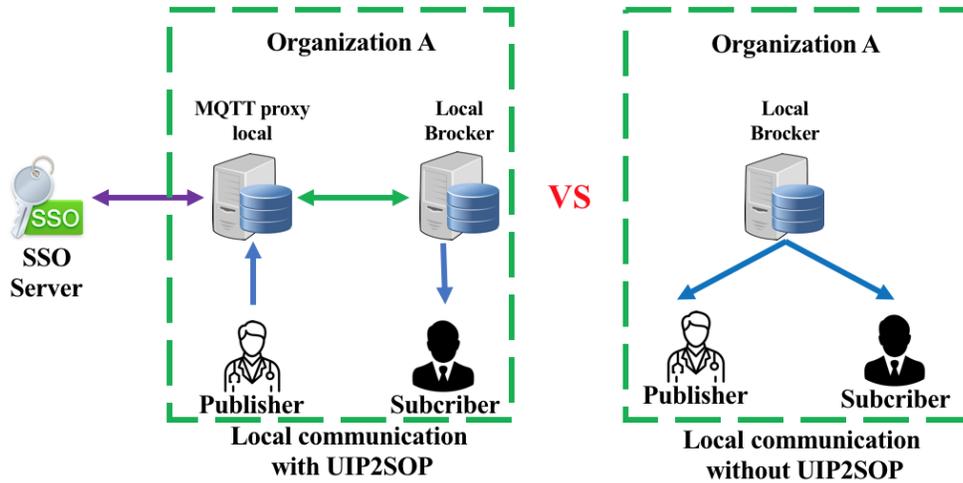


Fig. 8. Test Scenario for the Performance of Local Communication.

TABLE II. TEST RESULTS OF INTERNAL COMMUNICATION VIA UIP2SOP PLATFORM

	50 CCU	150 CCU	250 CCU	350 CCU	450 CCU	550 CCU	650 CCU
Through put (request/s)	14	24.7	26.3	25.5	23.9	21.6	-
Error	0%	0%	0%	0%	0%	5.2%	-

TABLE III. TEST RESULTS OF INTERNAL COMMUNICATION VIA THE MQTT BROKER

	50 CCU	150 CCU	250 CCU	350 CCU	450 CCU	550 CCU	650 CCU
Through put (request/s)	38	48.3	49.1	50.6	50.1	49.6	45.4
Error	0%	0%	0%	0%	0%	0%	28.3%

TABLE IV. TEST RESULTS OF EXTERNAL COMMUNICATION

CCU	25	50	75	100	125	150	175	200
Through put (request/s)	11	20.5	26.8	32.3	34.6	33.9	33	32.1
Error (%)	0	0	0	0	0	0	0	15.3

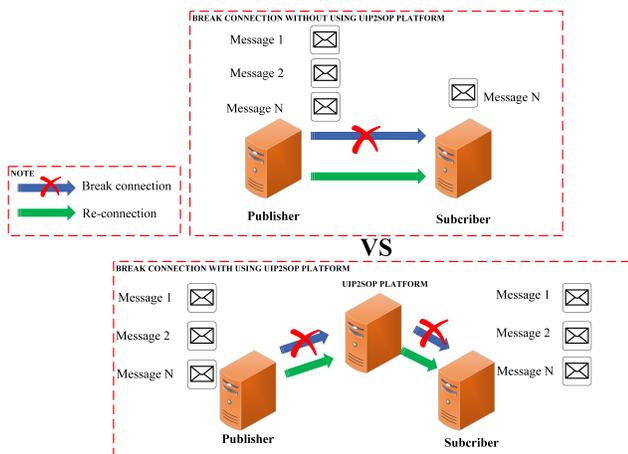


Fig. 9. Number of Receive Messages when System Recover after Broken Connection Issue.

[13] D. Mendez Mena, I. Papanagioutou, and B. Yang, "Internet of things: Survey on security," *Information Security Journal: A Global Perspective*, vol. 27, no. 3, pp. 162–182, 2018.

[14] L. Lundgren, "Light-weight protocol! serious equipment! critical implications!" *Defcon 24*, 2016.

[15] S. Wang, K. Gomez, K. Sithamparamanathan, M. R. Asghar, G. Russello, and P. Zanna, "Mitigating ddos attacks in sdn-based iot networks leveraging secure control and data plane algorithm," *Applied Sciences*, vol. 11, no. 3, p. 929, 2021.

[16] N. Zaidi, H. Kaushik, D. Bablani, R. Bansal, and P. Kumar, "A study of exposure of iot devices in india: Using shodan search engine," in *Information Systems Design and Intelligent Applications*. Springer, 2018, pp. 1044–1053.

[17] P. Fremantle, B. Aziz, J. Kopecký, and P. Scott, "Federated identity and access management for the internet of things," in *2014 International Workshop on Secure Internet of Things*. IEEE, 2014, pp. 10–17.

[18] L. Tawalbeh, F. Muheidat, M. Tawalbeh, M. Quwaider *et al.*, "Iot privacy and security: Challenges and solutions," *Applied Sciences*, vol. 10, no. 12, p. 4102, 2020.

[19] A. Subahi and G. Theodorakopoulos, "Detecting iot user behavior and sensitive information in encrypted iot-app traffic," *Sensors*, vol. 19, no. 21, p. 4777, 2019.

[20] V. Radha and D. H. Reddy, "A survey on single sign-on techniques," *Procedia Technology*, vol. 4, pp. 134–139, 2012.

[21] A. Rozik, A. Tolba, and M. El-Dosuky, "Design and implementation of the sense egypt platform for real-time analysis of iot data streams," *Advances in Internet of Things*, vol. 6, no. 4, pp. 65–91, 2016.

[22] P. Fremantle and B. Aziz, "Oauthing: privacy-enhancing federation for the internet of things," in *2016 Cloudification of the Internet of Things (CIoT)*. IEEE, 2016, pp. 1–6.

- [23] Á. Hugo, B. Morin, and K. Svantorp, "Bridging mqtt and kafka to support c-its: a feasibility study," in *2020 21st IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 2020, pp. 371–376.
- [24] L. N. T. Thanh *et al.*, "Toward a unique iot network via single sign-on protocol and message queue," in *International Conference on Computer Information Systems and Industrial Management*. Springer, 2021.
- [25] —, "Toward a security iot platform with high rate transmission and low energy consumption," in *International Conference on Computational Science and its Applications*. Springer, 2021.
- [26] H. X. Son and E. Chen, "Towards a fine-grained access control mechanism for privacy protection and policy conflict resolution," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019.
- [27] N. Duong-Trung, H. X. Son, H. T. Le, and T. T. Phan, "Smart care: Integrating blockchain technology into the design of patient-centered healthcare systems," in *Proceedings of the 2020 4th International Conference on Cryptography, Security and Privacy*, ser. ICCSP 2020. New York, NY, USA: Association for Computing Machinery, 2020, p. 105–109.
- [28] —, "On components of a patient-centered healthcare system using smart contract," in *Proceedings of the 2020 4th International Conference on Cryptography, Security and Privacy*. New York, NY, USA: Association for Computing Machinery, 2020, p. 31–35.
- [29] H. T. Le, N. T. T. Le, N. N. Phien, and N. Duong-Trung, "Introducing multi shippers mechanism for decentralized cash on delivery system," *money*, vol. 10, no. 6, 2019.
- [30] N. T. T. Le, Q. N. Nguyen, N. N. Phien, N. Duong-Trung, T. T. Huynh, T. P. Nguyen, and H. X. Son, "Assuring non-fraudulent transactions in cash on delivery by introducing double smart contracts," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, pp. 677–684, 2019.
- [31] N. M. Hoang and H. X. Son, "A dynamic solution for fine-grained policy conflict resolution," in *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, 2019, pp. 116–120.
- [32] H. X. Son and N. M. Hoang, "A novel attribute-based access control system for fine-grained privacy protection," in *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, 2019, pp. 76–80.
- [33] S. H. Xuan, L. K. Tran, T. K. Dang, and Y. N. Pham, "Rew-xac: an approach to rewriting request for elastic abac enforcement with dynamic policies," in *2016 International Conference on Advanced Computing and Applications (ACOMP)*. IEEE, 2016, pp. 25–31.
- [34] Q. N. T. Thi, T. K. Dang, H. L. Van, and H. X. Son, "Using json to specify privacy preserving-enabled attribute-based access control policies," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*. Springer, 2017, pp. 561–570.
- [35] H. X. Son, T. K. Dang, and F. Massacci, "Rew-smt: a new approach for rewriting xacml request with dynamic big data security policies," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*. Springer, 2017, pp. 501–515.
- [36] X. S. Ha, H. T. Le, N. Metoui, and N. Duong-Trung, "Dem-cod: Novel access-control-based cash on delivery mechanism for decentralized marketplace," in *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE, 2020, pp. 71–78.
- [37] X. S. Ha, T. H. Le, T. T. Phan, H. H. D. Nguyen, H. K. Vo, and N. Duong-Trung, "Scrutinizing trust and transparency in cash on delivery systems," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*. Springer, 2020, pp. 214–227.
- [38] H. X. Son, T. H. Le, N. T. T. Quynh, H. N. D. Huy, N. Duong-Trung, and H. H. Luong, "Toward a blockchain-based technology in dealing with emergencies in patient-centered healthcare systems," in *International Conference on Mobile, Secure, and Programmable Networking*. Springer, 2020, pp. 44–56.

A Hybrid Encryption Solution to Improve Cloud Computing Security using Symmetric and Asymmetric Cryptography Algorithms

Hossein Abroshan 

Department of Business Informatics and Operations Management, Ghent University
Ghent, Belgium

Abstract—Ensuring the security of cloud computing is one of the most critical challenges facing the cloud services sector. Dealing with data in a cloud environment, which uses shared resources, and providing reliable and secure cloud services, requires a robust encryption solution with no or low negative impact on performance. Thus, this study proposes an effective cryptography solution to improve security in cloud computing with a very low impact on performance. A complex cryptography algorithm is not helpful in cloud computing as computing speed is essential in this environment. Therefore, this solution uses an improved Blowfish algorithm in combination with an elliptic-curve-based algorithm. Blowfish will encrypt the data, and the elliptic curve algorithm will encrypt its key, which will increase security and performance. Moreover, a digital signature technique is used to ensure data integrity. The solution is evaluated, and the results show improvements in throughput, execution time, and memory consumption parameters.

Keywords—Cloud computing; security; cryptography; digital signature

I. INTRODUCTION

Many organisations and individuals use cloud computing services such as PAAS (Platform as a Service), IAAS (Infrastructure as a Service), and SAAS (Software as a Service). Cloud computing is a location-independent environment that shares calculations and resources to provide high-performance services [1]. Although cloud computing makes our lives easier, it comes with security challenges and the threat of cyberattacks such as the exploitation of authentication, sniffing, spoofing, resource manipulation, etc. [2]. These security challenges become more critical in cloud computing as resources and services are shared in an open environment between and within networks. Moreover, the user is storing data in third-party locations remotely, so its security is essential [3, 4]. Therefore, the users' data should be encrypted using cryptography algorithms to protect them against unauthorised accesses and to maintain the confidentiality and integrity of the data. Several cryptographic techniques have been used to protect data in different deployment models of cloud computing: public, private, and hybrid models and different cloud services models [5]. When cloud service providers and governments provide several concurrent services for large volumes of data (i.e., Big Data), security becomes even more critical, and guaranteeing data protection, especially its integrity, becomes a difficult task [6]. Thus, it is necessary to develop new techniques to improve

security with low impact on performance and high execution time [7]. However, as cryptographic techniques are usually complex and have an overload on the processes and cloud environment (e.g., data in transit), we need better encryption solutions with the lowest possible impact on performance, hence less negative impact on the services, while providing high security and data protection. Therefore, this study proposes a hybrid cryptography solution based on an improved Blowfish algorithm.

The paper is structured as follows: the second section explains the research background and is followed by the new solution, which is explained in the third section. Finally, the proposed solution is evaluated in the fourth section, and the paper culminates with a conclusion based on the discussion thus far.

II. RESEARCH BACKGROUND

Because of the importance of security in cloud computing, many studies have been conducted on this topic. For instance, AbdElminaam [8] proposed a hybrid cryptography technique using AES (Advanced Encryption Standard) and Blowfish. They suggest that using a combination of symmetric and asymmetric techniques will provide a high cloud security. Their results show a lower encryption time and better throughput than other techniques such as using combination of AES and RSA (Rivest–Shamir–Adleman) [9]. Thabit, et al. [10] proposed a lightweight cryptography technique for cloud computing security that uses a 128-bit block cipher and key to encrypt the data. They used Feistel and substitution permutation architectural methods to make the encryption more complex. The proposed algorithm has “flexibility in the length of the secret key and the number of turns”, and the results show a low encryption time. These solutions can improve security while decreasing the encryption time, but the problem with the proposed techniques is that they need to exchange the encryption key which can make confidentiality, and as a result the whole could computing security, risky. Another study [11] utilized different cryptography techniques to improve the cloud computing security. For this purpose, an AES based encryption and asynchronous key system is consolidate in their model. They also used an Elliptic curve cryptography to secure the communications between client and cloud storage system. The proposed solution randomly generates keys based on chaotic cryptography. However, it seems that the proposed technique did not consider the time needed for split/combine of data.

There are, however, studies that use other new techniques to improve the security of cloud computing. For example, Esposito, et al. [12] suggested using blockchain for security of cloud computing. In the proposed blockchain-based ecosystem for new data a new block is instantiated and distributed to all peers in the network, and the system will insert it in the chain when majority of the peers approve the new block. This solution can provide a good level of cloud computing security as blockchain is secure by design. Although the blockchain-based solution has many benefits such as high level of data integrity as, for instance, when data stored in the chain it is not possible to alter it, there are some challenges with this solution. For example, it might breach data protection regulations (such as GDPR¹) as personal sensitive data shall be stored for the “shortest time possible” and a time limit to erase the data must be established [13, 14]. Moreover, as it needs to be approved by more than half of the peers, it will not be possible to undo the unwanted changes or fix mistakes. Another issue with the proposed solution is that blockchain is designed to store transactional data which are small, so the solution does not perfectly work for securing all cloud computing data as many of them are large data such as images. Another study evaluated the performance and effectiveness (e.g. execution time and memory consumption) of different symmetric cryptographic algorithms [15]. There results show that Blowfish and DES (Data Encryption Standard) are better in terms of required encryption/decryption time and memory. Some studies reviewed the most important and efficient cryptographic method for cloud computing security [5, 16]. The results of the previous studies show a need for new cloud computing techniques that can improve security and at the same time have no or as less as possible negative effect on performance, to make cloud computing more secure with no effect on its services.

III. PROPOSED SOLUTION

The proposed solution is an effective technique that uses different cryptographic based data protection algorithms. The solution used symmetric and asymmetric encryption combined with an improved digital signature using an MD5² hashing function to ensure data integrity. It is necessary to consider the algorithm’s encryption speed, and the balance between performance and speed. Asymmetric key generation algorithms (i.e., public key) need more generation time than symmetric algorithms such as Blowfish, so they have a lower speed. Thus, we use an improved Blowfish (symmetric key generation) algorithm for encrypting the data in this solution. Fig. 1 illustrates a high-level diagram of the solution.

First, the hash code of the original data will be generated using MD5, which will be used to verify data integrity. In the second step, an Elliptic Curve (EC) algorithm will be used for digital signature generation and securing the MD5 code and private key. We need to have a high level of security and low length of the key and hash code, and as a result, the execution time of the EC will be decreased. Finally, the original data will be encrypted using an improved Blowfish algorithm, a

symmetric algorithm with low encryption/decryption time. In the following sections, each step of the solution is described in detail.

A. Digital Signature

This study used a digital signature together with a hashing function to assure data integrity. However, as the initial digital signature model is vulnerable to cyberattacks (i.e., the cyber attacker can create a fake digital signature by manipulating the digital signature verification process), in this solution, we first hash the message and subsequently sign the hashed message. Thus, creating a fake signature does not work for the attacker as the signature does not match the hashed message’s outcome, and the attacker cannot manipulate or damage the content of the message. Therefore, using a digital signature can be considered a robust tool for securing cloud computing [17]. Fig. 2 shows the flowchart of using a digital signature and hash code in the proposed solution.

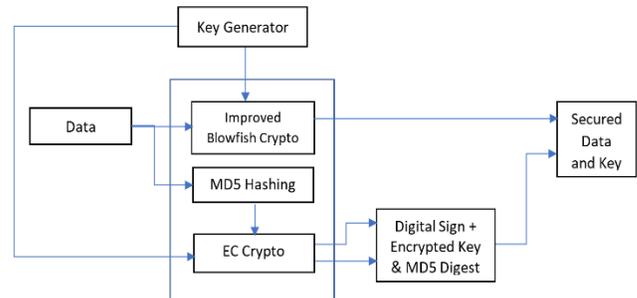


Fig. 1. High-Level Diagram of the Proposed Solution.

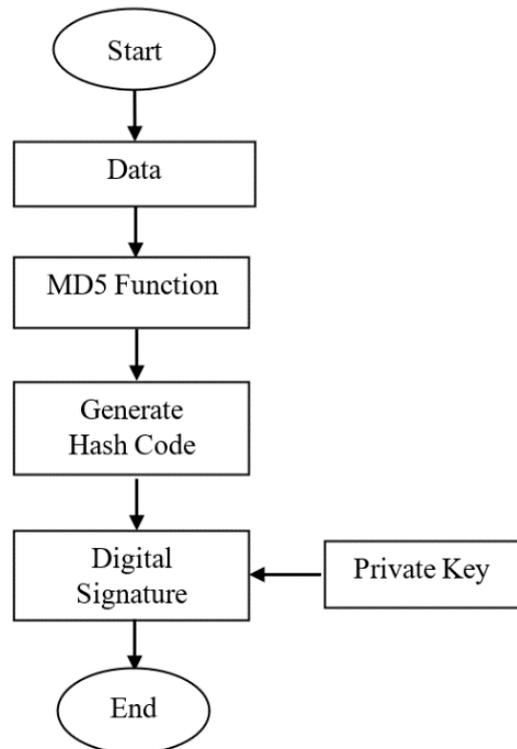


Fig. 2. Hash Code and Digital Signature Flowchart.

¹ General Data Protection Regulation is European personal data protection regulation.

² Message-digest algorithm

B. Data Encryption

This study aims to propose a secure cryptography technique with low execution time and high throughput, and we used an improved Blowfish algorithm in the core of the data encryption. This algorithm uses a variable key length from 32 up to 448-bits and used several subkeys [18]. The subkeys must be pre-computed before encryption/decryption of the data. This algorithm is much faster than other algorithms and requires less memory. The blowfish algorithm uses four 256 S-BOX containing a total of 1024 32-bit entries. The first byte from the first 32-bit entry will be used to find an entry in the first S-BOX, the second byte to find an entry in the second S-BOX and the same for all other entries (Modified blowfish algorithm). The same goes for the decryption procedure, starting with the cyphertext as an input, except that the subkeys will be used in a reverse way.

The F-function is the most time-consuming part of the encryption as in all rounds of the Blowfish algorithm, the F-function does the main calculation, including Adder and Rotation, in a modular format. Thus, in this study, we reduce the Blowfish's execution time by changing the F-function's module. Fig. 3 shows the overall process of the F-function module in a standard Blowfish.

To improve the algorithm, we decrease the complexity of execution time. Equation (1) calculates the value of F-function in a standard Blowfish.

$$F(X_L) = ((S_{1,a} + S_{2,b} \bmod 2^{32}) \text{ XOR } S_{3,c}) + S_{4,d} \bmod 2^{32} \quad (1)$$

We can change the F-function to equation (2), without reducing the security of the Blowfish algorithm.

$$F'(X_L) = (S_{1,a} + S_{2,b} \bmod 2^{32}) (S_{3,c} + S_{4,d} \bmod 2^{32}) \quad (2)$$

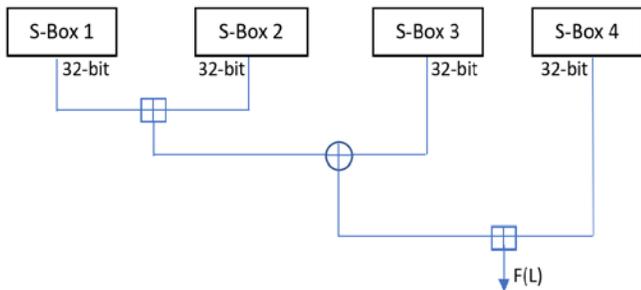


Fig. 3. F-Function Module in Blowfish.

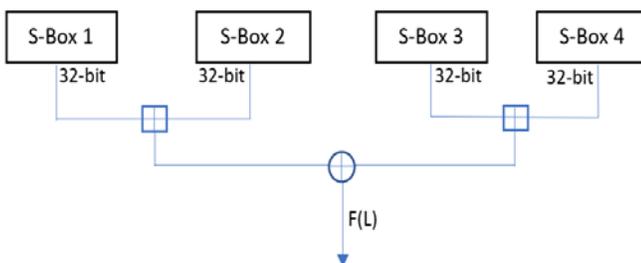


Fig. 4. The Improved F-Function Module.

This change will allow us to do both the $(S_{1,a} + S_{2,b} \bmod 2^{32})$ and $(S_{3,c} + S_{4,d} \bmod 2^{32})$ addition operations in parallel. This parallel operation will reduce the execution time as it will reduce the needed time for two operations to one operation. As Blowfish has 16 rounds, this change should improve the execution time 16 times in each encryption and decryption. Furthermore, as the security of Blowfish is related to its keys, this change will not negatively impact the algorithm's security. Fig. 4 illustrates the mentioned change.

Pseudocode:

Pseudocode for the Blowfish algorithm based on the improved F-function

```

Element, x.
Divide x into two 32-bit halves: xL, xR
For i = 1 to 16:
    xL = xL XOR Pi
    xR = F(xL) XOR xR
    Swap xL and xR
Next i
Swap xL and xR
xR = xR XOR P17
xL = xL XOR P18
Recombine xL and xR
Function F
Divide xL into four eight-bit quarters: a, b, c and d
F(xL) = (S1,a + S2,b mod 232) XOR (S3,c + S4,d mod 232)
    
```

First, the 64-bit entry splits into two left (L) and right (R) 32-bit. The next step is a XOR operation on the first 32-bit block (L). In the third step, the calculated 32-bit data will be transferred to F-function to be XOR'd with the other 32-bit block (R). Then, the L and R will be swapped to be used in the next Blowfish rounds. The decryption operation is the same except that the P1, P2, P18 will be used in a reverse way.

C. Securing the Symmetric-Key

Symmetric-key based algorithms, such as Blowfish, need to transfer the private key, which can be risky, as attackers might steal the key (e.g., man in the middle attacks). Thus, we use an elliptic-curve-based asymmetric cryptography algorithm to secure the private key and hash code in this solution. In this method, a 164-bit key will be used to provide a higher performance than other solutions [19]. The elliptic-curve-based cryptography algorithms have a high level of security and need low memory and bandwidth [20]. One of the benefits of this encryption is cryptography over finite fields. So, the assumption is that the p is a prime number, and the finite field contains a set of numbers less than p . Following is the explanation of the two-dimensional Elliptic curve (E).

If the $p > 3$ is an odd prime, and $a, \in Fp$ and the $4a^3 + 27b^2 \neq 0 \bmod p$, then the Elliptic curve $E(Fp)$ is equal to the equation (3).

$$y^2 = x^3 + ax + b \quad (3)$$

Using element Q , which can create a group in the $E(F_p)$ set, and considering $p=(x_1,y_1)$ and $\{p, Q\} \in E(F_p)$

For point adding $P + Q = (x_3, y_3)$ and $P \neq Q$

$$x_3 = \lambda^2 - x_1 - x_2$$

$$y_3 = \lambda(x_1 - x_3) - y_1$$

$$\lambda = (y_2 - y_1)/(x_2 - x_1)$$

For point doubling $P + P = 2P(x_3, y_3)$

$$x_3 = \lambda^2 - 2x_1$$

$$y_3 = \lambda(x_1 - x_3) - y_1$$

$$\lambda = (3x_1^2 + a)/2y_1$$

Adding two different points on the elliptic curve needs six additions, one square, two multiplications, and one inverse operation in F_p . Also, point doubling on the elliptic curve needs eight additions, two square, two multiplications, and one inverse operation. For cryptography using the elliptic curve, we first choose a random k number in a range that belongs to the field and consider it as the private key. Then, the public key Q will be calculated by $Q=kP$. In this situation, P is a point on the elliptic curve. The elliptic-curve cryptography is based on the difficulty of the discrete logarithm problem, so calculating k from the P and Q points has computational complexity, which is one of the main benefits of elliptic-curve cryptography. This method uses the following scalar multiplication.

$$Q = k.P = P + P + \dots + P \tag{4}$$

Because the operations in this method are based on the field, using Modular Multiplication can help to improve the performance of elliptic-curve cryptography. Fig. 5 shows the flowchart of this method, and the following steps describe the method in detail.

Step 1. Before the encryption, a hash code (using MD5) for the original data will be generated. This code will be used to improve the efficiency of the digital signature and data integrity verification.

Step 2. The hash code will be encrypted using the private key, and the digital signature of the data will be issued.

Step 3. The Blowfish's private key will be encrypted using elliptic-curve cryptography.

Step 4. The data will be encrypted by the symmetric cryptography algorithm (Blowfish). We use the symmetric key to encrypt the original data. Then, the encrypted data in this step and the previous steps will be transferred.

Step 5. The receiver decrypts the received data using a reverse process and the private key.

Step 6. The original data will be decrypted using Blowfish's private key, and the verification and validation process will be run using the hash function (digital signature).

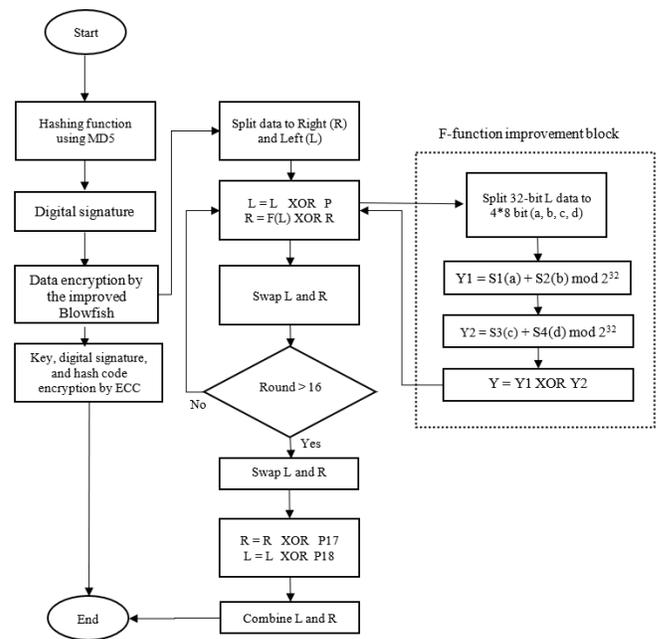


Fig. 5. The Proposed Solution's Flowchart.

IV. EVALUATION AND DISCUSSION

We implemented the proposed solution on Eclipse and by using Java development kit (JDK) version 7. The cryptography functions are available in two Java Cryptography Architecture (JCA) and Java Cryptography Extension (JCE) libraries in the JDK. The JCA provides most of the fundamental cryptography requirements, while the JCE can be used for advanced cryptography operations. The hardware used in the solution evaluation had an Intel Core2 Duo 2.5 GHz processor. All the evaluations were run in Windows 7.

Different parameters such as memory use, throughput, and execution time of the proposed solution were compared with AES, 3DES, DES, and RSA algorithms [21]. We initially evaluated the solution with a low size data (50 KB) to analyse its performance when the data is small and then evaluated the solution with larger data (i.e., 1024 KB and 2048 KB).

A. Evaluation Results – Throughput and Execution Time

The first evaluation is based on 50KB data. As fig. 6 illustrates, the execution time of the proposed technique (Hybrid) reduced 800, 550, 300 and 400 milliseconds in comparison with RSA, AES, DES, and 3DES (in order). This means that the encryption process took less time due to using the Blowfish algorithm (which is a fast algorithm) and changing the F-function. Furthermore, the asymmetric EC algorithm had no impact on the execution time as it was only used to encrypt the key and digest it.

In Fig. 7, the throughput of different techniques, with 50 KB data, are compared. As it is presented, the throughput of the proposed technique is better than the others.

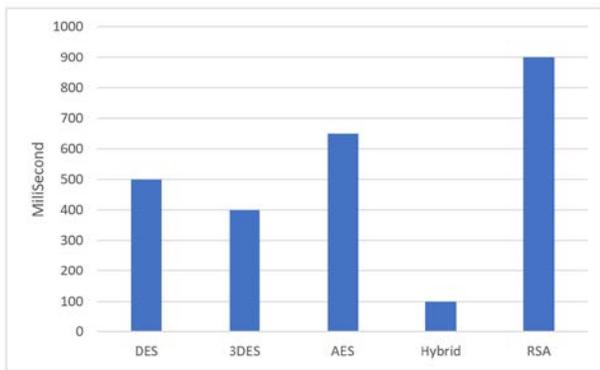


Fig. 6. Execution Time with 50KB Data.

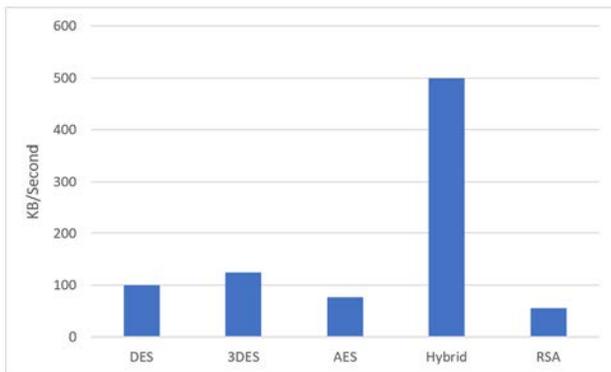


Fig. 7. Throughput with 50KB Data.

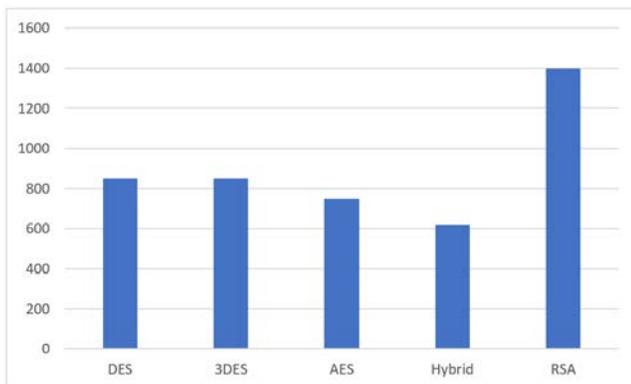


Fig. 8. Execution Time with 1024KB Data.

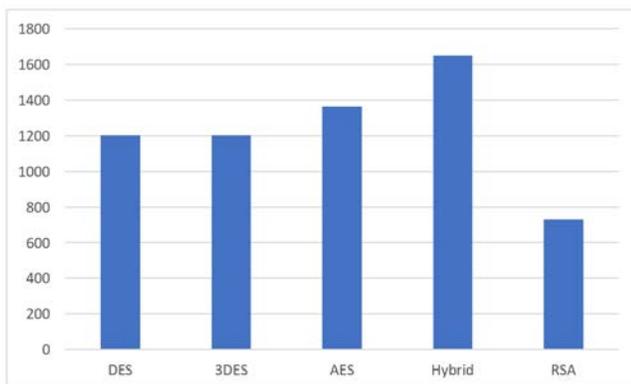


Fig. 9. Throughput with 1024KB Data.

Because the data size can impact the execution time, we increase the data size to 1024 KB to evaluate the solution with larger data. As shown in Fig. 8, the proposed solution has a lower execution time than the other techniques, even with a larger data size (i.e., the 1024 KB). The proposed technique (Hybrid) reduced 230, 230, 130 milliseconds in comparison with symmetric algorithms (i.e., DES, 3DES, and AES), and 731 milliseconds in comparison with the asymmetric algorithm (i.e., RSA). This shows that the proposed technique is faster than the other techniques.

Fig. 9 shows the throughputs of the selected techniques and the proposed solution with 1024 KB data. As shown, the proposed solution has better throughputs, which could be expected as the execution time of the solution is lower. The solution's throughput is, on average, 16% higher than the symmetric algorithms (AES, 3DES, and DES) and over 50% higher than the asymmetric algorithm (RSA).

Finally, we evaluate the selected algorithms with a 2049 data size. Fig. 10 and 11 show the execution time and throughput results. Increasing the data size increases the execution time, but as it was shown, the proposed solution has a lower execution time than DES, 3DES, and RSA. The execution time is 470ms lower than 3DES, 320ms lower than DES, and 670ms lower than RSA. However, the proposed solution's execution time does not work better than the AES. However, considering the other benefits such as the digital signature and hashing, we might ignore the increased execution time, as the time difference is likely to have a very low effect in the cloud computing environment. Future studies should evaluate the proposed solution in the cloud computing environment to determine how it works with larger data and if the mentioned benefits can overshadow the execution time (compared to the AES).

As Fig. 11 shows, the throughput of the proposed solution is higher than DES, 3DES, and RSA. This is because of the changes we applied to the F-function and, in fact, parallel processing in the core of Blowfish encryption. Furthermore, the data encryption and key encryption are separated in this solution, so parallel usage of Blowfish and EC did not affect the execution time and throughput as only private key hash codes are encrypted by EC, and because of their small size, this had almost no effect on the execution time. Moreover, the original data, which has a larger size, is encrypted by the improved Blowfish algorithm, which is faster and has a high performance.

Several studies proposed hybrid and multilevel encryption solutions and/or modified cryptography algorithms to improve cloud security [7, 22-26]. However, our proposed solution benefits using both symmetric and asymmetric algorithms together with a digital signature. Combining the utilised algorithms and improving the Blowfish's F-function provides high data protection and signature integrity with high-performance encryption. However, like other cloud security methods, our proposed solution should be tested and evaluated in cloud computing environments with extensive data.

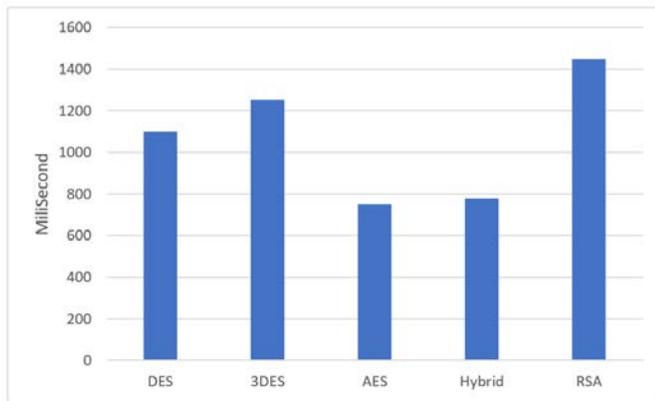


Fig. 10. Execution Time with 2048KB Data.

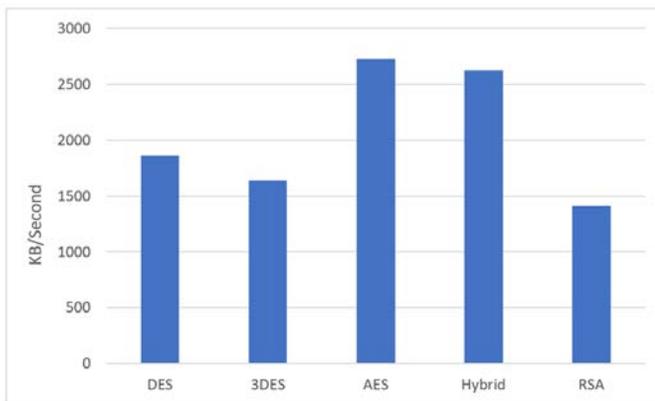


Fig. 11. Throughput with 2048KB Data.

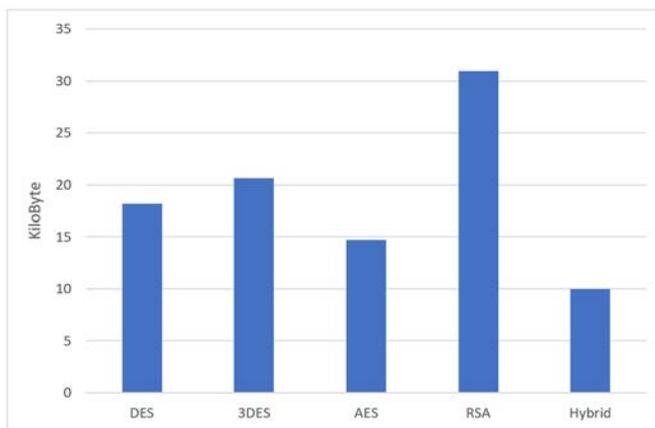


Fig. 12. Memory Usage.

B. Memory Consumption

Memory consumption is one of the important metrics in evaluating encryption techniques. Fig. 12 shows the used memory in the proposed solution and other selected solutions. The proposed solution used lower memory in comparison with the other techniques. The main factors that impact memory usage are the number and type of operations, size of the key and initialisation vectors [21], so we tested the memory consumption with 50KB data only. The lower memory usage of the proposed solution is because of using the Blowfish algorithm as one of the most memory-efficient cryptography

algorithms [21]. The other reason for lower memory usage is applied changes in the core of Blowfish in the proposed solution. In addition, using EC, which uses a 164-bit key that needs a low memory [27], will allow us to use the proposed solution in infrastructures with low memory resources, such as old ones.

V. CONCLUSION

This study developed an improved cryptography solution that can provide efficient and high-performance, secure cloud computing. Because the core of this solution is based on an improved Blowfish (a symmetric algorithm) which decreases the needed time for encryption, the execution time and throughput of the solution are improved. Moreover, using the EC asymmetric cryptography algorithm to encrypt the key, reduces the security challenges of symmetric key exchange algorithms, such as stealing the key in transit. In addition to the mentioned benefits, the proposed solution uses MD5-based digital signatures to provide data integrity. The solution evaluation shows an overall improvement in comparison with AES, DES, 3DES, and RSA. The throughput, memory usage, and execution time of the proposed solution were, on average, better than the other solutions. However, we found that AES was slightly faster and had higher throughput when we increased the data size. We suggest that future studies test the solution in different infrastructures, especially cloud computing environments, with a high volume of data.

REFERENCES

- [1] N. Subramanian and A. Jeyaraj, "Recent security challenges in cloud computing," *Computers & Electrical Engineering*, vol. 71, pp. 28-42, 2018.
- [2] H. Tabrizchi and M. K. Rafsanjani, "A survey on security challenges in cloud computing: issues, threats, and solutions," *The journal of supercomputing*, vol. 76, no. 12, pp. 9493-9532, 2020.
- [3] Z. Balani and H. Varol, "Cloud Computing Security Challenges and Threats," in *2020 8th International Symposium on Digital Forensics and Security (ISDFS)*, 2020: IEEE, pp. 1-4.
- [4] A. Oroboade, T. Aderonke, A. Boniface, and A. J. Gabriel, "Cloud application security using hybrid encryption," *Communications*, vol. 7, pp. 25-31, 2020.
- [5] V. Agarwal, A. K. Kaushal, and L. Chouhan, "A Survey on Cloud Computing Security Issues and Cryptographic Techniques," in *Social Networking and Computational Intelligence: Springer*, 2020, pp. 119-134.
- [6] S. Ray, K. N. Mishra, and S. Dutta, "Big Data Security Issues from the Perspective of IoT and Cloud Computing: A Review," *Recent Advances in Computer Science and Communications*, vol. 12, pp. 1-00, 2020.
- [7] K. Sajay, S. S. Babu, and Y. Vijayalakshmi, "Enhancing the security of cloud data using hybrid encryption algorithm," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-10, 2019.
- [8] D. S. AbdElminaam, "Improving the security of cloud computing by building new hybrid cryptography algorithms," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 40-48, 2018.
- [9] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120-126, 1978.
- [10] F. Thabit, S. Alhomdy, A. H. Al-Ahdal, and S. Jagtap, "A new lightweight cryptographic algorithm for enhancing data security in cloud computing," *Global Transitions Proceedings*, vol. 2, no. 1, pp. 91-99, 2021.
- [11] A. Hussain, C. Xu, and M. Ali, "Security of cloud storage system using various cryptographic techniques," *Int. J. Math. Trends Technol.*, vol. 60, no. 1, pp. 45-51, 2018.

- [12] C. Esposito, A. De Santis, G. Tortora, H. Chang, and K.-K. R. Choo, "Blockchain: A panacea for healthcare cloud-based data security and privacy?," *IEEE Cloud Computing*, vol. 5, no. 1, pp. 31-37, 2018.
- [13] EC. "For how long can data be kept and is it necessary to update it?" https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr/how-long-can-data-be-kept-and-it-necessary-update-it_en (accessed 18/05/2021, 2021).
- [14] ICO. "Principle (e): Storage limitation." <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/storage-limitation> (accessed 18/05/2021, 2021).
- [15] A. R. Wani, Q. Rana, and N. Pandey, "Performance Evaluation and Analysis of Advanced Symmetric Key Cryptographic Algorithms for Cloud Computing Security," in *Soft Computing: Theories and Applications*: Springer, 2019, pp. 261-271.
- [16] B. UMAPATHY and D. KALPANA, "A SURVEY ON CRYPTOGRAPHIC ALGORITHM FOR DATA SECURITY IN CLOUD STORAGE ENVIRONMENT," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 09, p. 2020.
- [17] I. Ahmed, "A brief review: security issues in cloud computing and their solutions," *Telkomnika*, vol. 17, no. 6, 2019.
- [18] M. R. Asassfeh, M. Qataweh, and F. M. AL-Azzeh, "Performance evaluation of blowfish algorithm on supercomputer iman1," *International Journal of Computer Networks & Communications (IJCNC)*, vol. 10, no. 2, 2018.
- [19] S. Chandra, S. Paira, S. S. Alam, and G. Sanyal, "A comparative survey of symmetric and asymmetric key cryptography," in 2014 international conference on electronics, communication and computational engineering (ICECCE), 2014: IEEE, pp. 83-93.
- [20] T. Wollinger, J. Pelzl, V. Wittelsberger, C. Paar, G. Saldamli, and Ç. K. Koç, "Elliptic and hyperelliptic curves on embedded μP ," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 3, no. 3, pp. 509-533, 2004.
- [21] P. Patil, P. Narayankar, D. Narayan, and S. M. Meena, "A comprehensive evaluation of cryptographic algorithms: DES, 3DES, AES, RSA and Blowfish," *Procedia Computer Science*, vol. 78, pp. 617-624, 2016.
- [22] A. PanimalarS, N. Dharani, R. Aiswarya, and P. Shailesh, "Cloud Data Security Using Elliptic Curve Cryptography," 2017.
- [23] V. K. R. Gangireddy, S. Kannan, and K. Subburathinam, "Implementation of enhanced blowfish algorithm in cloud environment," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-7, 2020.
- [24] U. Gupta, M. S. Saluja, and M. T. Tiwari, "Enhancement of Cloud Security and removal of anti-patterns using multilevel encryption algorithms," *International Journal of Recent Research Aspects*, vol. 5, no. 1, pp. 55-61, 2018.
- [25] A. Chauhan and J. Gupta, "A novel technique of cloud security based on hybrid encryption by Blowfish and MD5," in 2017 4th International conference on signal processing, computing and control (ISPCC), 2017: IEEE, pp. 349-355.
- [26] V. Saranya and K. Kavitha, "A modified blowfish algorithm for improving the cloud security," *Elsiyum J*, vol. 4, no. 3, pp. 1-6, 2017.
- [27] D. Mahto and D. K. Yadav, "Performance Analysis of RSA and Elliptic Curve Cryptography," *IJ Network Security*, vol. 20, no. 4, pp. 625-635, 2018.

Evil Twin Attack Mitigation Techniques in 802.11 Networks

Raja Muthalagu¹, Sachin Sanjay²

Department of Computer Science

Birla Institute of Technology and Science Pilani – Dubai Campus, Dubai, UAE

Abstract—Evil Twin Wi-Fi attack is an attack on the 802.11 IEEE standard. The attack poses a threat to wireless connections. Evil Twin Wi-Fi attack is one of the attacks which has been there for a long time. Once the Evil Twin Wi-Fi attack is performed this acts as a gateway to many other attacks such as DNS spoofing, SSL Strip, IP Spoofing, and many more attacks. Thus, preventing the attack is essential for privacy and data security. This paper will be going through in detail how the attack is performed and different measures to prevent the attack. The proposed algorithm sniffs for fake AP using the whitelist in all the channels, once an unauthorized AP is detected the user has an option to de-authenticate any user in the unauthorized network in case any clients do connect to it by accident also the algorithm will be checking if any de-authentication frame is being sent to any of the AP to know which of the AP is being compromised. The efficiency of proposed approach is verified by simulating and mitigating the evil-twin attack.

Keywords—Evil twin Wi-Fi attack; DNS spoofing; SSL strip; IP spoofing

I. INTRODUCTION

In today's world, wireless connections are becoming more important day by day. Almost all the appliance in use gets connected to the internet to customize, improve, and make our lives simpler, these devices are called the Internet of Things (IoT) [1]. The IoT protocol stack and some of the most used protocols and comparison of it from security aspect is discussed. Also, the paper attempts to present various challenges related to security from aspect of protocols, network, and device [2]. Various security gap in most of the current IoT technologies are reported [3-4]. The number of IoT devices will be increase in future and it is important to find specific standard and mechanism to get security in the IoT [4]. The majority of IoT devices use wireless connections to connect to the internet, this is usually IEEE 802.11 standard connection [5]. The IEEE 802.11 is specifically designed for wireless connections, the first 802.11 has been developed in 1997 and is still being improved on. This type of wireless standard is used with frequencies 2.4, 5, 60 GHz, the Evil Twin Wi-Fi attack typically targets this standard [6-9]. To detect evil-twin attack, a real-time client-side detection technique is proposed [7]. A rogue Access Point (AP) detection technique based on the round-trip time (RTT) on the mobile-user side is proposed [8]. A Convolutional Neural Network (CNN) based Evil-Twin Attack (ETA) detection technique is proposed which uses the preamble of the WiFi signal as the feature and uses it to train a CNN based classification model [9]. A novel ETS detection is proposed [10] which make use of the commonly

used network diagnostic tool traceroute. A passive ETA detection scheme is proposed [11] by using channel state information in physical layer. a scan-based self anomaly detection (SSAD), which is a client-side solution to detect and mitigate channel-based man-in-the-middle attacks using access point (AP) scans [12-13]. Even though ETA is an old attack measure taken to prevent this attack is not common and is not implemented by a majority of the population. This attack is usually performed on public Wi-Fi networks where there are multiple users connected to the public network because it is open and multiple clients are connected to it, these types of networks make it a perfect candidate to initiate and perform the attack.

The attack begins with a dummy access point being created and fooling the client's device into thinking the fake access point is a better connection, this can be done either by increasing the signal strength of the connection or by disconnecting the client from the current access point (more on this later). The attack can also be performed in another way where the attacker sniffs the previously connected Wi-Fi connections and makes a fake access point with the same configurations, the device actively searches for previously connected access point and when it finds the fake access point with the same configuration as the legitimate access point the phone connects to the fake access point (AP). Once the device has connected to the attackers AP many man-in-the-middle attacks (MITM) [14] can be performed on this device to compromise the privacy and data security of the victim. In the below sections will be going through different methods to prevent such an attack and a custom method to counter this attack which includes the algorithm and the code explained in detail.

The detection and mitigation of evil twin attacks have recently received huge attention from researchers around the world. A virtual private network (VPN) is one solution available to counter-attack the evil twin attack but it is an inappropriate solution. In general, detecting the evil twin attack can be classified into two categories: client-based solutions [15, 16] and administrator-based solutions [17-18]. Although the majority of approaches being used belong to administrator-based mechanisms, they cannot detect evil twin in real-time and have a heavy overhead. A passive user-side solution, called Wi-Fi legal access point (AP) finder (LAF) is proposed in [19], to the notorious evil twin access point problem, which provide solution to various security problems. Discrete Event System (DES) based approach for Intrusion Detection System (IDS) for evil twin attacks in a Wi-Fi network is discussed [20].

Designing a client-side approach to detect evil twins is not easy for clients because users have limited resources and without an authorization list.

This work analyzes the feature of evil twin attacks and proposes a novel algorithm for mitigating the evil twin attack in a Wi-Fi network. Our scheme can make a deterministic and accurate decision on whether the client device is connecting to an evil twin AP. The implement and evaluate the proposed scheme with real experiments is done in this paper.

The organization of the paper is as follows. The various method of performing an Evil Twin Wi-Fi attack is reported Section II. The proposed system model and algorithm discussed in depth in Section III. The outcome of the proposed methodology is illustrated in Section IV followed by conclusion in Section V.

II. HOW DOES EVIL TWIN WI-FI ATTACK WORK?

An evil twin is a fraudulent Wi-Fi access point that appears to be legitimate but is set up to eavesdrop on wireless communications [1]. To perform this attack few to no additional hardware is required as shown in Fig. 1. In the test run a laptop with a wireless card (supports monitor mode). Alternatively, one can use a raspberry pi with a wireless card and a power bank to perform the attack remotely and discretely. For simplicity, it is recommended use premade tools such as airodump-ng, airdrop-ng and aireplay-ng which can be found preinstalled in offensive operating systems such as Kali Linux, Parrot OS, and many more.

First set the card to monitor mode this can be done either through “iwconfig” or using the preinstalled tool called “airmon-ng”. Once the wireless card is in monitor mode then it can move to the next phase where it scans for the wireless connection, we want to perform the attack on. After obtain required information such as SSID, channel and other things, a malicious AP with the same configuration can be created.

Now there are two ways to do the next part one is to wait patiently for the victim’s device to connect to the attacker’s AP or we use another method, which is a flaw in the 802.11b and some other standards where the management frames are not protected or validated. This method involves sending a crafted “Death frame” to the targeted AP as depicted in Fig. 2. The AP receives the frame and forwards it to the respective clients, the clients then receive the forwarded frame and are forced to disconnect from the targeted AP. The counterfeit access point may be given the same SSID and BSSID as a nearby Wi-Fi network.

The evil twin can be configured to pass Internet traffic through to the legitimate access point while monitoring the victim’s connection [2]. What makes this attack so dangerous is that if a client does connect to this network intentionally or unintentionally, the client has no way of knowing that the device connected to the attacker’s AP and his privacy and security is compromised.

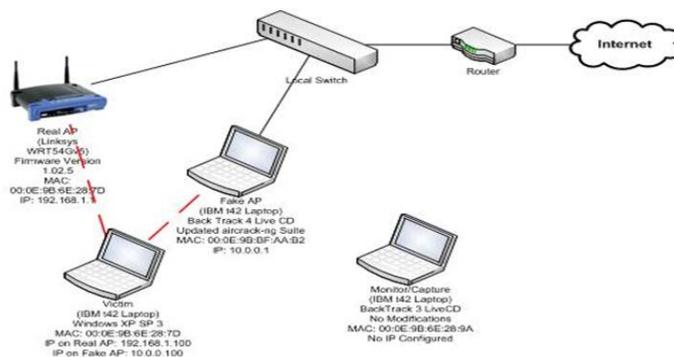


Fig. 1. Evil Twin Wi-Fi Attack Scenario.

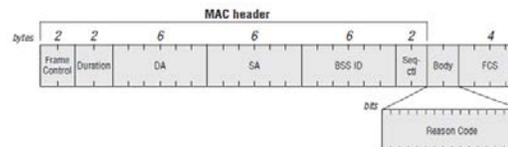


Fig. 2. Structure of Deauthentication Frame.

III. PROPOSED SYSTEM MODEL AND ALGORITHM

The proposed system in this paper is a counterattack against the Evil Twin Wi-Fi attack. The algorithm requires a list of known BSSID for whitelisting. The algorithm sniffs for fake AP using the whitelist in all the channels, once an unauthorized AP is detected the user has an option to de-authenticate any user in the unauthorized network in case any clients do connect to it by accident also the algorithm will be checking if any de-authentication frame is being sent to any of the AP to know which of the AP is being compromised. The program checks each channel for about 2 seconds and switches to the next channel also it scans for all 14 channels available in the 2.4GHz range.

As seen in Fig. 3, there is an Evil Twin Access point that is trying to mimic the actual access point. In Fig. 4, one of the victims connects to the evil twin, the counterattack then figures out which access point gets attacked and informs the Administrator about the attack. The administrator initiates the counterattack to disconnect the user from the Evil Twin Access point. First, the wireless card must set in monitor mode. Once the mode is set, then the whitelist, network name, and wireless interface as the arguments for the program needs to be provide. The program initiates by scanning networks in all the channels and finding any BSSID which is not in the whitelist this is done by checking all the wireless beacons. In the wireless beacon, check for its BSSID if the BSSID does not exist within our whitelist then mark it.

The program also checks if there is any death frame sent this similar to the whitelist where scanning of the wireless packets is done and we specifically check if the packet contains an 802.11 death frame, if it does this packet might indicate that the transmitter of the packet is the targeted AP. For the next step, a death packet is will send to the unauthorized AP and this is done by crafting a wireless death frame and specifying the transmitter address as target AP BSSID and death reason as 7. This program can be a good counterattack against unauthorized AP. It can be used with a raspberry pi and a wireless card and placed in separate locations. This can also be used to monitor wireless networks and can trigger an attack remotely.

The Fig. 5 is the complete flow of the code done in a flowchart. Here the scanning of channel for death frames and the quit command run in parallel this is to prevent any sort of interference to the code while scanning.

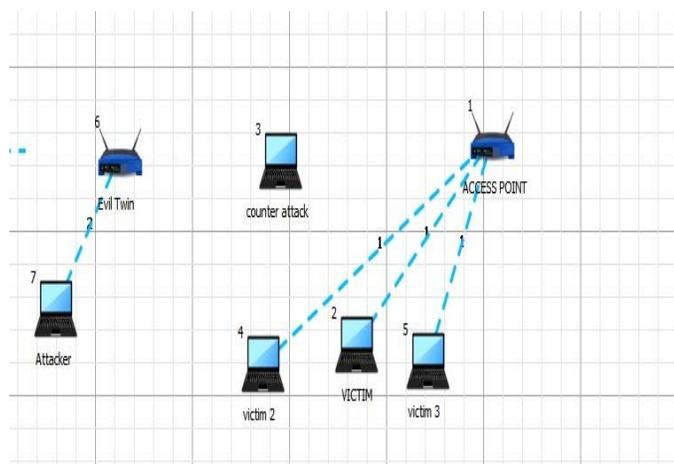


Fig. 3. Evil Twin Access Point which is trying to mimic the Actual Access Point.

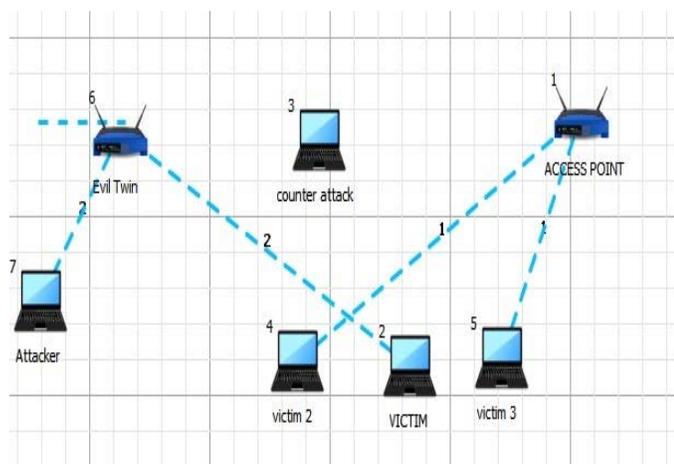


Fig. 4. One of the Victims Connects to the Evil Twin.

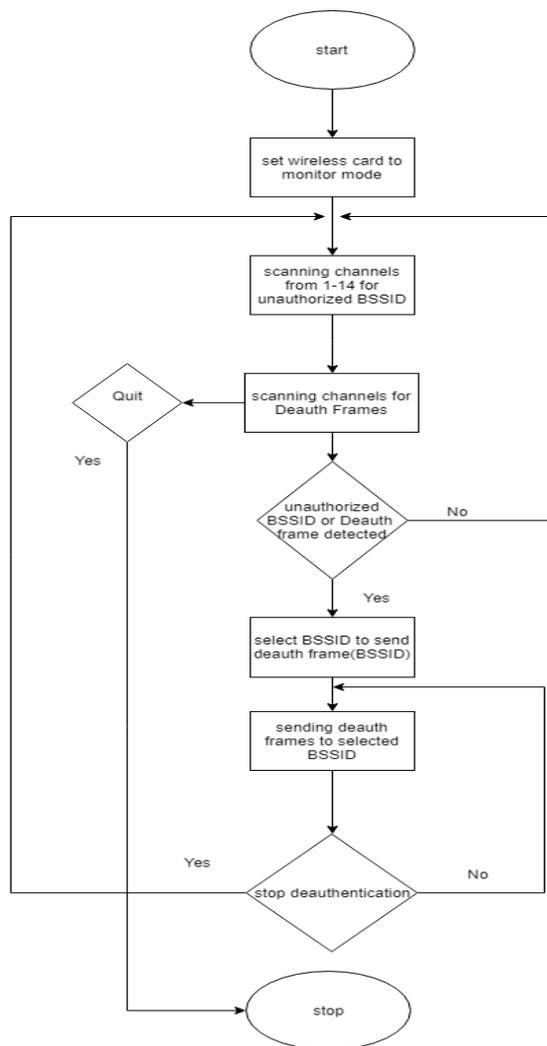


Fig. 5. Flow of the Code.

IV. RESULTS AND DISCUSSIONS

This form of counterattack provides an easier interface and simplicity as opposed to manual methods. Also, it is compulsory to spread awareness of the attack, as more people know the attack less likely they fall for the attack. There are many other methods to mitigate this attack, for example using a virtual private network whenever connecting to an open Wi-Fi connection. Another method is by the new WPA3 standards which is released couple of months ago. This standard protects the wireless management frame thereby indirectly preventing the Evil Twin Wi-Fi attack.

In the above Fig. 6, one can see that only at higher education levels that the people are more informed about the attack. A study in UK has been conducted by XIRRUS Wi-Fi Network [4], there are 8.5 million Wi-Fi hotspots available in UK. XIRRUS has surveyed over 300 respondents between 18 and 75 about how they connect to the internet, what they do when they are connected, and how they connect .79% of respondents said that they did not trust security measures on public Wi-Fi, while 62% of respondents said they still connect to public networks.

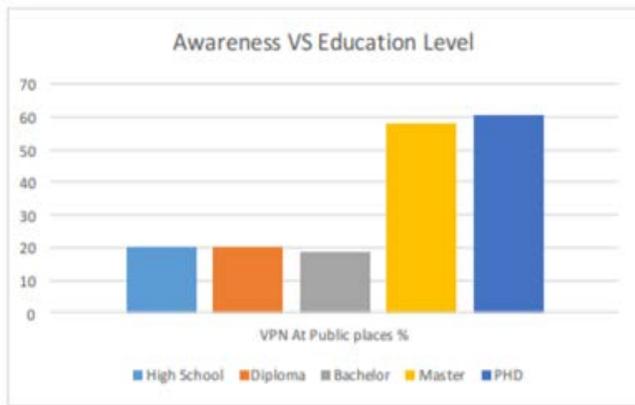


Fig. 6. Awareness Vs Education Level.

The study also showed that 58% of the survey respondents admitted using public Wi-Fi but only 7% had installed VPN service on their devices to protect their confidential data [3]. In conclusion Evil Twin Wi-Fi attack poses a serious threat to the users currently; therefore methods must be developed to prevent this attack. One of the methods which can be used to prevent such attacks is to use our proposed method where we can form an inexpensive method to counter the Evil Twin Wi-Fi attack.

V. CONCLUSION

There is always a network security threat that wireless users are vulnerable to be attacked by evil twins in WLANs so, in this paper, a novel approach to mitigating the evil twin attack to protect wireless users from the hackers is proposed. The attack is simulated, and different measures are recommended to prevent the attack. The proposed algorithm sniffs for fake AP using the whitelist in all the channels, once an unauthorized AP is detected the user has an option to de-authenticate any user in the unauthorized network in case any clients do connect to it by accident also the algorithm will be checking if any de-authentication frame is being sent to any of the AP to know which of the AP is being compromised. The efficiency of proposed approach is verified by simulating and mitigating the evil-twin attack.

REFERENCES

- [1] H. Suo, J. Wan, C. Zou and J. Liu, "Security in the Internet of Things: A Review," 2012 International Conference on Computer Science and Electronics Engineering, Hangzhou, , pp. 648-651, 2012.
- [2] Sardeshmukh, H., & Ambawade, D. "Internet of Things: Existing protocols and technological challenges in security". International Conference on Intelligent Computing and Control (I2C2), pp. 1-7., 2017.
- [3] Sain, M., Kang, Y. J., & Lee, H. J. "Survey on security in Internet of Things: State of the art and challenges". 19th International Conference on Advanced Communication Technology (ICACT), pp. 699-704, Feb. 2017.
- [4] Yang Y., Wu L., Yin G., Li L., Zhao H. "A survey on security and privacy issues in internet-of-things", IEEE Internet of Things Journal, vol.4, no.5 , pp. 1250-1258, 2017.

- [5] N. Baghaei and R. Hunt, "IEEE 802.11 wireless LAN security performance using multiple clients," Proceedings. 2004 12th IEEE International Conference on Networks (ICON 2004) (IEEE Cat. No.04EX955), Singapore, pp. 299-303, 2004.
- [6] Agarwal, M., Biswas, S. & Nandi, S. "An Efficient Scheme to Detect Evil Twin Rogue Access Point Attack in 802.11 Wi-Fi Networks". Int J Wireless Inf Networks, vol. 25, pp.130-145, 2018.
- [7] O. Nakhila and C. Zou, "User-side Wi-Fi evil twin attack detection using random wireless channel monitoring," IEEE Military Communications Conference, Baltimore, MD, pp. 1243-1248, 2016.
- [8] Kitisiworapan, S., Jansang, A. & Phonphoem, A. "Client-side rogue access point detection using a simple walking strategy and round-trip time analysis". J Wireless Com Network, 252, 2020.
- [9] Yinghua Tian*, Sheng Wang and Long Zhang "Convolutional neural network based evil twin attack detection in WiFi networks" 2nd International Conference on Computer Science Communication and Network Security, 7 pages, Feb. 2021
- [10] A. Burns, L. Wu, X. Du and L. Zhu, "A Novel Traceroute-Based Detection Scheme for Wi-Fi Evil Twin Attacks," GLOBECOM 2017 - 2017 IEEE Global Communications Conference, Singapore, pp. 1-6, 2017,
- [11] Wang C. et al. "Detecting Evil-Twin Attack with the Crowd Sensing of Landmark in Physical Layer". Algorithms and Architectures for Parallel Processing. ICA3PP 2018. Lecture Notes in Computer Science, vol 11337. Springer, Cham, 2018.
- [12] Sheng Gong, Hideya Ochiai, Hiroshi Esaki, "Scan-Based Self Anomaly Detection: Client-Side Mitigation of Channel-Based Man-in-the-Middle Attacks Against Wi-Fi", Computers Software and Applications Conference (COMPSAC) 2020 IEEE 44th Annual, pp. 1498-1503, 2020.
- [13] Said Abdul Ahad Ahadi, Nitin rakesh, Sudeep varshney, "Overview On Public Wi-Fi Security Threat Evil Twin Attack Detection", Advent Trends in Multidisciplinary Research and Innovation (ICATMRI) 2020 IEEE International Conference on, pp. 1-6, 2020.
- [14] Ahmad F, Adnane A, Franqueira VNL, Kurugollu F, Liu L. "Man-In-The-Middle Attacks in Vehicular Ad-Hoc Networks: Evaluating the Impact of Attackers' Strategies." Sensors (Basel). Vol. 18 no. 11, 4040, 2018
- [15] H. Han, B. Sheng, C. C. Tan, Q. Li, and S. Lu. "A timing-based scheme for rogue ap detection". IEEE Transactions on Parallel & Distributed Systems, vol. 22, no. 11 pp.1912-1925, 2011.
- [16] C. D. Mano, A. Blaich, Q. Liao, Y. Jiang, D. A. Cieslak, D. C. Salyers, and A. Striegel. "Ripps: Rogue identifying packet payload slicer detecting unauthorized wireless hosts through network traffic conditioning". ACM Transactions on Information & System Security, vol. 11, no. 2, pp-1-23, 2008.
- [17] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. "Wireless device identification with radiometric signatures." ACM International Conference on Mobile Computing and NETWORKING, pp. 116-127, 2008.
- [18] N. T. Nguyen, G. Zheng, Z. Han, and R. Zheng. "Device fingerprinting to enhance wireless security using nonparametric bayesian method" In INFOCOM, 2011 Proceedings IEEE, pp. 1404-1412, 2011.
- [19] Fu-Hau Hsu, Chuan-Sheng Wang, Chih-Wen Ou, Yu-Liang Hsu, "A passive user-side solution for evil twin access point detection at public hotspots", International Journal of Communication Systems, vol. 23, no.4, pp. e4460, 2020.
- [20] Nirmal S Selvarathinam, Amit Kumar Dhar, Santosh Biswas, "Evil Twin Attack Detection using Discrete Event Systems in IEEE 802.11 Wi-Fi Networks", Control and Automation (MED) 2019 27th Mediterranean Conference on, pp. 316-321, 2019.

Solving Reactive Power Scheduling Problem using Multi-objective Crow Search Algorithm

Surender Reddy Salkuti

Department of Railroad and Electrical Engineering, Woosong University
Daejeon, Republic of Korea

Abstract—This paper presents the solution of multi-objective based optimal reactive power dispatch (MO-ORPD) problem by optimizing the system power losses and voltage stability enhancement index (VSEI)/L-index objectives. ORPD problem is considered as an important issue from system security and operational point of view for optimal steady-state operation of power system. Here, single-objective based ORPD problem is solved using Crow Search Algorithm (CSA) and multi-objective based ORPD problem is solved using multi-objective CSA (MO-CSA). The CSA is considered as an efficient and robust algorithm which determines the global optimal solution for solving the non-linear and discontinuous objective functions. Two standard test systems, i.e., IEEE 30 and 57 bus systems are considered to show the effectiveness, suitability and robustness of CSA and MO-CSA for solving the ORPD problem.

Keywords—Optimal power flow; crow search algorithm; evolutionary algorithms; loss minimization; reactive power scheduling; multi-objective optimization

NOMENCLATURE

N_l	Number of lines
G_k	Conductance of k^{th} line
N_B	Number of buses
N_G	Number of generator/PV buses.
Q_{Gi}	Reactive power generation from the generator and VAR source at i^{th} bus
Q_{Di}	Reactive power demand at i^{th} bus
P_{Gi}	Power generation at i^{th} bus
N_c	Number of switchable capacitors/VAR sources.
$P_{inj,i}$	Active power injection at i^{th} bus
$Q_{inj,i}$	Reactive power injection at i^{th} bus
N_L	Number of load buses
N_T	Number of tap changing transformers
S_{Li}	Apparent power flowing through i^{th} transmission line

I. INTRODUCTION

Solution to the optimal reactive power dispatch (ORPD)/scheduling is required to meet the load demands at the nominal frequency as the reactive powers to maintain

nominal voltage throughout the system. The ORPD is required to obtain secured, efficient and reliable operating condition of power system. Reactive power is localized to the point where the voltage needs regulating. Generally controlling of generator voltages is done automatically at corresponding generation power plant location. Series and shunt reactive devices are on the grid also provides the required local corrections. The energy management system of power system processes the raw data to provide reliable data base for the analysis of the power network [1]. Basically it includes topology processing and state estimation. Topology processing unit uses the dynamic status of the switches, i.e. breakers and isolators, to translate the network model into a node branch representation that can be analyzed by matrix methods. State estimation calculates the complex bus voltages and branch flows using the network model parameters, and values telemetered via Supervisory Control and Data Acquisition. The state estimator output provides a reliable database for other analysis applications such as optimal power flow (OPF), ORPD, etc. From planning perspective, the OPF is used for various applications for the shunt volt ampere reactive (VAR) planning [2], transfer capability studies, series capacitor planning, loss optimization studies, and reactive interchange studies. The literature related to ORPD problem is described next.

A. Literature Review

Review of various optimization approaches developed for the solution and analysis of ORPD with the main goal of enhancing efficiency, performance and security of power system is presented in [3]. The author in [4] solves the ORPD problem with voltage stability enhancement, and power loss, voltage deviation minimizations using the evolutionary based Artificial Bee Colony (ABC) algorithm. The author in [5] solves the ORPD problem using meta-heuristic based Crow Search Algorithm (CSA) and finds the optimal control variables settings. An enhanced and efficient differential evolution (DE) with new mutation approach is described in [6, 7] to solve ORPD problem with HVDC transmission link. A methodology based on entropy evolution approach is proposed in [8] using Gravitational Search technique and fractional particle swarm optimization (PSO) for the solution of ORPD problem. The author in [9] proposes the solution to ORPD considering the system losses and voltage deviation minimization as objectives. A mathematical model of ORPD considering the energy loss minimization at the costs of adjusting control equipment and the current time interval is proposed in [10].

The author in [11] solves the ORPD problem using the marine predators' algorithm by handling uncertainties concerning to the output of solar, wind units and load demand. Differential evolution (DE) based productive stochastic search approach for solving the ORPD problem is presented in [12]. A fractional evolutionary based method that achieves the objectives of ORPD problem by including the FACTS controllers is presented in [13]. The author in [14] proposes a stochastic fractal search methodology for the solution of ORPD problem considering the voltage stability index (VSI), voltage deviation and power losses objectives. A Levy Interior Search technique is proposed in [15] for the solution of multi-objective (MO) based ORPD problem. A two-step based methodology for the solution of MO based ORPD problem is proposed in [16].

B. Motivation and Contribution

From literature review section, it can be concluded that there is a research gap for solving the single and multiple objective based ORPD problems and determining the global optimal solution using the multi-objective based CSA. The conflicting nature of the system power loss and voltage stability margins, when treated as objective functions in the ORPD problem reveal that these objective functions shouldn't be treated independently. When the power system efficiency, security, and reliability are of equal concern, the two considered important and conflicting objectives are to be minimized simultaneously using multi-objective optimization (MOO) techniques. This provided motivation to use MO based CSA for simultaneous optimization of two conflicting ORPD objectives. The major contributions of this work are:

- ORPD is formulated as a non-linear, discrete optimization problem which is solved by using evolutionary based crow search algorithm (CSA).
- ORPD problem is solved considering minimizations of power losses and L-index objectives.
- The proposed problem with these two important and conflicting objectives must be solved using the MOO techniques. In this work, multi-objective CSA is used as a MOO algorithm.
- The solution of ORPD problem has been tested on IEEE 30 and 57 bus systems.
- Obtained results depict robustness, effectiveness and superiority over the other algorithms reported in the literature.

The remainder of this work is arranged as: Mathematical modeling of ORPD is illustrated in Section II. Section III describes the simulation results on two standard systems. Section IV summarizes the contributions and conclusions of this paper.

II. PROBLEM FORMULATION

The goal of ORPD is to set optimal values to the settings of control variables such that optimal values of objective function can be met. The control variables considered for this ORPD are the voltage magnitudes of generator buses, tap settings of transformers and reactive power generation from switchable VAR/capacitor banks [17]. Fig. 1 depicts the control variables considered in the proposed ORPD problem. ORPD is solved initially as single-objective optimization (SOO) problem by selecting each objective independently, and then it is solved as MOO problem by selecting all objectives simultaneously [18]. Here, two distinct objectives are selected for solving ORPD problem and they are presented below:

A. Objective 1: Power Loss Minimization

For the ORPD problem, minimization of power loss is selected as an objective (J_1), and it is a non-linear function of system bus voltages and phase angles [19]. This objective function can be expressed as,

$$J_1 = P_{loss} = \sum_{k=1}^{N_l} G_k [V_i^2 + V_j^2 - 2V_i V_j \cos(\delta_i - \delta_j)] \quad (1)$$

The above power loss equation can be rewritten as,

$$J_1 = P_{loss} = \frac{1}{2} \sum_{i=1}^{N_B} \sum_{\substack{j=1 \\ i \neq j}}^{N_B} G_{ij} [V_i^2 + V_j^2 - 2V_i V_j \cos(\delta_i - \delta_j)] \quad (2)$$

B. Objective 2: Voltage Stability Enhancement Index (VSEI)/ L-Index Minimization

For the ORPD problem, the L-index is selected as another objective (J_2) and it utilizes the information from general load flow (LF) solution. L-index is determined at all the load buses, and it illustrates proximity of power system to the point of voltage collapse [20]. L-index is determined by,

$$L_j = \left| 1 - \sum_{i=1}^{N_G} \bar{F}_{ji} \frac{E_i}{E_j} \right| \quad j = N_G + 1, \dots, N_B \quad (3)$$

In equation (3), the \bar{F}_{ji} values are determined from the $[Y_{Bus}]$ as,

$$\begin{bmatrix} \bar{I}^G \\ \bar{I}^L \end{bmatrix} = \begin{bmatrix} Y^{GG} & Y^{GL} \\ Y^{LG} & Y^{LL} \end{bmatrix} \begin{bmatrix} E_G \\ E_L \end{bmatrix} \quad (4)$$

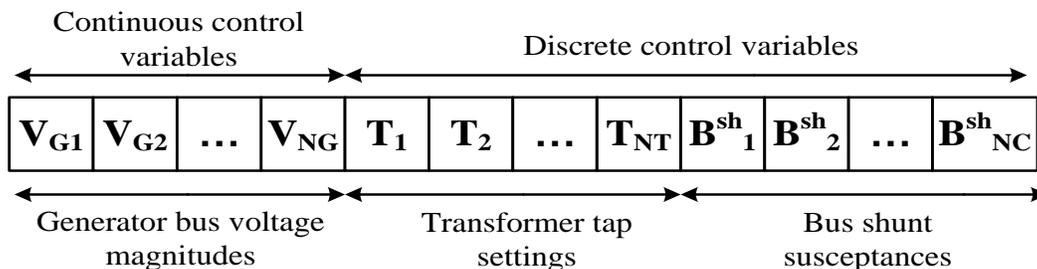


Fig. 1. Control Variables Considered in ORPD Problem.

Where E_G , E_L , \bar{I}^G , \bar{I}^L are voltage and current at generator and load nodes respectively. Rearranging the above equation [21],

$$\begin{bmatrix} E_L \\ \bar{I}^G \end{bmatrix} = \begin{bmatrix} \bar{Z}^{LL} & \bar{F}^{LG} \\ \bar{K}^{GL} & \bar{Y}^{GG} \end{bmatrix} \begin{bmatrix} \bar{I}^L \\ E_G \end{bmatrix} \quad (5)$$

This method is based on calculation of $[\bar{F}^{LG}]$, which is obtained after reordering the generator and load buses. Where $\bar{F}^{LG} = -[\bar{Y}^{LL}]^{-1}[\bar{Y}^{LG}]$. Bus with maximum L-index is the suitable location for the reactive compensation [22]. This minimization of L-index objective can be formulated as sum of squared values of L-indices, and it can be expressed as,

$$J_2 = \text{VSEI} = \sum_{j=N_G+1}^{N_B} L_j^2 \quad (6)$$

C. Equality Constraints

These are general load flow (LF) expressions, and they are represented by [23],

$$P_{Gi} - P_{Di} = P_{inj,i} = V_i \sum_{j=1}^{N_B} V_j [G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}] \quad i = 1, 2, \dots, N_B \quad (7)$$

$$Q_{Gi} - Q_{Di} = Q_{inj,i} = V_i \sum_{j=1}^{N_B} V_j [G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}] \quad i = 1, 2, \dots, N_B \quad (8)$$

Where $Y_{ij} = G_{ij} + jB_{ij}$.

D. Inequality Constraints

They represent the operational and technical limits on the power network.

1) *Generator constraints*: The amount of power output from generators must be within the specific limits [24, 25] and they can be expressed as,

$$P_{Gi}^{min} \leq P_{Gi} \leq P_{Gi}^{max} \quad i = 1, 2, 3, \dots, N_G \quad (9)$$

$$Q_{Gi}^{min} \leq Q_{Gi} \leq Q_{Gi}^{max} \quad i = 1, 2, 3, \dots, N_G \quad (10)$$

Voltage magnitudes of generators are limited by [26],

$$V_{Gi}^{min} \leq V_{Gi} \leq V_{Gi}^{max} \quad i = 1, 2, 3, \dots, N_G \quad (11)$$

Where ‘min’ and ‘max’ represent the lower and upper limits of respective variables.

2) *Reactive power source/switchable VAR source constraints*: This constraint is limited by VAR source capacity limit [27] and it can be expressed as,

$$Q_{ci}^{min} \leq Q_{ci} \leq Q_{ci}^{max} \quad i = 1, 2, \dots, N_c \quad (12)$$

3) *Transformer constraints*: The transformer tap position is limited by [28],

$$T_i^{min} \leq T_i \leq T_i^{max} \quad i = 1, 2, \dots, N_T \quad (13)$$

4) *Security constraints*: The limits on voltage magnitudes of load buses and transmission lines are considered as security constraints [29, 30], and they are expressed as,

$$V_{Li}^{min} \leq V_{Li} \leq V_{Li}^{max} \quad i = 1, 2, \dots, N_L \quad (14)$$

$$S_{Li} \leq S_{Li}^{max} \quad i = 1, 2, \dots, N_l \quad (15)$$

In this work, the SOO based ORPD problems are solved by using crow search algorithm (CSA), and MOO based ORPD problems are solved by using MO-CSA. Description of CSA and MO-CSA is presented in [31]-[33]. The detailed flow chart of MO-CSA is shown in Fig. 2.

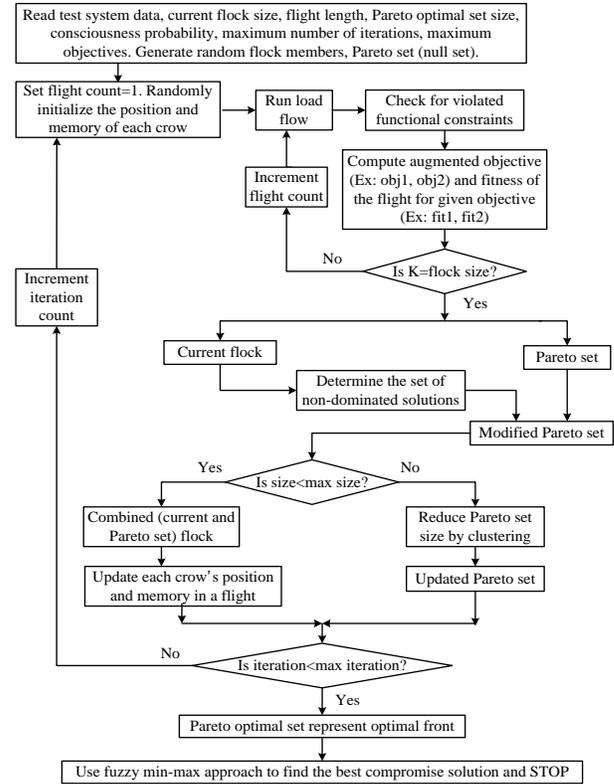


Fig. 2. Flow Chart of the MO-CSA.

III. RESULTS AND DISCUSSION

In this paper, to show the robustness and performance of proposed ORPD problem, CSA and MO-CSA algorithms are used and they are examined on IEEE 30 bus and 57 bus systems [34-37]. In this ORPD problem, voltages at generator buses, settings of transformer taps and values of shunt capacitance are considered as control variables. Here, three cases are implemented, and they are:

- Case 1: System power loss minimization
- Case 2: L-index/ VSEI minimization
- Case 3: Simultaneous optimization of active power losses and L-index

A. Simulation Results on 30 Bus System

This system has 6 generator/PV buses, 24 demands, 41 lines, 4 tap changing transformers and 9 shunt capacitance/compensation buses. In this test system, 19 control variables (6 voltage magnitudes of generator buses, 4 tap changing transformers and 9 bus shunt capacitance) are considered. Lower and upper limits of generator bus voltages, tap settings and shunt susceptance are (0.95 pu and 1.1 pu), (0.9 pu and 1.1 pu), and (0.0 pu and 0.05 pu).

TABLE I. SETTINGS OF CONTROL VARIABLES FOR 30 BUS SYSTEM

Control variables	Single Objective Optimization (SOO)								MOO
	Case 1				Case 2				Case 3
	GA	GSA	DE	CSA	GA	PSO	DE	CSA	MO-CSA
V _{G1} (pu)	1.0	1.07165	1.1	1.0926	1.045	1.052	1.0511	1.0532	1.0853
V _{G2} (pu)	0.999	1.02219	1.0944	1.0815	1.0573	1.034	1.0345	1.0465	1.0619
V _{G5} (pu)	0.974	1.04	1.0749	1.0629	1.0718	1.0465	1.0426	1.0356	1.0562
V _{G8} (pu)	1.007	1.0507	1.0768	1.0532	1.0423	1.0321	1.0415	1.0422	1.0485
V _{G11} (pu)	1.0894	0.97712	1.0999	1.0986	1.0250	1.0252	1.0326	1.0624	1.0826
V _{G13} (pu)	1.088	0.96765	1.0999	1.0413	1.045	1.0563	1.0442	1.0315	1.0385
T _{6,9} (pu)	NA	1.0984	1.0465	1.05	0.925	0.90	0.95	0.925	0.9375
T _{6,10} (pu)	NA	0.9824	0.9097	0.9875	0.9875	0.95	1.0125	0.9375	0.9875
T _{4,12} (pu)	NA	1.0959	0.9867	1.0215	1.0215	0.9875	1.025	0.9875	1.025
T _{28,27} (pu)	NA	1.0593	0.9689	1.075	1.05	1.075	0.95	1.025	1.0625
b _{sh10} (pu)	NA	0.016537	0.05	0.04	0.04	0.02	0.04	0.05	0.03
b _{sh12} (pu)	NA	0.043722	0.05	0.05	0.03	0.05	0.05	0.03	0.04
b _{sh15} (pu)	NA	0.01199	0.05	0.05	0.04	0.03	0.03	0.03	0.05
b _{sh17} (pu)	NA	0.020876	0.05	0.05	0.05	0.04	0.03	0.05	0.04
b _{sh20} (pu)	NA	0.03577	0.04406	0.03	0.04	0.03	0.03	0.04	0.03
b _{sh21} (pu)	NA	0.02602	0.05	0.02	0.03	0.04	0.04	0.03	0.04
b _{sh23} (pu)	NA	0.0	0.028004	0.04	0.02	0.03	0.05	0.02	0.05
b _{sh24} (pu)	NA	0.013839	0.05	0.05	0.05	0.03	0.02	0.04	0.04
b _{sh29} (pu)	NA	0.03	0.025979	0.03	0.05	0.04	0.03	0.05	0.03
Power loss (MW)	4.2716	4.5143	4.555	4.1023	---	---	---	6.724	4.927
VSEI	---	---	---	0.1298	0.1134	0.1105	0.1126	0.1103	0.1188

Table I shows optimal settings of control variables for 30 bus system. Here, first the loss minimization objective is optimized independently (i.e., Case 1) using the genetic algorithm (GA), DE, gravitational search algorithm (GSA) and CSA, and obtained results are presented in Table I [38-39]. The power loss obtained using CSA is 4.1023 MW and the obtained corresponding VSEI value is 0.1298. Later, the VSEI/L-index objective is optimized independently (i.e., Case 2) using GA, PSO, DE and CSA, and obtained results are reported in Table I. The optimum L-index value obtained by using the CSA is 0.1188, and obtained power loss value is 4.927 MW.

From these case studies, it is clear that when the power loss is minimized then L-index has been deviated from optimum, and vice-versa. Hence, these two objectives are optimized simultaneously (i.e., Case 3) by using the MO-CSA. Pareto optimal front (POF) obtained for Case 3 has been depicted in Fig. 3. The best-compromised solution can be determined from the POF using fuzzy satisfaction method, and it has system power loss of 4.927 MW, and L-index of 0.1188.

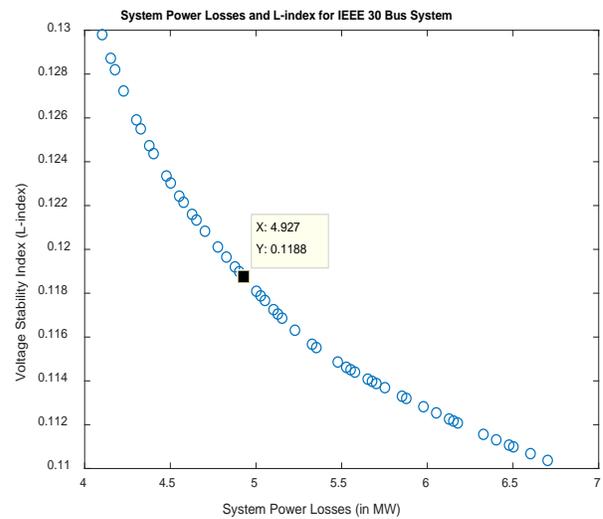


Fig. 3. Pareto Optimal Front for Case 3 of 30 Bus Systems.

B. Simulation Results on 57 Bus System

This system has 7 PV buses, 42 loads, 80 lines, 15 tap changing transformers and 3 shunt capacitance/compensation buses. In this test system, 25 control variables (7 generator bus voltages, 15 tap changing transformers and 3 bus shunt capacitance) are considered. Lower and upper Minimum and maximum limits of generator bus voltages, tap settings and shunt susceptance are (0.95 pu and 1.1 pu), (0.9 pu and 1.1 pu), and (0.0 pu and 0.05 pu).

Table II shows optimal settings of controls for 57 bus system. Here, first the loss minimization objective is optimized independently (i.e., Case 1) using the GA, ABC, GSA and CSA, and the obtained results are reported in Table II.

The power loss obtained using CSA is 24.1022 MW and the obtained corresponding VSEI value is 0.1849. Later, the VSEI/L-index objective is optimized independently (i.e., Case 2) by using GA, PSO and CSA, and the obtained values are reported in Table II [40-41]. The values of optimum L-index obtained by using CSA is 0.1363, and obtained power loss value is 38.682 MW.

From these case studies, it is clear that when the power loss is optimized then L-index has been deviated from optimum value, and vice-versa. Hence, these two objectives are optimized simultaneously (i.e., Case 3) by using the MO-CSA. The POF obtained for Case 3 has been depicted in Fig. 4. The trade-off solution can be obtained from the POF using fuzzy satisfaction method, and it has system power loss of 29.952 MW, and L-index of 0.1513.

TABLE II. SETTINGS OF CONTROL VARIABLES FOR 57 BUS SYSTEM

Control variables	Single Objective Optimization (SOO)							MOO
	Case 1				Case 2			Case 3
	GA	ABC	GSA	CSA	GA	PSO	CSA	MO-CSA
V _{G1} (pu)	1.06	1.0532	1.06	1.063	1.0712	1.0752	1.0786	1.0724
V _{G2} (pu)	1.06	1.0587	1.0582	1.055	1.0685	1.0596	1.0652	1.0507
V _{G3} (pu)	1.0483	1.052	1.0466	1.046	1.0567	1.0615	1.0451	1.0520
V _{G6} (pu)	1.0423	1.0358	1.0409	1.021	1.0466	1.0590	1.0480	1.0477
V _{G8} (pu)	1.06	1.0518	1.0587	1.053	1.0591	1.0620	1.0589	1.0535
V _{G9} (pu)	1.0432	1.0436	1.0417	1.015	1.0488	1.0532	1.0674	1.0480
V _{G12} (pu)	1.0387	1.0455	1.0377	1.064	1.0432	1.5589	1.0483	1.0551
T _{4,18} (pu)	0.9	1.02	0.944	0.95	1.0	0.95	0.9875	0.925
T _{4,18} (pu)	1.1	1.03	1.0182	0.9625	0.95	0.9125	0.9125	0.9875
T _{21,20} (pu)	1.0314	0.95	1.0207	1.05	0.9125	0.9625	0.95	1.075
T _{24,26} (pu)	1.0097	1.03	1.0110	1.0	0.9625	1.0	0.9	1.05
T _{7,29} (pu)	0.9754	0.98	0.9744	0.9375	0.95	1.0125	1.05	0.9125
T _{34,32} (pu)	0.9746	1.05	0.9721	0.9	0.925	1.0375	0.9625	0.9625
T _{11,41} (pu)	0.9	0.95	0.9015	0.95	0.925	0.9875	0.9125	1.0375
T _{15,45} (pu)	0.9726	0.98	0.9723	1.025	0.95	1.075	0.9875	0.975
T _{14,46} (pu)	0.9538	0.96	0.9537	1.0	0.9125	0.9125	1.0125	0.9625
T _{10,51} (pu)	0.9680	0.99	0.9664	0.9875	1.075	1.0625	1.05	1.0375
T _{13,49} (pu)	0.9264	1.04	0.9269	1.025	0.9	0.9875	1.1	1.05
T _{11,43} (pu)	1.1	1.08	0.9645	0.95	0.95	0.9125	1.0375	0.9125
T _{40,56} (pu)	1.0624	0.99	0.9943	1.0125	1.075	1.05	0.95	0.9625
T _{39,57} (pu)	1.0265	0.97	0.9737	0.9875	1.0	0.925	0.975	0.975
T _{9,55} (pu)	0.9764	1.02	0.9750	0.95	0.9125	0.975	0.9875	1.0625
b _{sh18} (pu)	0.0999	0.0785	0.0928	0.05	0.03	0.05	0.05	0.03
b _{sh25} (pu)	0.059	0.05656	0.0589	0.04	0.04	0.05	0.04	0.03
b _{sh53} (pu)	0.063	0.04953	0.0628	0.05	0.04	0.04	0.02	0.05
Power loss (MW)	24.3826	24.1025	24.2619	24.1022	38.247	38.584	38.682	29.952
VSEI	0.1846	---	---	0.1849	0.1422	0.1408	0.1363	0.1513

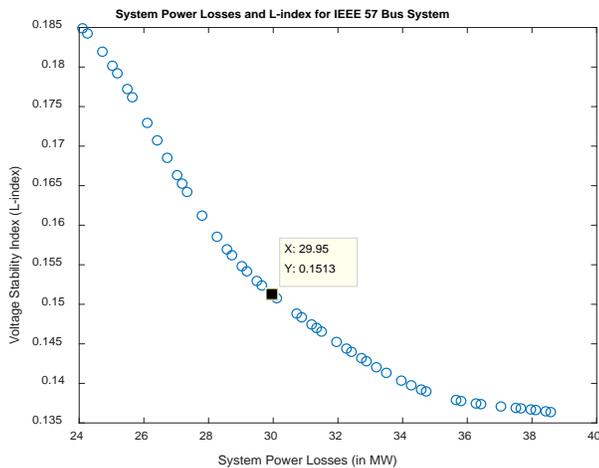


Fig. 4. Pareto Optimal Front for Case 3 of IEEE 57 Bus System.

IV. CONCLUSIONS

This paper proposed solution to the multi-objective based nonlinear, discrete ORPD problem. System power losses and L-index/VSEI are selected as objective functions, and they are solved by using the CSA. This work also considers the discrete variables, i.e., settings of transformer taps and switchable VAR/shunt capacitors. For this ORPD problem, these two objectives are observed as the two conflicting objectives and there is a need to solve these two objectives jointly. Therefore, the MO-CSA algorithm is used to solve this multi-objective ORPD problems. The proposed MO-CSA gives the Pareto optimal set with the trade-off/best-compromised solutions, in a single run. System operator selects one trade-off solution using the higher level information or by fuzzy satisfaction method. The solution of ORPD problem gives optimum values of objective function and control variable values. The proposed CSA and MO-CSA algorithms are examined on 30 and 57 bus systems. The obtained results depict the robustness, superiority, and effectiveness over other algorithms reported in the literature.

ACKNOWLEDGMENT

This research work was funded by “Woosong University’s Academic Research Funding – 2021”.

REFERENCES

- [1]. K. Nuaekaw, P. Artrit, N. Pholdee, S. Bureerat, “Optimal reactive power dispatch problem using a two-archive multi-objective grey wolf optimizer”, *Expert Systems with Applications*, vol. 87, pp. 79-89, 2017.
- [2]. N. Wang, J. Li, X. Yu, D. Zhou, W. Hu, Q. Huang, Z. Chen, F. Blaabjerg, “Optimal active and reactive power cooperative dispatch strategy of wind farm considering levelised production cost minimization”, *Renewable Energy*, vol. 148, pp. 113-123, 2020.
- [3]. Y. Muhammad, R. Khan, M.A.Z. Raja, F. Ullah, N.I. Chaudhary, Y. He, “Solution of optimal reactive power dispatch with FACTS devices: A survey”, *Energy Reports*, vol. 6, pp. 2211-2229, 2020.
- [4]. M. Ettappan, V. Vimala, S. Ramesh, V.T. Kesavan, Optimal reactive power dispatch for real power loss minimization and voltage stability enhancement using Artificial Bee Colony Algorithm, *Microprocessors and Microsystems*, vol. 76, 2020.
- [5]. M. Lakshmi, A. Ramesh Kumar, “Optimal Reactive Power Dispatch using Crow Search Algorithm”, *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 3, pp. 1423-1431, Jun. 2018.

- [6]. M.A.M. Shaheen, H.M. Hasanien, A. Alkuhayli, “A novel hybrid GWO-PSO optimization technique for optimal reactive power dispatch problem solution”, *Ain Shams Engineering Journal*, vol. 12, no. 1, pp. 621-630, 2021.
- [7]. S. Sayah, “Modified differential evolution approach for practical optimal reactive power dispatch of hybrid AC–DC power systems”, *Applied Soft Computing*, vol. 73, pp. 591-606, 2018.
- [8]. R. Jamal, B. Men, N.H. Khan, M.A.Z. Raja, Y. Muhammad, “Application of Shannon Entropy Implementation Into a Novel Fractional Particle Swarm Optimization Gravitational Search Algorithm (FPSOGSA) for Optimal Reactive Power Dispatch Problem”, *IEEE Access*, vol. 9, pp. 2715-2733, 2021.
- [9]. N.H. Khan, Y. Wang, D. Tian, M.A.Z. Raja, R. Jamal, Y. Muhammad, “Design of Fractional Particle Swarm Optimization Gravitational Search Algorithm for Optimal Reactive Power Dispatch Problems”, *IEEE Access*, vol. 8, pp. 146785-146806, 2020.
- [10]. Y. Zhang, Z. Ren, “Optimal reactive power dispatch considering costs of adjusting the control devices”, *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1349-1356, Aug. 2005.
- [11]. M. Ebeed, A. Alhejji, S. Kamel, F. Jurado, “Solving the Optimal Reactive Power Dispatch Using Marine Predators Algorithm Considering the Uncertainties in Load and Wind-Solar Generation Systems”, *Energies*, vol. 13, 2020.
- [12]. V. Suresh, S.S. Kumar, “Optimal reactive power dispatch for minimization of real power loss using SBDE and DE-strategy algorithm”, *Journal of Ambient Intelligence and Humanized Computing*, 2020.
- [13]. Y. Muhammad, R. Akhtar, R. Khan, F. Ullah, M.A.Z. Raja, J.A.T. Machado, “Design of fractional evolutionary processing for reactive power planning with FACTS devices”, *Scientific Reports*, vol. 11, 2021.
- [14]. T.L. Duong, M.Q. Duong, V.D. Phan, T.T. Nguyen, “Optimal Reactive Power Flow for Large-Scale Power Systems Using an Effective Metaheuristic Algorithm”, *Journal of Electrical and Computer Engineering*, vol. 2020, pp. 1-11, 2020.
- [15]. K. Nagarajan, A.K. Parvathy, A. Rajagopalan, “Multi-Objective Optimal Reactive Power Dispatch using Levy Interior Search Algorithm”, *International Journal on Electrical Engineering and Informatics*, vol. 12, no. 3, pp. 547-570, Sept. 2020.
- [16]. M. Zhang, Y. Li, “Multi-Objective Optimal Reactive Power Dispatch of Power Systems by Combining Classification-Based Multi-Objective Evolutionary Algorithm and Integrated Decision Making”, *IEEE Access*, vol. 8, pp. 38198-38209, 2020.
- [17]. M. Basu, “Multi-objective optimal reactive power dispatch using multi-objective differential evolution”, *International Journal of Electrical Power & Energy Systems*, vol. 82, pp. 213-224, 2016.
- [18]. M.S. Saddique, A.R. Bhatti, S.S. Haroon, M.K. Sattar, S. Amin, I.A. Sajjad, S.S. Haq, A.B. Awan, N. Rasheed, “Solution to optimal reactive power dispatch in transmission system using meta-heuristic techniques—Status and technological review”, *Electric Power Systems Research*, vol. 178, 2020.
- [19]. S. Jaiswal, T. Ghose, “Optimal real power dispatch of centralized micro-grid control operation”, *International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, Kollam, India, 2017, pp. 1-7.
- [20]. M. Hashemi, M.H. Zarif, “A novel hierarchical distributed framework for optimal reactive power dispatch based on a system of systems structure”, *Computers & Electrical Engineering*, vol. 78, pp. 162-183, 2019.
- [21]. M. Alramlawi, E. Mohagheghi, P. Li, “Predictive active-reactive optimal power dispatch in PV-battery-diesel microgrid considering reactive power and battery lifetime costs”, *Solar Energy*, vol. 193, pp. 529-544, 2019.
- [22]. P.P. Biswas, P.N. Suganthan, R. Mallipeddi, G.A.J. Amarantunga, “Optimal reactive power dispatch with uncertainties in load demand and renewable energy sources adopting scenario-based approach”, *Applied Soft Computing*, vol. 75, pp. 616-632, 2019.
- [23]. SK.M. Shareef, R.S. Rao, “Optimal reactive power dispatch under unbalanced conditions using hybrid swarm intelligence”, *Computers & Electrical Engineering*, vol. 69, pp. 183-193, 2018.

- [24]. R.N.S. Mei, M.H. Sulaiman, Z. Mustafa, H. Daniyal, "Optimal reactive power dispatch solution by loss minimization using moth-flame optimization technique", *Applied Soft Computing*, vol. 59, pp. 210-222, 2017.
- [25]. J. Li, N. Wang, D. Zhou, W. Hu, Q. Huang, Z. Chen, F. Blaabjerg, "Optimal reactive power dispatch of permanent magnet synchronous generator-based wind farm considering levelised production cost minimization", *Renewable Energy*, vol. 145, pp. 1-12, 2020.
- [26]. A.M. Shaheen, R.A. El-Sehiemy, S.M. Farrag, "Optimal reactive power dispatch using backtracking search algorithm", *Australian Journal of Electrical and Electronics Engineering*, vol. 13, no. 3, pp. 200-210, 2016.
- [27]. S.R. Salkuti, "Optimal Reactive Power Scheduling Using Cuckoo Search Algorithm", *International Journal of Electrical and Computer Engineering*, vol. 7, no. 5, pp. 2349-2356. Oct. 2017.
- [28]. T.T. Nguyen, D.N. Vo, H.V. Tran, L.V. Dai, "Optimal Dispatch of Reactive Power Using Modified Stochastic Fractal Search Algorithm", *Complexity*, vol. 2019, pp. 1-28, 2019.
- [29]. Y. Zhang, C. Chen, C. Lee, "Solution of the optimal reactive power dispatch for power systems by using novel charged system search algorithm," *7th International Symposium on Next Generation Electronics (ISNE)*, Taipei, Taiwan, 2018, pp. 1-4.
- [30]. P.K. Roy, S.P. Ghoshal, S.S. Thakur, "Optimal Reactive Power Dispatch Considering Flexible AC Transmission System Devices Using Biogeography-based Optimization", *Electric Power Components and Systems*, vol. 39, no. 8, pp. 733-750, 2011.
- [31]. A. Meddeb, N. Amor, M. Abbas, S. Chebbi, "A Novel Approach Based on Crow Search Algorithm for Solving Reactive Power Dispatch Problem", *Energies*, vol. 11, 2018.
- [32]. A. Saha, A. Bhattacharya, P. Das, A.K. Chakraborty, "Crow search algorithm for solving optimal power flow problem", *Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, 2017, pp. 1-8.
- [33]. S.R. Salkuti, Multi-objective based Optimal Network Reconfiguration using Crow Search Algorithm, *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 12, no. 3, pp. 86-95, 2021.
- [34]. S.S. Reddy, P.R. Bijwe, A.R. Abhyankar, "Faster Evolutionary Algorithm Based Optimal Power Flow using Incremental Power Flow Model", *International Journal of Electrical Power & Energy Systems*, vol. 54, pp. 198-210, 2014.
- [35]. M.S. Kumari, S. Maheswarapu, "Enhanced Genetic Algorithm based computation technique for multi-objective Optimal Power Flow solution", *International Journal of Electrical Power & Energy Systems*, vol. 32, no. 6, pp. 736-742, 2010.
- [36]. S.S. Reddy, P.R. Bijwe, "Differential evolution-based efficient multi-objective optimal power flow", *Neural Computing and Applications*, vol. 31, no. 1, pp. 509-522, Jan. 2019.
- [37]. M. Mehdinejad, B.M. Iyatloo, R.D. Bonab, K. Zare, "Solution of optimal reactive power dispatch of power systems using hybrid particle swarm optimization and imperialist competitive algorithms", *International Journal of Electrical Power & Energy Systems*, vol. 83, pp. 104-116, 2016.
- [38]. S.R. Salkuti, "Optimal power flow using hybrid differential evolution and harmony search algorithm", *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 5 pp. 1077-1091, May 2019.
- [39]. C.M. Huang, Y.C. Huang, "Combined Differential Evolution Algorithm and Ant System for Optimal Reactive Power Dispatch", *Energy Procedia*, vol. 14, pp. 1238-1243, 2012.
- [40]. S.S. Reddy, Ch.S. Rathnam, "Optimal Power Flow Using Glowworm Swarm Optimization", *International Journal of Electrical Power & Energy Systems*, vol. 80, pp. 128-139, 2016.
- [41]. S. Mouassa, T. Bouktir, A. Salhi, "Ant lion optimizer for solving optimal reactive power dispatch problem in power systems", *Engineering Science and Technology, an International Journal*, vol. 20, no. 3, pp. 885-895, 2017.

Controlling a Wheelchair using a Brain Computer Interface based on User Controlled Eye Blinks

Sebastián Poveda Zavala¹, Sang Guun Yoo^{2*}

Facultad de Ingeniería de Sistemas
Escuela Politécnica Nacional
Quito, Ecuador
Smart Lab
Escuela Politécnica Nacional
Quito, Ecuador

David Edmigio Valdivieso Tituana³
Facultad de Ingeniería Eléctrica y Electrónica
Escuela Politécnica Nacional
Quito, Ecuador
Smart Lab
Escuela Politécnica Nacional
Quito, Ecuador

Abstract—Data published by different organizations such as the United Nations indicates that there are a large number of people who suffer from different types of movement disabilities. In many cases, the disability is so severe that they cannot have any kind of movements. Faced with this situation, Brain Computer Interface technology has taken up the challenge of developing solutions that allow delivering a better quality of life to those people, and one of the most important areas has been the mobility solutions, which includes the brain computer interface enabled electric wheelchairs as one of the most helpful solutions. Faced with this situation, the present work has developed a Brain Computer Interface solution that allows users to control the movement of their wheelchairs using the brain waves generated when blinks their eyes. For the creation of this solution, the Incremental Prototyping methodology has been used to optimize the development process by generating independent modules. The solution is made up of several components i.e. EEG System (OpenBCI), Main Controller, Wheelchair Controller and Wheelchair that allows to have a modularity to carry out updates (improvements) of their functionalities in a simple way. The developed system has shown that it requires a low amount of training time and has a real world applicable response time. Experimental results show that the users can perform different tasks with an acceptable grade of error in a period of time that could be considered as acceptable for the system. Considering that the prototype was created for people with disabilities, the system could grant them a certain level of independence.

Keywords—BCI; EEG; brain computer interface; OpenBCI; eye blink detection

I. INTRODUCTION

Brain Computer Interface (BCI) is a technology in which the brain waves of a person are used to control an external object. It acts as a mediator between the brain and a device [1]. BCI based systems record the electrical activity of a human brain via different technologies, such as electroencephalography (EEG) [2], functional magnetic resonance imaging (fMRI) [3], and functional near infrared (fNIR) [4]. Among these alternatives, EEG is the most popular solution because of its low cost, high temporal resolution, and non-invasiveness features [5]. This technology can be applied in different fields, being one of the most important the health area [6][7][8]. BCI brings a capable and efficient way of aiding users with Motor Neuron Diseases (MND) such as Amyotrophic Lateral

Sclerosis (ALS), Progressive Bulbar Palsy (PBP), Primary lateral sclerosis (PLS), among others.

According to the United Nations, the number of individuals in the world above the age of 60 is expected to increase rapidly, and this is especially significant in developing countries where the proportion of older individuals will increase from 9% in 2015 to 16% by 2040 [9]. Additionally, according to the ALS association, most people develop ALS between the ages of 40 and 70, with an average age of 55 at the time of its diagnosis [10]. In many cases, the disability is so severe that the patients cannot have any kind of movements.

Faced with this situation, BCI technology has taken up the challenge of developing solutions that allow delivering a better quality of life to those people, and one of the most important areas has been the mobility technologies, which includes the BCI enabled electric wheelchairs as one of the most helpful solutions. In this situation, the present work has proposed and developed a system for controlling an electric wheelchair utilizing the brain signals generated by the user when blinks his/her eyes. The developed prototype will allow people with severe movement disabilities to have a higher level of independence in terms of mobility.

In summary, the objective of the present work is to design, develop and test a BCI system prototype which will allow people with movement disabilities to control an electric wheelchair using the brain signals generated when they blink their eyes.

The rest of this work is organized as follows. First, the previous works related to BCI wheelchair systems are studied. Then, the proposed solution and the used methodology are described with details. Once, explained the architecture and implementation of the proposed solution, the results of the experiments performed using the implemented prototype are analyzed. Finally, the conclusions reached in the present work are presented.

II. BACKGROUND

In the last decade, a variety of BCI technologies has been developed to help people with MND. Some of these technologies have allowed to upgrade the movement of people improving their lifestyles and bringing back their sense of

*Corresponding Author

independence which are considered as navigation assistive technologies [11]. Among the different assistive technologies, one of the solutions that could help to people with movement disabilities is the wheelchairs controlled by BCI systems. In this aspect, several alternatives for controlling wheelchairs have been proposed so far.

Turnip et al. have proposed an electric wheelchair control system utilizing Steady state visually evoked potential (SSVEP) responses. This is done by classifying 6-9Hz frequencies which are used to execute the turn left, turn right, go back, and go forward commands. These signals are processed and classified using a sixth order Bandpass Filter (BPF) and an Adaptive Neuro-Fuzzy Inference System (ANFIS) [12].

In another similar work, Li et al. developed a hybrid system for controlling a wheelchair combining P300 and SSVEP. The control is done by positioning 4 groups of flickering buttons; each group contains a large button in the center and 8 smaller surrounding it which flashes at a fixed frequency with the objective to evoke SSVEP responses. At the same time, the buttons were intensified through shape and color change in a random order to produce P300. Based on these two classifications, an output is reached. To classify both signals at the same time, they used different methods; in the P300 case, they used Support Vector Machine (SVM) and for the SSVEP, they used power ratios [13].

Additionally, Kim et al prototyped a BCI based system that allows the control of a wheelchair using motor imagery. They utilized five commands (left, left-diagonal, right, right-diagonal and forward) which are detected using a classifier that detects the imagination of left or right hand movements whilst the user simultaneously performs foot movement imagination. The signals are first spatially filtered using a common average reference and then band-pass filtered between 8 and 32 [14].

On the other hand, Y. Yu et al. proposed a hybrid BCI wheelchair system where combines motor imagery and P300 potential [15], while K. Sureshbai et al. proposed a SSVEP based BCI wheelchair where SSEVEP responses are generated by looking at four LEDs which are used to deliver forward, reverse, left, and right movements of an electrical wheelchair [16].

In another work, a novel EEG brain-controlled wheelchair with five steering behaviors is developed, which is based on two-stage control strategies combining sustained and brief motor imagery BCIs [17].

As shown in the previous works, there are a large number of BCI enabled wheelchair solutions. However, most of them uses dense techniques such as SSVEP and Motor Imagery which brings several limitations. First of all, in solutions based on SSVEP responses, the user has to look at several visual stimuli commonly implemented on blinking LEDs which could cause severe fatigue when the wheelchair has to be used by the user for a considerable period of time [18].

On the other hand, the Motor Imagery requires long time of training by users since it is based on brain signals generated by the user when simulates a physical action with his/her mind; additionally, the motor imagery commonly does not deliver high level of precision in the signal classification [19][20].

In this situation, the present work has the intention of developing a BCI enabled wheelchair which is controlled by using the brain signals generated when the user blinks his/her eyes. This method of brain signal classification has the advantage of having high level of precision and easiness of use; additionally, the solution has the benefits of not requiring a large amount of time in training processes [21, 22].

III. METHODOLOGY

In this work, Incremental Prototyping Methodology was applied for the development of the proposed solution [23]. This approach is useful when you want to create a quick prototype that is thoroughly and easily tested. In our case, we worked with two smaller increments, the first was the main controller, which manages the user interface and the processing of brain waves. The second increment (module of the prototype) was the wheelchair controller, which handles the movement of the wheelchair. The implementation of both modules creates a BCI system that can handle the movements of the wheelchair (see Fig. 1).

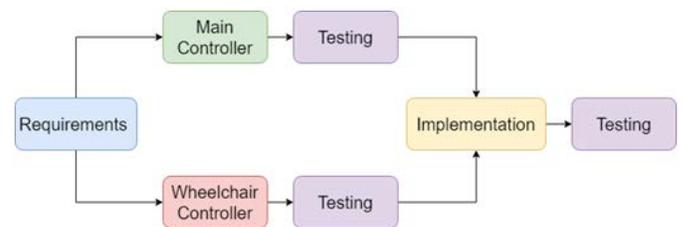


Fig. 1. Diagram of the Applied Methodology.

IV. PROPOSED SOLUTION

A. Architecture of the Proposed Solution

Since the purpose of the present work is to generate a BCI solution for controlling an electrical wheelchair, we have generated a general architecture of the prototype to be developed (see Fig. 2). It is also important to indicate that the proposed solution will use the brain signals generated when the user blinks his/her eyes. This method of brain signal classification has the advantage of having high level of precision, easiness of use, and low amount of time in training processes [21, 22].

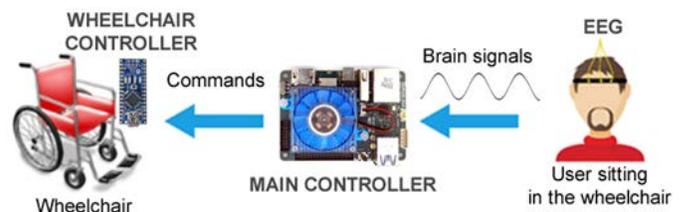


Fig. 2. General Architecture of the Proposed BCI System.

In the proposed architecture, the brain signals of the user are gathered by the EEG device connected to the user, and such data is transmitted to the Main Controller which executes a BCI application that filters and classifies the raw brain signals. Once the brain signals are classified into the commands that the user wants to perform, this data is communicated to the Wheelchair Controller, which generates the movement of the wheelchair.

B. Implementation of the Proposed Solution

The architecture shown in the previous section has been implemented as follows (see Fig. 3). In the proposed system, the OpenBCI V3 Cyton biosensing board was used to acquire the brain signals. We have chosen this EEG device for its advantages over other EEG systems such as modularity, relatively low cost, open architecture, and compatibility with different research platforms such as OpenViBE, EEGLAB, BCILAB and MATLAB, and since it provides APIs for different software development languages such as Python, C# and Java [24]. Once acquired the signals, such data is sent to the Main Controller which has a Single Board Computer (SBC) called Odroid XU-4 [25] as its main component. This component has the responsibility of performing the feature extraction and feature translation processes. The SBC also has the function of communicating with the Wheelchair Controller which was implemented using an Arduino Nano. In addition, the SBC also implements the functionality of showing the user interface through an LCD Screen.

1) *EEG system:* As indicated previously, the EEG signals (brain signals) are obtained by using the Cyton biosensing board. We also have used one active reusable flap snap electrode (TDE-202) located on the surface of the scalp (FP1) following the extended 10-20 system [26][27] with a sampling frequency of 250 Hz. Once the signals are gathered by the Cyton biosensing board, they are sent to the Main Controller using the Bluetooth communication. On the other side, the signals are received by the OpenBCI dongle of the Main Controller.

2) *Main controller:* This component has the function of executing the BCI application responsible for processing the signal, delivering the commands to the wheelchair controller, and working as the user interface of the solution. The main component of the controller is the SBC called Odroid XU-4 which implements the Ubuntu 18.04 operating system. The user interface of the application is shown in a 7-inch HDMI display-C with a resolution of 1046 x 600. We have chosen this SBC because it provides higher processing capabilities compared to traditional SBCs such as Raspberry Pi while maintaining the small size [19].

The user interface has been developed with a simple design to maintain an efficient interaction with the user (see Fig. 4). The interface shows to the user when the wheelchair is in movement and which direction it is going, as well as the commands that the user is inputting. This software was developed using an open-source framework called Electron which allows the creation of lightweight cross-platform applications using JavaScript, HTML and CSS.

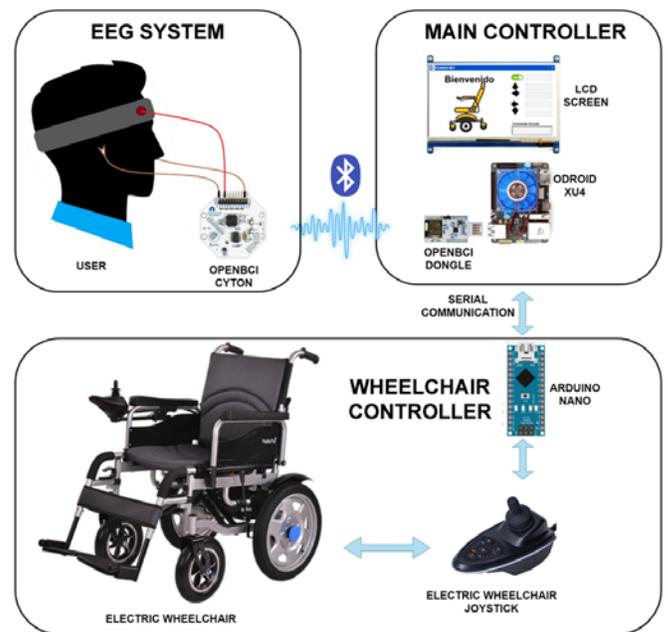


Fig. 3. Complete Diagram of the Proposed BCI System.



Fig. 4. Interface of the Developed Application.

As indicated in the general architecture, the SBC of the Main Controller includes a BCI application that receives the brain signals of the user to then convert the signals into wheelchair controlling commands. Once classified the commands, this data is sent to the Wheelchair Controller. The process of converting the raw brain signals into understandable wheelchair controlling commands are executed into the following steps:

a) *Feature extraction:* Once received the raw brain signals, the application applies a 0.5 Hz high pass filter, a 50 Hz notch filter, and a 5-25 Hz band pass filter, following the directions delivered from one of our previous works [21]. Fig. 5 shows the difference between the raw data and the filtered data.

b) *Feature translation:* After executing the feature extraction process, the acquired signals are classified into two types of blinks i.e. short blinks and composed long blinks [22]. According to S. Poveda Zavala et al. a short blink is detected when a user blinks normally; when a short blink is

performed by the user, the processed signal drops below -75 uV and it rises above 85 uV in a small period of time of 1.5 seconds (see Fig. 6). On the other side, the signals are classified as a composed long blink when the user closes their eyes from 2 to 5 seconds, opens them and then blinks quickly again; when this happens, the signal drops below -75 uV, it then returns to a normal state (between -25 and 25 uV), rises above 100 uV and then drops below -75 uV, all of this in a time frame of 2 to 5 seconds (see Fig. 7).

c) *Signal classification:* The classification of the signal is done by using a software developed in Python 3.7. We have chosen this programming language since it is compatible with the OpenBCI software library.

Once established the usage of short blinks and composed long blinks, the present work has generated an encoding system to classify a combination of different type of blinks into the commands that will be used to control the movements of the wheelchair. The proposed encoding system is shown in Table I, where a short blink is represented as a dot (.) and a composed long blink is represented as a dash (-). At this point, it is important to indicate that, since this encoding system must not interfere with normal blinking patterns of the user, the present work has implemented the usage of four consecutive composed long blinks as the activation/deactivation command for the whole system. This combination was selected for this purpose since this pattern is not presented in a normal blink sequence of a person.

TABLE I. ENCODING SYSTEM USED IN THE PROPOSED SOLUTION

Command	Description	Blink Combination
Activate /Deactivate	Activates or deactivates the system	----
Go forward	Moves the wheelchair forward for 1 second	---
Turn left	Turns the wheelchair left for 1 seconds	---.
Turn right	Turns the wheelchair right for 1 seconds	---.
Go backwards	Moves the wheelchair backwards for 1 second	---

3) *Wheelchair controller:* To control the electric wheelchair, it was necessary to rewire the control scheme of the device's joystick (see Fig. 8). The controller of the joystick includes an input device that allows to send information about the angle or direction of the movement of the wheelchair.

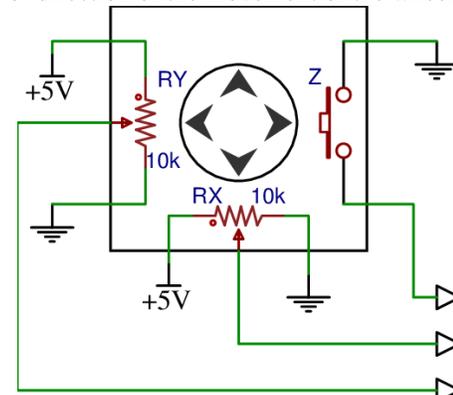


Fig. 8. Resistance Diagram of the Wheelchair's Joystick.

This device is powered by a 5 volts direct current. The operation is based on a variable resistor, which changes the output voltage depending on the movement of the joystick lever. In idle state, the outgoing voltage signals from the Y to X axis are 2.5 volts. The joystick sends the voltage signals depending on the position of the lever (see Fig. 9). The wheelchair controller has reproduced the aforementioned signals by using a system created in an Arduino Nano.

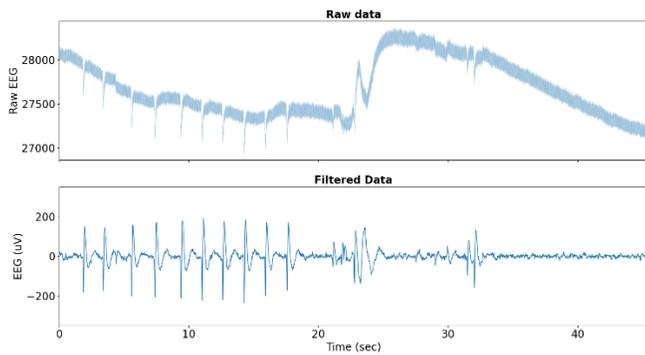


Fig. 5. Raw Extracted Data vs Filtered Data.

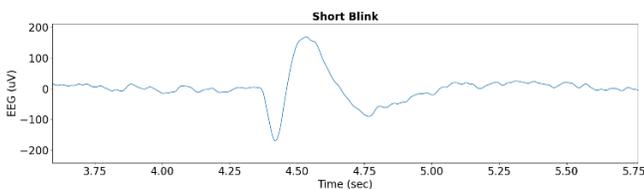


Fig. 6. Signal Generated when the user Performs a Short Blink.

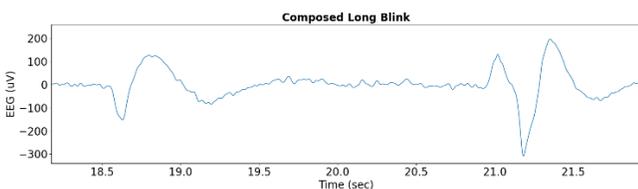


Fig. 7. Signal Generated when the user Performs a Composed Long Blink.

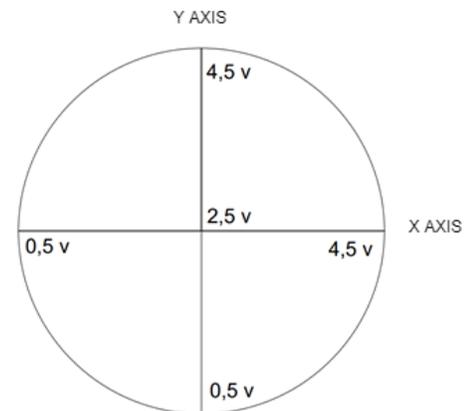


Fig. 9. Signal Diagram.

Since the Arduino Nano does not have a Digital to Analog Converter (DAC) system and since it only can generate Pulse Width Modulation (PWM) signals, the Wheelchair Controller has implemented an RC filter to obtain a DC voltage signal with low ripple (see Fig. 10). This solution allowed the generation of 0-to-5-volt signals through the analog terminal present in the Arduino Nano. The connection between the wheelchair power pins (see Fig. 11) and the Arduino Nano cables are present in the Table II and Fig. 12.

The Arduino Nano was powered by using the VCC and GND terminals of the wheelchair control system. The Arduino Nano receives the commands from the Main Controller through a serial communication. Once received the commands, the Arduino Nano interprets them and move the wheelchair to different directions. To simulate the real signals that the variable resistors deliver, a voltage ramp was implemented, which also allows to avoid sudden movements in the wheelchair. The commands that control the movement of the wheelchair are present in the following Table III.

TABLE II. CONNECTIONS BETWEEN WHEELCHAIR POWER PIN AND MICROCONTROLLER (ARDUINO NANO) CABLES

Wheelchair control system connections	Arduino cable connections
X AXIS	Blue cable
GND	White-orange cable
VCC	Orange cable
Y AXIS	White-blue Cable

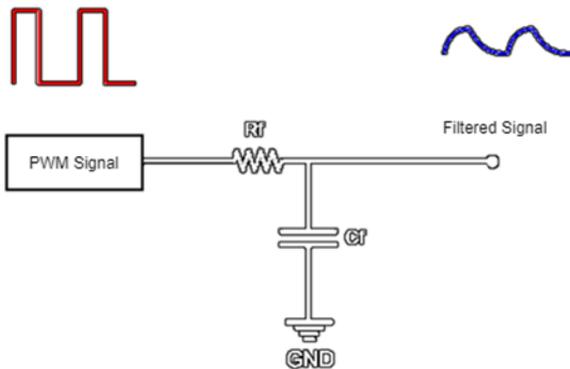


Fig. 10. PWM Signal Filtering.

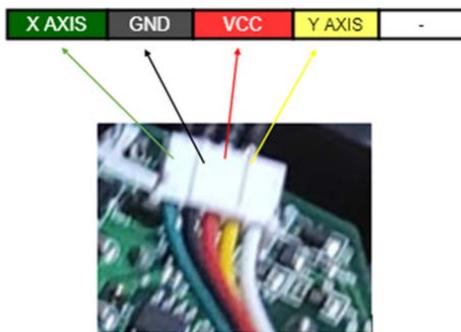


Fig. 11. Wheelchair Control System Connections.



Fig. 12. Power System - Microcontroller Connection.

TABLE III. COMMANDS USED IN WHEELCHAIR CONTROLLER

Command	Data sent by the Main Controller
Forward	w
Backwards	x
Right	d
Left	a

The amount of time that the wheelchair realizes its movements can be customized depending on user requirements. In the present prototype, such variable was set with the period of one second since it was the most beneficial frame for the experimentation process. The speed at which the wheelchair moves can also be configured, in the present prototype; such variable was set with the slowest option in order to guarantee the safety of the user.

V. EXPERIMENTATION AND RESULTS

For the testing of the developed prototype, ten 20 to 45-year-old healthy people participated in the experiment. Before the experiment, each participant used the system for a total of 2 hours to get accustomed to its functionality.

A. Experiment

The aim of the experiment was to test the speed in which the subjects could perform two tasks (see Fig. 13). In the first task, the participants would have to complete the following steps: (1) activate the system, (2) move forward, (3) turn right, (4) move forward and (5) deactivate the system. On the other hand, the second task consisted of the following steps: (1) activate the system, (2) turn right, (3) move forward, (4) turn left, (5) move forward, (6) turn left, (7) move forward, (8) turn right, (9) move forward, and (10) deactivate the system.

We have asked the participants to complete these two tasks 10 times to validate the consistency of the results. If the participant has failed completing a command, he would start all over again and have one attempt added to his total for that task, all participants would start with 0 attempts. Table IV shows the average number of attempts the participant needed to complete the first task. We believe that 0.82 attempts at 64.73 seconds

for the first task and 0.85 attempts at 140.59 seconds for the second task are acceptable for this developed prototype. The working prototype is shown in Fig. 14.

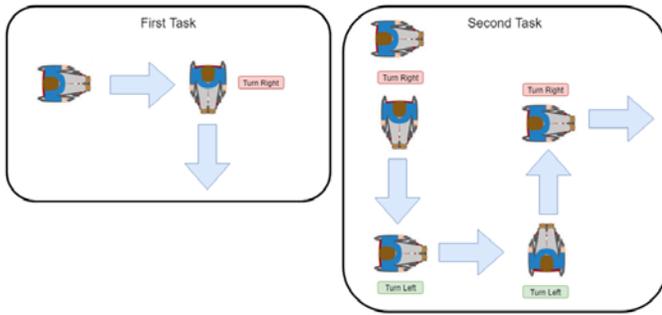


Fig. 13. First and Second Tasks Performed during the Experimentation Process.

TABLE I. AVERAGE NUMBER OF ATTEMPTS AND TIME PER ATTEMPT FOR TASK 1 AND 2.

Particip- pant	Task 1		Task 2	
	Average Attempts	Average Elapsed Time (s)	Average Attempts	Average Elapsed Time (s)
1	0.5	61.25	0.9	134.40
2	1.2	67.26	1.0	144.64
3	0.6	62.50	0.6	129.54
4	0.7	60.50	0.8	132.57
5	1.3	70.16	1.2	161.28
6	0.8	63.56	1.0	155.64
7	0.8	64.98	0.5	125.09
8	0.2	60.04	0.6	134.54
9	1.5	72.89	0.9	138.23
10	0.6	64.14	1.0	149.89
Average Overall	0.82	64.73	0.85	140.58



Fig. 14. Working Prototype.

VI. CONCLUSIONS

In the present work, a BCI based wheelchair control system utilizing eye blinks was developed. The solution is made up of several components i.e. EEG System, Main Controller, Wheelchair Controller and Wheelchair that allows having a modularity to carry out updates (improvements) of their functionalities in a simple way. The developed system has shown that it requires a low amount of training time and has a

real world applicable response time. Experimental results show that the users can perform different tasks with an acceptable grade of error in a period of time that could be considered as acceptable for the system. Considering that it was created for people with disabilities (e.g. people that suffer from MND), the system could grant them a certain level of independence. As future work, we will work in analyzing and classifying other brain signals for developing new solutions for people with disabilities.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by the Escuela Politécnica Nacional, for the development of the project PVS-2018-022 – “Silla de ruedas eléctrica controlado por ondas cerebrales”.

REFERENCES

- [1] A. Vallabhaneni, T. Wang, B. He, “Brain—Computer Interface,” *Neural Engineering*, 2005, pp. 85-121.
- [2] R. Abiri et al., “A comprehensive review of EEG-based brain–computer interface paradigms,” *Journal of Neural Engineering*, Vol 16, Number 1, 2019.
- [3] J. Simon., P. Fishbein, I. Zhu, M. Roberts, I. Martin, “Functional Magnetic Resonance Imaging-Based Brain Computer Interfaces,” *Neural Interface Engineering*. 2020. Doi: 10.1007/978-3-030-41854-0_.
- [4] N. Naseer, K.S. Hong, “fNIRS-based brain-computer interfaces: a review,” *Frontiers in Human Neuroscience*, 9:3, 2015. doi: 10.3389/fnhum.2015.00003.
- [5] B. He, S. Gao, H. Yuan, and J. R. Wolpaw, “Brain–computer interfaces,” in *Neural Engineering*, Springer US, 2013, pp. 87–151.
- [6] R.M. Aileni, G. Suciuc, V. Suciuc, J. Ciurea, P. Sever, “Assistive Mobile Technologies for Health Monitoring and Brain–Computer Interface for Patients with Motor Impairments,” *Mobile Solutions and Their Usefulness in Everyday Life*, EAI/Springer Innovations in Communication and Computing. Springer, Cham, 2019, doi: 10.1007/978-3-319-93491-4_11.
- [7] J. Cantillo-Negrete et al., “Motor Imagery-Based Brain-Computer Interface Coupled to a Robotic Hand Orthosis Aimed for Neurorehabilitation of Stroke Patients,” *Journal of Healthcare Engineering*, Vol. 2018, Article ID 1624637, 2018. doi: 10.1155/2018/1624637.
- [8] C. Lim et al., “A randomized controlled trial of a brain-computer interface based attention training program for ADHD,” *PLoS ONE*, Vol. 14(5): e0216225, 2019. doi: 10.1371/journal.pone.0216225, 2019.
- [9] United Nations, “World Population Ageing, 2014,” *Dep. Econ. Soc. Aff. Popul. Div.*, p. 73, 2014, [Online]. Available: <http://books.google.com/books?hl=en&lr=&id=9WoK26zWCyIC&pgis=1>.
- [10] ALS Association, “Who Gets ALS? - The ALS Association,” p. 90, 2011, [Online]. Available: <http://www.alsa.org/about-als/who-gets-als.html>.
- [11] S. Poveda Zavala, J. Luis Murillo López, K. Ortíz Chicaiza, and S. Guun Yoo, “Review of Steady State Visually Evoked Potential Brain-Computer Interface Applications: Technological Analysis and Classification,” *J. Eng. Appl. Sci.*, vol. 15, no. 2, 2019, pp. 659–678. doi: 10.36478/jeasci.2020.659.678.
- [12] M. Turnip et al., “An application of online ANFIS classifier for wheelchair based brain computer interface,” *Proc. 2015 Int. Conf. Autom. Cogn. Sci. Opt. Micro Electro-Mechanical Syst. Inf. Technol. ICACOMIT 2015*, pp. 134–137. doi: 10.1109/ICACOMIT.2015.7440192.
- [13] Y. Li, J. Pan, F. Wang, Z. Yu, “A Hybrid BCI System Combining P300 and SSVEP and Its Application to Wheelchair Control,” *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, 2013, pp. 3156–3166, doi: 10.1109/tbme.2013.2270283.
- [14] K.-T. Kim, T. Carlson, and S.-W. Lee, “Design of a Robotic Wheelchair with a Motor Imagery based Brain-Computer Interface,” *Int. Winter*

- Work. Brain-Computer Interface, 2013, pp. 46–48. doi: 10.1109/IWW-BCI.2013.6506625.
- [15] Y. Yu et al., “Self-Paced Operation of a Wheelchair Based on a Hybrid Brain-Computer Interface Combining Motor Imagery and P300 Potential,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 25: 12, 2017. doi: 10.1109/TNSRE.2017.2766365.
- [16] K. Sureshbai Mistry et al., “An SSVEP based brain computer interface system to control electric wheelchairs,” *Proceedings of 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, May 2018. doi: 10.1109/I2MTC.2018.8409632.
- [17] H. Wang, A. Bezerianos, “Brain-controlled wheelchair controlled by sustained and brief motor imagery BCIs,” *Electronic Letters*, Vol 53: 17, 2017, doi: 10.1049/el.2017.1637.
- [18] T. Cao et al., “Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain-computer interfaces,” *Biomed Eng Online*, 13(1):28, 2014. doi: 10.1186/1475-925X-13-28.
- [19] S. Abdulkader, A. Atia, M. Mostafa, “Brain computer interfacing: Applications and challenges,” *Egyptian Informatics Journal*, Vol. 16, Issue 2, 2015, pp. 213-230.
- [20] N. Padfield et al., “EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenge,” *Sensors*, Vol. 19(6): 1423, 2019. doi: 10.3390/s19061423.
- [21] S. Poveda Zavala et al., “EEG Signal Processing Model for Eye Blink Detection,” *J. Eng. Appl. Sci.*, vol. 15, no. 7, 2020, pp. 1671–1675. doi: 10.36478/jeasci.2020.1671.1675.
- [22] S. Poveda Zavala et al., “BCI Based Home Automation using User Controlled Blinks,” *J. Eng. Appl. Sci.*, vol. 15, no. 6, 2020, pp. 1377–1384. doi: 10.36478/jeasci.2020.1377.1384.
- [23] F. Kordon and J. Henkel, “An overview of rapid system prototyping today,” *Design automation for embedded systems*, vol. 8, 2003, pp. 275-282.
- [24] OpenBCI, “Open Source Biosensing Tools (EEG, EMG, EKG, and more),” 2019. [Online]. Available: <https://openbci.com/>.
- [25] ODROID-XU4, “ODROID,” 2019. [Online]. Available: <https://www.hardkernel.com/shop/odroid-xu4-special-price/>.
- [26] G. H. Klem, H. O. Lüders, H. H. Jasper, and C. Elger, “The ten-twenty electrode system of the International Federation,” *Electroencephalogr Clin Neurophysiol*, vol. 52, no. 3, 1999, pp. 3–6.
- [27] S. Dutta et al., “Development of a BCI-based gaming application to enhance cognitive control in psychiatric disorders,” *Innovations in Systems and Software Engineering*, Vol. 17, 2021, pp. 99 – 107.

IoT based Smart Water Quality Prediction for Biofloc Aquaculture

Md. Mamunur Rashid¹

Department of CSE, University of Liberal Arts Bangladesh
Dhaka, Bangladesh

Al-Akhir Nayan², Md. Obaidur Rahman⁵

Department of CSE, European University of Bangladesh
Dhaka, Bangladesh

Joyeta Saha⁴

Department of ECE, North South University
Dhaka, Bangladesh

Sabrina Afrin Simi³

Dept. of Human Computer Interaction
University of Siegen, Siegen, Germany

Muhammad Golam Kibria⁶

Department of CSE, IoT Lab, University of Liberal Arts
Bangladesh, Dhaka, Bangladesh

Abstract—Traditional fish farming faces several challenges, including water pollution, temperature imbalance, feed, space, cost, etc. Biofloc technology in aquaculture transforms the manual into an advanced system that allows the reuse of unused feed by converting them into microbial protein. The objective of the research is to propose an IoT-based solution to aquaculture that increases efficiency and productivity. The article presented a system that collects data using sensors, analyzes them using a machine learning model, generates decisions with the help of Artificial Intelligence (AI), and sends notifications to the user. The proposed system has been implemented and tested to validate and achieve a satisfactory result.

Keywords—Smart aquaculture system; biofloc technology; machine learning; life below water

I. INTRODUCTION

In biofloc aquaculture, it is inevitable to be more intelligent to monitor the water quality in real-time and feed accurately. However, due to real-time water quality monitoring, the balance of bacteria in the aquaculture environment might be harmed; hence fish's disease-resistant ability is decreased. It is impossible to measure the water quality accurately based on experience only [1, 2]. An intelligent system could help the farmers by reading the water parameters on time to monitor and maintain the quality accordingly. Hence, identifying the water parameters suitable for the biofloc aquaculture, a water quality prediction model for the dynamic changes in water parameters, and accordingly are essential.

Biofloc technology can reduce food costs, while a smart aquaculture system can reduce labor cost. It is a good option that is cheaper and beneficial to fish's health [3, 4]. Being a low-lying country, several natural calamities like floods, cyclones, etc., have a significant effect on aquaculture at both the ponds and marine waters. Fish farmers must bear a substantial loss due to the polluted water and increasing salinity of the coastal water for those disasters. Traditional fish farming leads to several other problems, such as water

pollution caused by carbon dioxide, ammonia, and nitrogen. External filtration is needed for detoxifying, which is costly and time-consuming. Biofloc technology is an excellent alternative to the cost-effective traditional aquaculture system. Biofloc itself helps to purify the water naturally, hence the use of external tools or ingredients might be reduced. Maintaining water quality can ensure increasing production, decreasing the death of fish. Water quality parameters are the most important factors to maintain a fish farm using Bioflocs.

This article mainly focuses on water level parameters. An automated system has been implemented to collect data through sensors, analyze them using a machine learning method, analyze the water quality. The application of IoT makes it easier to monitor the water and maintain the ecology in biofloc aquaculture.

In this article, data collected from the fish farm has been used for training and testing purposes. A machine learning method has been applied to develop the model. The water quality parameters, including pH, have been analyzed, and the correlation between them is obtained. The water quality prediction model is trained based on the collected data.

II. RELATED WORKS

Deep learning (DL) technology is used in numerous fields. X. Yang et al. focused on DL applications in aquaculture. They worked on identifying live and dead fish, classified species, performed behavioral analysis, and feeding decisions. The algorithm and the results of the method were applied to the smart fish farm. The findings showed that the deep learning method could extract features automatically. They made the most valuable contribution to the field of agriculture. But the technique was failed to address complex data in aquaculture [5].

S. Liu et al. did an experiment on "Ras Carpio" using the Recirculating Aquaculture System (RAS) [6]. In 2011, RAS was an intelligent alternative to traditional aquaculture in

ponds. The water parameters were being continually monitored, and whenever the parameter's value got out of the fish's versatile range, the water was recirculated. For this purpose, there were two drainage systems. DO, pH, and the temperature was monitored by WATT TriO Matic 700IQ (SW), WATT Sensolyt 700 IQ (SW), and WATT Tri oxyTherm type sensors. Though the system had many benefits and was replaced rapidly with a regular aquaculture system, it had some disadvantages. It needs water exchange which is a lengthy and costly process.

M. I. Dzulqornain et al. implemented an innovative IoT-based system on the IFTTT model [7]. They used dissolved oxygen, water temperature, the potential of hydrogen (pH) as parameters. The water level was sensed with sensors, and for controlling the system, an aerator system was utilized integrating with microcontroller NodeMCU v3, relay, power supply, and propeller. The sensor data was uploaded to the cloud, and the client could visualize them from anywhere. The application was both web-based and android-based. The system was well enough, but its process was manual.

A. A. Nayan et al. worked on measuring river water quality for agriculture and fishing purposes [8] and identified fish diseases by detecting the changes in water quality [9]. They used a machine learning technique that evaluated the water quality and processed intelligent suggestions. They utilized pH, DO, BOD, COD, TSS, TDS, EC, PO4³⁻, NO₃-N, and NH₃-N to calculate the water quality and predicted the result using boosting technique. But the study only focused on big water sources like rivers and did not suggest any solution for small water sources like ponds.

To minimize the gap of previous studies and to provide a better understanding of the current state of the art of DL in aquaculture, we have worked on this project. The work offers good support for deploying applications for smart fish farming which is entirely new compared with other results. An automatic system with Biofloc technology has been introduced. We have tried to decrease feed costs by reducing FCR (feed conversion ratio). The nutrients used in this technology can be recycled and reused easily.

III. MATERIALS AND METHODS

A. Biofloc Technology (BFT)

Biofloc is a new technology introduced in aquaculture for low-cost fish farming. Bioflocs are used to make reusable food from organic waste nutrients. It is a thin layer made up of beneficial bacteria, microorganisms, and algae that filters the water. Bacteria is cultured for this technology because it produces flocs or algae, breaking ammonia to minimize water pollution. The biofloc method of fish farming can be helpful to grow vegetables and fish together in the yard. This method requires tanks, oxygen supply pumps, and round-the-clock electricity. It needs less amount of food and reduces the chances of the disease. Being an eco-friendly system, it reduces the impact on the environment and improves productivity. Water must be exchanged to minimal or zero in this system. It is cost-effective by reducing the usage of protein-rich feed [10, 11, 12]. Fig. 1 shows a general image captured from a biofloc farm.



Fig. 1. Biofloc Technology in Aquaculture.

B. Hardware Components

IoT innovation has brought "Sensor Development" to a new stage. IoT systems operate and use a range of sensors to provide different kinds of information and data. It helps to gather information, drive, and distribute it to a network of similar gadgets. The collected data allows it possible for the devices to work autonomously, and every day the whole world turns to be "smarter." By integrating sensors, microcontrollers, and other smart gadgets, the project was implemented. We have used the following hardware components to run the project.

- Arduino UNO [13].
- White Breadboard.
- pH sensor [14].
- Temperature Sensor [15].
- Total Dissolved Solids (TDS) Sensor [16].
- Computer.
- Wires.

C. System Architecture

Fig. 2 shows the architecture of our Smart Aquaculture Water Monitoring System. The system monitors continuously and sends notifications through a Wi-Fi [17] module to an android application. The project's primary function is to check the water parameters: pH, temperature, and TDS. We have collected samples from different farms that use Biofloc technology. Processing data from the samples, we trained the artificial neural network. The sensors connected with the system provide a continuous reading of the water parameters. The system generates output evaluating the trained data and the reading provided by the sensors. It predicts the water quality, determines the situation, and process wise decision. It sends the result and decision as a notification to the user through an android application. Fig. 3 shows the flow diagram of the project.

D. Hardware Connection

We collected the required hardware, tools and connected them according to the diagram shown in Fig. 4. The figure mentions the pin diagram of Arduino Uno with Temperature sensor and pH sensor.

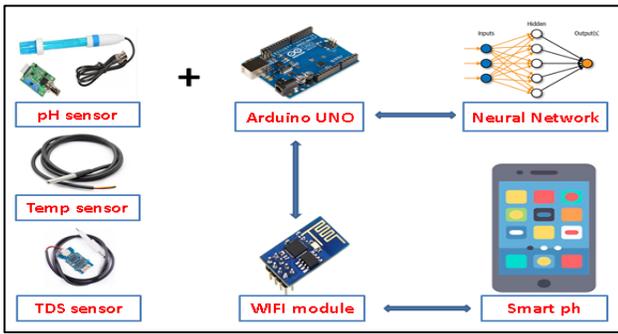


Fig. 2. System Architecture.

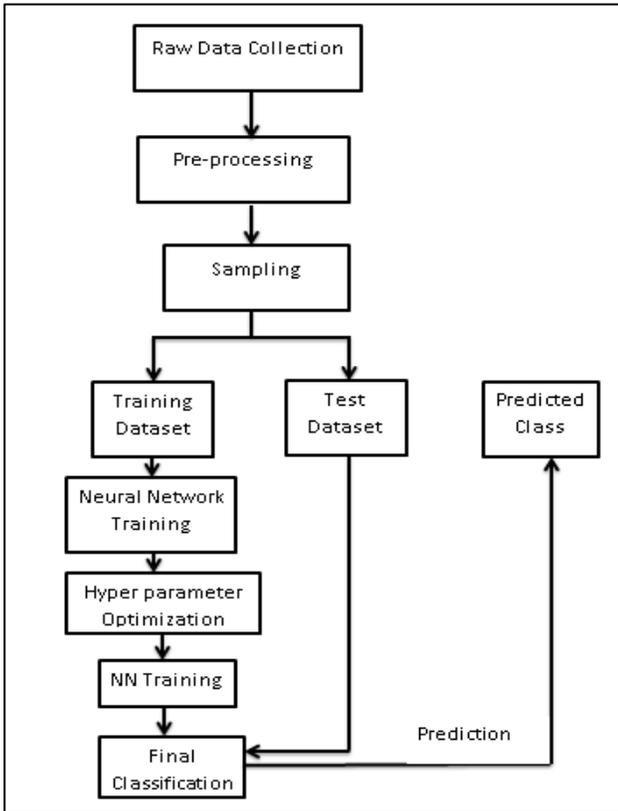


Fig. 3. Flow Diagram.

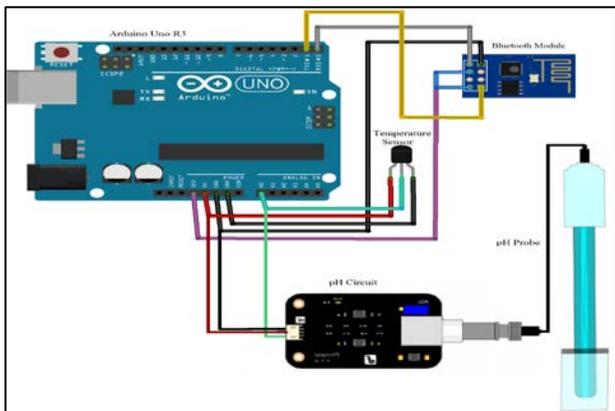


Fig. 4. Hardware Connection.

E. Study Area

Few farms have already adopted biofloc technology. We found many projects in Bangladesh and visited there. The first center named “Biotech Aquaculture” introduced the technology in Bangladesh. The center is situated at Dakshinkhan, Uttara, Dhaka where Tilapia, Golsha, Pabda, Koi, and Shing (Bengali name of the fishes) are cultivated. The project has been established on the “No Water Exchange” principle. We visited there in September 2020. Investigating their working procedure, we found that they examine the water two times every day and collect the pH, water temperature, total dissolved oxygen (TDS), ammonia (NH₃), and floc (molasses). Fig. 5 shows a complete picture of the study area.

Suppose the water parameters get out of the suitable range. In that case, they control them by taking necessary steps for example: filtering out the excess bioflocs (as it tends to increase generally), adding baking soda in a safe amount, and removing the fish from the tank before raising the pH. The tank's water is matured, so the parameters do not change frequently, and the water is adaptable for the fish species.

Another center is located at Bosila, Dhaka named “Matsabid Biofloc Aquaculture Farm”. They are using the same technology but do not follow the “No Water Exchange” method. It is a big project of 15 large tanks and two large ponds. The whole system is continuously monitored manually by a family residing there. The aerator is mandatory for the bioflocs to survive, so the aerator is turned on 24 hours a day. Only three water parameters (pH, salinity, and the number of flocs) are monitored in the farm and whenever one gets increased or decreased, they change the water. For this exchange, they have a drainage system and a water pump. So, the water is not matured here. The image of the study area is shown in Fig. 6 and 7.

F. Data Collection

We collected samples from different centers. We bought instruments and sensors for data collection purposes. We used pH, temperature, and TDS sensors to measure the values. Sample collection was more manageable, but the data processing was difficult and time-consuming. We worked for more than three months to process the necessary data from the samples. Tables I and II show the details of the collected data.



Fig. 5. Study Area 1 at “Biotech Aquaculture”.



Fig. 6. Study Area 2 at “Matshyabid Biofloc Aquaculture Farm”



Fig. 7. Fish Tank at “Matshyabid Biofloc Aquaculture Farm”.

TABLE I. COLLECTED DATA FROM STUDY AREA 1

Day	Date	pH	Temp	TDS	NH3	Floc
1	08/06/20	7	30	1.75	2	10 ml
2	09/06/20	7	30	1.6	4	20 ml
3	10/06/20	7	28	1.3	2	8 ml
4	11/06/20	7	29	1.35	2	10 ml
5	12/06/20	7	28	1.32	0.25	13 ml
6	15/06/20	7	28	1.32	1	10 ml
7	16/06/20	7	30	1.59	0.25	8 ml
8	18/06/20	7	29	1.58	2	10 ml
9	19/06/20	7	28	1.33	2	20 ml
10	20/06/20	7	28	1.56	0.5	15 ml
11	21/06/20	7	30	1.60	1	20 ml
12	22/06/20	7	30	1.55	0	12 ml
13	23/06/20	7	29	1.54	0	20 ml
14	25/06/20	7	30	1.75	0	8 ml

TABLE II. COLLECTED DATA FROM STUDY AREA 2

Date	Time	pH	TDS	Floc
18/10/20	7.00 am	8.2	652	45 gm
18/10/20	8.00 pm	8.2	684	55 gm
19/10/20	7.00 am	8.2	684	43 gm
20/10/20		8.3	684	54 gm
21/10/20	7.00 am	8.3	659	22 gm
21/10/20	8.00 pm	8.4	654	25 gm
22/10/20	7.00 am	8.5	682	25 gm
22/10/20	8.00 pm	7.1	678	50 gm
23/10/20	7.00 am	7.9	654	45 gm
25/10/20	9.00 am	8.3	658	30 gm
26/10/20	7.00 am	8.3	684	29 gm
26/10/20	8.00 pm	8.4	681	35 gm
27/10/20	9.00 am	7.9	652	30 gm
28/10/20	7.00 am	8.3	685	33 gm
28/10/20	8.00 pm	8.2	682	45 gm

G. Data Preprocessing

We collected data from water using different instruments and stored it in a CSV file. The CSV file contains six various labels, including temperature, TDS, pH, and flocs. pH is considered as the output data by which the model was tested and the rest five are regarded as the input data by which the model was trained. A glimpse of our dataset is shown in Table III.

H. Machine Learning Algorithm for Decision Making

Artificial Neural Networks (ANN) function as the neurons of the human brain. The network contains nodes that receive the input signal and pass it to the previous nodes as the synapse of a nerve cell does [18, 19]. The covariates and input variables are weighted, and these weighted signals are then passed through activation functions. Let y be output signal and be the activation function, the mathematical expression of signal processing in ANN is:

$$y(x) = \Phi(\sum_{i=1}^n w_i \cdot x_i)$$

The network contains an input layer, an output layer, and one or more hidden layers. The hidden layers are responsible for the performance of the model. We used five hidden layers for faster execution. More layers slow down the training and testing process.

I. Dependencies

We trained the model on Ubuntu 20.04 LTS and used python 3.8. These were dependencies and libraries. We installed and imported the following libraries to run the code and train the model.

- Tensor flow
- Keras
- Pandas

- Matplot
- Numpy
- Boxplot Analysis

J. Parameters and Algorithms

The whole dataset was splatted into two different parts that are training and testing. We took 80% of the data for training, and the rest 20% was used for testing purposes. The model was trained several times using different epoch sizes. We encountered the overfeeding condition. Lastly, utilizing batch size 32 and epoch size 150, the model achieved the best accuracy. The model consists of 5 hidden layers. The layers help to increase the model’s performance. The workflow of the layers is shown in Fig. 8.

To vanish the gradient problem and allow the model to run faster and perform well, we used ReLu (Rectified Linear Unit) [20, 21] with the hidden layers. Softmax was utilized as an activation function in the output layer. The number of classes was 4 for the output [22]. 0 denotes a shallow DO level, 1 represents a low DO level, 2 denotes an average DO level, and 3 denotes a high DO level.

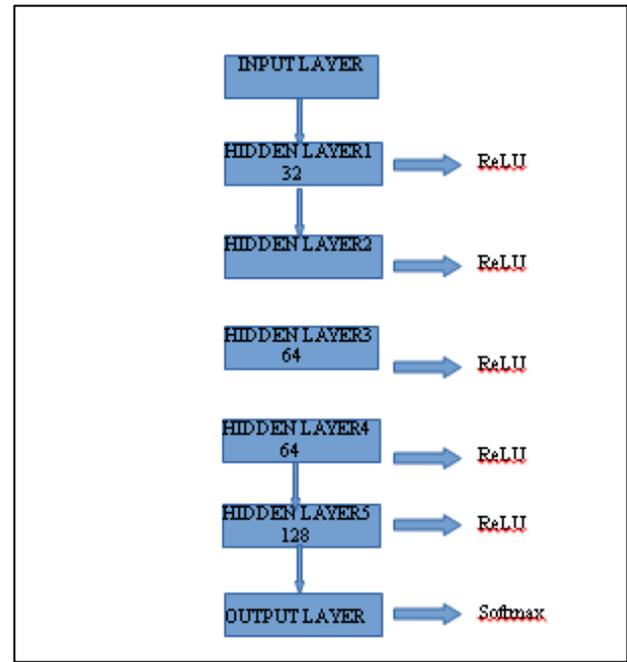


Fig. 8. Workflow of the Layers.

TABLE III. PROCESSED DATA

	A	B	C	D	E
1	Temp	D.O. (mg)	pH	TDS	Flocs (ml)
2	29.5	6.3	6.9	1.7	10
3	29.7	5.7	6.9	3.8	50
4	29.5	5.8	7.3	1.9	40
5	30	5.5	7.4	1.5	10
6	29.2	6.1	6.7	1.4	30
7	29.1	7.3	7	1	10
8	28.7	7	6.9	1.2	30
9	28.7	7.3	6.7	1.5	10
10	29.5	7.2	6.8	1.2	30
11	29	5.3	6.4	1.6	120
12	30.5	6.3	7.5	1.5	10
13	29.1	5.5	6.3	1.4	10
14	30.1	7.3	7.8	2	30
15	29.2	6.5	7.9	1.5	40
16	25.1	7.2	7.7	4.9	30
17	29.6	6.6	7.8	5	60
18	27.4	6.9	7.3	5.2	30
19	27.8	6.8	7.9	4.9	180
20	30.6	6.7	7.6	10.3	190
21	25	5.1	7.6	3.6	60
22	28.1	5.6	7.7	4.6	70
23	28.6	6.3	6.9	4.7	40
24	26.9	6.6	6.8	5	30
25	28.2	6.8	6.5	5.2	60

IV. SIMULATION AND RESULT

A. Training and Testing

After collecting and processing the dataset, we trained the model. The model was trained with the 80% data and later tested with the rest 20% data. We utilized a completely different type of data for training and testing purposes. The model scored 0.773 testing accuracy, which was well enough to maintain good performance. The loss was calculated with the increasing number of epochs. After 55 epochs, the loss was minimized rapidly. We calculated the loss between 0 to 1.2 range. The minimum testing loss was 0.5, and the training loss was 0.7. The training and testing accuracy and loss have been shown in Fig. 9 and 10.

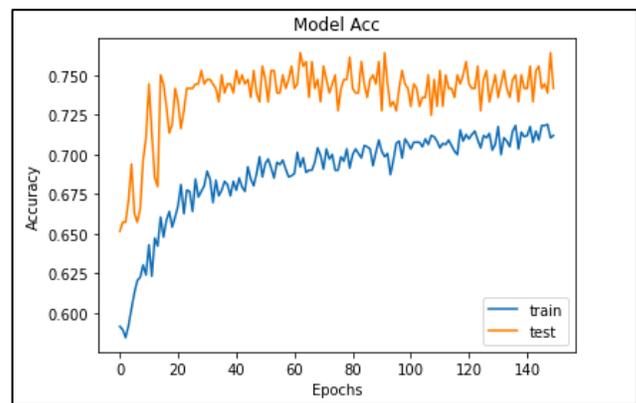


Fig. 9. Training and Testing Accuracy.

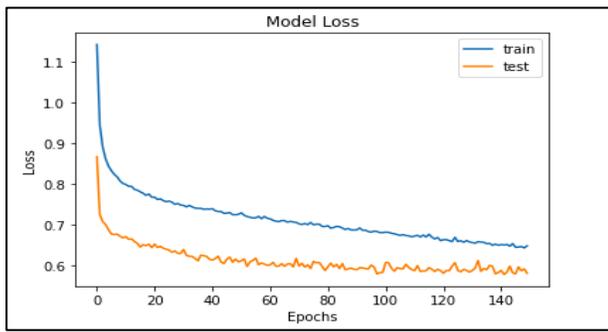


Fig. 10. Training and Testing Loss.

B. System Output

Evaluating the parameter’s value, the system provides an output. Depending on the DO level (shallow, low, average, or high) in the water, the system generates a message and provides a decision analyzing the current output. To make the system more user-friendly, we have designed a smartphone application to notify the user about the result and determination that the machine generates.

A user needs the android application, an ubuntu droplet, and a droplet’s IP address for the push notification. We used GCM (Google Cloud Messaging) [23] for android, where we enabled the API for our project first and then linked the android App through it. We deployed an Ubuntu droplet and set up a python GCM simple server on Ubuntu. Lastly, a push notification was displayed on the Android app generated by the system. A snapshot of the smartphone application is shown in Fig. 11.

C. Performance Comparison

Many researchers have worked on intelligent biofloc technology. pH, DO, BOD, COD, TSS, TDS, EC, PO43-, NO3-N, and NH3-N are the standard parameters utilized by most researchers for measuring water quality and its changes. The Artificial Neural Network (ANN) [24], Group Data Handling Method (GMDH) [25], Support Vector Machine (SVM) [26], Least-Squares Support Vector Regression (LSSVR) [27] and Long Short-Term Memory (LSTM) [28] are

commonly used methods. Besides, some projects used microcontrollers and sensors only. The projects are cost-effective, but such projects do not predict the water quality, fish health condition and cannot process automatic solutions. A caretaker is needed to perform the manual functions. In this article, we have proposed and implemented a system that will help fish farmers grow fish production minimizing the feeding cost. It needs not to use big ponds or a wider area. Our project can help anyone to produce plenty of fishes in a small container or house at a small cost. We have compared our approach with other existing techniques. For comparing, we provided significant importance to the proposed method, accuracy, cost reduction rate, parameters, real-time monitoring capability, prediction capability, decision-making capability, and user satisfaction level. The comparison is shown in table 4. The information mentioned in the table is collected from different published articles. Here we did not compare among the methods. We have reached the performances of different approaches only.



Fig. 11. Notification to the Smartphone Application.

TABLE IV. PERFORMANCE COMPARISON

Method Used	Testing Accuracy (Percentage)	Real-time Monitoring	Parameters	Automatic Solution	Decision-Making Ability	User satisfaction
AAN	72 %	Monitors 24 x 7	Temperature, DO, TDS, pH, BOD, COD, TSS	Does not perform automatic solution	Can predict and process smart decision	Medium
GMDH	74%	Monitors 24 x 7	Temperature, DO, TDS, pH	Does not perform automatic solution	Better prediction and decision-making system	High
SVM	70%	Monitors 24 x 7	Temperature, DO, TDS, pH	Does not perform automatic solution	Lower quality prediction	Low
LSSVR	76%	Monitors 24 x 7	Temperature, DO, TDS, pH, EC, PO43-, NO3-N	Does not perform automatic solution	Can predict and process smart decision	Medium
LSTM	82 %	Monitors 24 x 7	pH, temperature, DO	Does not perform automatic solution	Better prediction and decision-making system	High
Our Approach	77 %	Monitors 24 x 7	Temperature, DO, TDS, pH, Floc	Performs automatic solution	Can predict and process smart decision	Android application is available for monitoring from anywhere. User satisfaction is measured as “Very High”

V. CONCLUSION

Aquaculture farmers have been surviving from economic constraints, high-paid and even unavailability of human resources, timely monitoring of water quality, and sudden increase in toxicity for decade after decade. The IoT-based water quality monitoring system monitors the water quality in real-time and reduces the cost of production, increases efficiency, reduces human dependency, and thus ensures sustainable development economically and socially. The proposed system monitors the water quality in real-time and sends a notification to the user instantly, which reduces the risk. A machine learning technique has been applied to trace the water quality. To validate the proposed model, experiments on the implemented functionalities have been performed. The experiments show 0.773 as testing accuracy which was well enough to maintain good performance. Currently, the implementation of identified functionalities has been carried out. In the future, we wish to improve the model to achieve higher accuracy and evaluate the performance in terms of the fish population.

ACKNOWLEDGMENT

The IoT Lab has supported this research work, a state-of-art specialized research facility of its kind situated at ULAB, implemented, and supported by the Bangladesh Hi-Tech Park Authority of ICT Division.

REFERENCES

- [1] S. Saha, R. Hasan and S. Kabir, "IoT Based Smart Farm Monitoring System", *International Journal of Recent Technology and Engineering*, vol. 8, no. 4, pp. 5490-5494, 2019. Available: 10.35940/ijrte.d8806.118419 [Accessed 1 November 2020].
- [2] R. Crab, T. Defoirdt, P. Bossier and W. Verstraete, "Biofloc technology in aquaculture: Beneficial effects and future challenges", *Aquaculture*, vol. 356-357, pp. 351-356, 2012. Available: 10.1016/j.aquaculture.2012.04.046 [Accessed 1 December 2020].
- [3] B. Ghose, "Fisheries and Aquaculture in Bangladesh: Challenges and Opportunities.", *Aquaculture Research*, 2014. Bashar, Abul. (2018). Biofloc Aquaculture: prospects and challenges in Bangladesh. 10.13140/RG.2.2.13233.94560.
- [4] D. SK, "Biofloc Technology (BFT): An Effective Tool for Remediation of Environmental Issues and Cost Effective Novel Technology in Aquaculture", *International Journal of Oceanography & Aquaculture*, vol. 2, no. 2, 2018. Available: 10.23880/ijoac-16000135 [Accessed 3 November 2020].
- [5] X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong and C. Zhou, "Deep learning for smart fish farming: applications, opportunities and challenges", *Reviews in Aquaculture*, vol. 13, no. 1, pp. 66-90, 2020. Available: 10.1111/raq.12464 [Accessed 27 December 2020].
- [6] S. Liu et al., "Prediction of dissolved oxygen content in river crab culture based on least squares support vector regression optimized by improved particle swarm optimization", *Computers and Electronics in Agriculture*, vol. 95, pp. 82-91, 2013. Available: 10.1016/j.compag.2013.03.009 [Accessed 20 November 2020].
- [7] U. Ahmed, R. Mumtaz, H. Anwar, A. Shah, R. Irfan, and J. García-Nieto, "Efficient Water Quality Prediction Using Supervised Machine Learning", *Water*, vol. 11, no. 11, p. 2210, 2019. Available: 10.3390/w11112210 [Accessed 20 November 2020].
- [8] A. A. Nayan, M. G. Kibria, M. O. Rahman and J. Saha, "River Water Quality Analysis and Prediction Using GBM," 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), Dhaka, Bangladesh, 2020, pp. 219-224, DOI: 10.1109/ICAICT51780.2020.9333492.
- [9] A. A. Nayan, J. Saha, A. N. Mozumder, K. R. Mahmud, A. K. A. Azad, M. G. Kibria "Early Detection of Fish Diseases by Analyzing Water Quality Using Machine Learning Algorithm", *Walailak Journal of Science and Technology*, vol. 18, Article 11740 (12 pages), 2021.
- [10] D. Marisda, "Application of Tilapia Cultivation Biofloc Technology For Utilization of Non-Productive House Yards", *SEWAGATI*, vol. 3, no. 3, 2019. Available: 10.12962/j26139960.v3i3.6111.
- [11] J. Choi, J. Park, H. Kim, J. Hwang, D. Lee, and J. Lee, "Assessment Of Water Quality Parameters During A Course Of Applying Biofloc Technology(Bft)", *Journal Of Fishries And Marine Sciences Education*, Vol. 32, No. 6, pp. 1632-1638, 2020. Available: 10.13000/Jfmse.2020.12.32.6.1631.
- [12] J. Jung, J. Hur, K. Kim, and H. Han, "Evaluation of floc-harvesting technologies in biofloc technology (BFT) system for aquaculture", *Bioresource Technology*, vol. 314, p. 123719, 2020. Available: 10.1016/j.biortech.2020.123719.
- [13] "Arduino - ArduinoBoardUno", *Arduino.cc*, 2021. [Online]. Available: <https://www.arduino.cc/en/Main/arduinoBoardUno>>. [Accessed: 08-Mar- 2021].
- [14] "Sensor for pH determination", *Sensor Review*, vol. 20, no. 2, 2000. Available: 10.1108/sr.2000.08720bad.016.
- [15] "New low temperature ultrasonic ranging sensor", *Sensor Review*, vol. 20, no. 1, 2000. Available: 10.1108/sr.2000.08720aad.012.
- [16] L. Zhang, "Effects of electrolyte total dissolved solids (TDS) on performance and anodic microbes of microbial fuel cells", *AFRICAN JOURNAL OF BIOTECHNOLOGY*, vol. 10, no. 74, 2011. Available: 10.5897/ajb11.1993.
- [17] X. Kuang and H. Huo, "A Design of WIFI Wireless Transmission Module Based on MCU", *Applied Mechanics and Materials*, vol. 442, pp. 367-371, 2013.
- [18] "Machine Learning and Artificial Neural Network", *International Journal of Recent Trends in Engineering and Research*, vol. 4, no. 3, pp. 660-668, 2018. Available: 10.23883/ijrter.2018.4179.tdms.
- [19] R. CVS and N. Pardhasaradhi, "Analysis of Artificial Neural-Network", *International Journal of Trend in Scientific Research and Development*, vol. -2, no. -6, pp. 418-428, 2018. Available: 10.31142/ijtsrd18482.
- [20] "Pattern Recognition Of Numeric Numbers Using Artificial Neural Network", *International Journal of Recent Trends in Engineering and Research*, pp. 242-244, 2018. Available: 10.23883/ijrter.conf.20171201.048.kzv9.
- [21] C. Banerjee, T. Mukherjee and E. Pasilio, "Feature representations using the reflected rectified linear unit (RReLU) activation", *Big Data Mining and Analytics*, vol. 3, no. 2, pp. 102-120, 2020. Available: 10.26599/bdma.2019.9020024.
- [22] K. Kaiho, "Benthic foraminiferal dissolved-oxygen index and dissolved-oxygen levels in the modern ocean", *Geology*, vol. 22, no. 8, p. 719, 1994.
- [23] V. Dubey, "Android Device to Device Messaging using Google Cloud Messaging (GCM)", *SMART MOVES JOURNAL IJOSCIENCE*, vol. 1, no. 1, 2015. Available: 10.24113/ijoscience.v1i1.6.
- [24] A. Kadam, V. Wagh, A. Muley, B. Umrikar and R. Sankhua, "Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India", *Modeling Earth Systems and Environment*, vol. 5, no. 3, pp. 951-962, 2019. Available: 10.1007/s40808-019-00581-3.
- [25] O. Varis, "Water Quality Models: Tools For The Analysis Of Data, Knowledge, and Decisions", *Water Science and Technology*, vol. 30, no. 2, pp. 13-19, 1994. Available: 10.2166/wst.1994.0024.
- [26] A. Haghiabi, A. Nasrolahi and A. Parsaie, "Water quality prediction using machine learning methods", *Water Quality Research Journal*, vol. 53, no. 1, pp. 3-13, 2018. Available: 10.2166/wqrj.2018.025.
- [27] G. Tan, J. Yan, C. Gao and S. Yang, "Prediction of water quality time series data based on least squares support vector machine", *Procedia Engineering*, vol. 31, pp. 1194-1199, 2012. Available: 10.1016/j.proeng.2012.01.1162.
- [28] Z. Hu et al., "A Water Quality Prediction Method Based on the Deep LSTM Network Considering Correlation in Smart Mariculture", *Sensors*, vol. 19, no. 6, p. 1420, 2019. Available: 10.3390/s19061420.

A Novel High-order Linguistic Time Series Forecasting Model with the Growth of Declared Word-set

Nguyen Duy Hieu¹

Faculty of Natural Sciences and Technology
Tay Bac University
Son La, Vietnam

Pham Dinh Phong^{2*}

Faculty of Information Technology
University of Transport and Communications
Hanoi, Vietnam

Abstract—The existing researches have shown that the fuzzy forecasting methods based on the high-order fuzzy time series are better than the fuzzy forecasting methods based on the first-order fuzzy time series. The linguistic forecasting method based on the first-order linguistic time series which can handle directly the word-set of the linguistic variable have been examined by Hieu et al. This paper examines a novel model of high-order linguistic time series with the growth of declared word-set. A procedure for forecasting the enrollments of University of Alabama and Lahi crop production of India is developed based on the proposed model. In the proposed forecasting method, the high-order linguistic logical relationship groups will be established and utilized to calculate the forecasted values based on the quantitative semantics of the used words generated by the hedge algebras structure. The experimental results show that the forecasting accuracy of the proposed high-order forecasting method is better than their counterparts and the growth of the word-set of the linguistic variable is significant in increasing the accuracy of the forecasted results.

Keywords—Linguistic time series; high-order linguistic time series; linguistic logical relationship; hedge algebras; time series forecasting

I. INTRODUCTION

The subject of the time series forecasting model has attracted substantial attention in recent years. For almost the realistic cases, the forecasting problem is uncertainty. In other words, it seems impossible to predict the future values of things and phenomena with one hundred percent of accuracy. However, we can examine different forecasting models and find the best fit model for a certain dataset.

The living of human is a sequence of the decision-making process. In reality, they always use their linguistic experience to make decisions or in other words, they save their experiences of the real world by linguistic words and use them to make decisions in the future. In computer science, researchers have found many ways to simulate the human behaviors into the machines. The most successful one is fuzzy set theory approach. Fuzzy set theory proposed by Zadeh [1] uses fuzzy sets to represent the computational semantics of linguistic words. The utilization of fuzzy set theory to time series forecasting problem first proposed by Song and Chissom in 1993 [2]–[4]. For example, we can use linguistic words such

as “small”, “quite small”, “very small”, “large”, “rather large”, “extremely large”, etc. to describe the number of students who enroll to a university by years. By the fuzzy time series, the mentioned linguistic words can be represented by fuzzy sets. There are many publications of the fuzzy forecasting methods based on the fuzzy time series which can be found in the literature.

Chen in [5] examined the high-order fuzzy time series forecasting model where the n^{th} -order fuzzy logical relationship can be derived from the fuzzified historical data. In recent two decades, among so many improvements of the fuzzy time series forecasting methods, a number of high-order fuzzy time series forecasting methods has been proposed such as [6]–[23]. Among those studies, Singh [17] proposed a computational method of forecasting for the fuzzy time series model using high-order fuzzy logical relationships of differences of consecutive values. Most of other improvements of high-order fuzzy time are the proposals of the hybrid forecasting models with modern computational techniques using artificial neural networks [11], [12], [16], [19], machine learning [13], [14], particle swarm optimization [23], genetic algorithm [6], granular computing [8], etc.

Despite many important achievements, in the computation model of fuzzy time series, there is no formal connection between the fuzzy sets and the inherent semantics of the linguistic words represented by the fuzzy sets. In other words, the linguistic words are used in the fuzzy time series merely as the linguistic labels of fuzzy sets. Therefore, the relationships between the qualitative semantics of linguistic words and their fuzzy sets based semantics have not been thoroughly examined in the fuzzy time series forecasting models.

In the other side of natural language processing method, hedge algebras (HA) was introduced by Ho and Wechler in 1990 [24]. In hedge algebras, the linguistic variables can be modeled by an axiomatic way using inherent quantitative semantics of linguistic values. Many applications of hedge algebras can be found in the field of fuzzy control [25]–[31], data classification and regression using fuzzy rule-based systems [32], [33], image processing [34] and so on. In 2020, Hieu et al. [35], [36] proposed linguistic time series (LTS) forecasting model handling the linguistic words immediately based on the theory of hedge algebras. It is the first time that

*Corresponding Author

This research is funded by Ministry of Education and Training of Vietnam under a project led by Nguyen Duy Hieu from 2022 to 2023.

the constructed linguistic time series can simulate the way human users describe numerical time series in terms of their linguistic words.

In this paper, inspired by the mentioned researches of fuzzy time series, we propose a novel high-order linguistic time series forecasting model. In the proposed model, the current linguistic value of the linguistic variable \mathcal{X}_L depend not only on the adjacent previous word but also the series of preceding words. In addition, the forecasting linguistic variable \mathcal{X}_L is considered as the linguistic counterpart of the numerical forecasting one \mathcal{X}_N . So, the word-set of \mathcal{X}_L is infinite and just limited words of it are selected for a specific application, so-called declared word-set. In our proposed high-order LTS forecasting models, we assume that there is not any restriction on the declared word set. The accuracy of the proposed high-order LTS forecasting models can considerably increase when the declared word-set of \mathcal{X}_L grows. The proposed high-order LTS forecasting models are applied to the problems of the enrollments of University of Alabama [4] and the Lahi crop production of India [17] to illustrate the effectiveness of the proposed forecasting models and the scalability of the word-set of \mathcal{X}_L . Besides, this study would like to present a comprehensive formalized methodology for handling linguistic words directly and simulate the way human users making a forecasting process in terms of their words. Therefore, we do not apply any optimization techniques which may improve the accuracy of forecasted results.

The rest of the paper will be organized as follows. Section II briefly introduces some fundamental knowledge of hedge algebras and the linguistic time series forecasting model. The proposed novel high-order linguistic time series forecasting model is also presented in this section. The experimental studies will be introduced and discussed in Section III. Finally, the paper will be covered with the conclusions.

II. LINGUISTIC TIME SERIES AND ITS FORECASTING MODEL

A. Some basic Knowledges of Hedge Algebras

In the human society, people utilize their historical experiences to communicate one to another by linguistic words in natural language. Linguistic words themselves reveal their inherent semantics. Because of their qualitative semantics, every word has its semantic order relation with each other in the word-domain of a linguistic variable. That motivated Ho and Wechler [24], [37] to propose hedge algebras which can deal with linguistic words directly.

Let $\mathcal{AX} = (X, G, C, H \leq)$ be a hedge algebras (HA), where X is the word-domain of a given linguistic variable \mathcal{X}_L ; G is the set of two opposite generators, the negative one denoted by c^- and the positive one denoted by c^+ ; C is the set of three constants $0, W, 1$ which are the smallest word, the neutral and the largest ones (e.g., such specific words are “absolutely small”, “medium” and “absolutely large”); H is the set of hedges such as “little”, “rather”, “very”, “extremely”, and so on; \leq is the semantic order relation in the word-domain.

Each linguistic value in HA structure has its own semantic. In order to model the quantitative semantics of all words in the

word-domain, it is necessary to construct two notions of HA which are semantically quantifying mapping (SQM) and fuzziness measure, denoted by $v : X \rightarrow [0,1]$ and $m : X \rightarrow [0,1]$, respectively. SQM and fuzziness measure of HA structure can be defined as follows.

Definition 1. [38] A function $m : X \rightarrow [0,1]$ is said to be *fuzziness measure* of terms in X provided that:

- $m(c^-) + m(c^+) = 1$;
- $\sum_{h \in H} m(hu) = m(u)$, for $\forall u \in X$;
- $m(0) = m(W) = m(1) = 0$, fuzziness measure of all constants is 0;
- For $\forall x, y \in X, \forall h \in H, \frac{m(hx)}{m(x)} = \frac{m(hy)}{m(y)}$, this proportion does not depend on any specific elements and, hence, it is called fuzziness measure of the hedge h and denoted also by $\mu(h)$ for simplicity.

The fuzziness measure m of a linguistic variable \mathcal{X}_L has the following properties:

$$m1) m(hx) = \mu(h)m(x), \text{ for } \forall x \in X;$$

$$m2) m(c^-) + m(c^+) = 1;$$

$$m3) \sum_{-q \leq i \leq p, i \neq 0} m(h_i x) = m(x), \forall x \in X;$$

$$m4) \text{ Set } \sum_{-q \leq i \leq -1} \mu(h_i) = \alpha, \sum_{1 \leq i \leq p} \mu(h_i) = \beta, \text{ where } \alpha, \beta > 0 \text{ and } \alpha + \beta = 1.$$

In the computational model of HA, each linguistic word or hedge has its sign defined as follows.

Definition 2. [38] A function $sign : X \rightarrow \{-1,1\}$ is a mapping which is defined recursively as follows. For $h, h' \in H$ and $c \in \{c^-, c^+\}$:

- 1) $sign(c^-) = -1, sign(c^+) = +1$;
- 2) $sign(hc) = -sign(c)$ ($= +sign(c)$) if h is negative (positive) with respect to c ;
- 3) $sign(h'hx) = -sign(hx)$ if $h'hx \neq hx$ and h' is negative with respect to h ;
- 4) $sign(h'hx) = +sign(hx)$ if $h'hx \neq hx$ and h' is positive with respect to h .

Definition 3. [38] For given values of the fuzziness parameters of a variable \mathcal{X}_L , its corresponding SQM $v : X \rightarrow [0, 1]$ is defined as follows:

$$v(W) = \theta = m(c^-); \tag{1}$$

$$v(c^-) = \theta - \alpha m(c^-) \text{ and } v(c^+) = \theta + \alpha m(c^+); \tag{2}$$

$$v(h_j x) = v(x) + sign(h_j x) \left\{ \sum_{i=sign(j)}^j m(h_i x) - \omega(h_j x) m(h_j x) \right\} \tag{3}$$

$$\text{where } \omega(h_j x) = \frac{1}{2} [1 + sign(h_j x) sign(h_p h_j x) (\beta - \alpha)] \in \{\alpha, \beta\}.$$

SQM and fuzziness measure are two important notions of HA. More specifically, for given α and θ parameters, all the components of the HA's computational model can be obtained.

B. Linguistic Time Series (LTS)

Humans use their own language capacity to collect the data. Their brain saves all information in the real world by linguistic experience of specific owner. Human also use their own experience to make decision. In the time series forecasting problem, simulating the behavior of human's decision making process, the model also uses preceding data to predict the future value of time series data. When all the points of historical time series data are linguistic values, we have the LTS.

Definition 4. [35] Let \mathbb{X} be a set of linguistic words in the natural language of a variable \mathcal{X} defined on the universe of discourse $U_{\mathcal{X}}$ to describe its numeric quantities. Then, any series $L(t), t = 0, 1, 2, \dots$, where $L(t)$ is a finite collection of words of \mathbb{X} , is called a linguistic time series.

Definition 5. [35] Let T be a LTS. In T , if every time $t = k - 1$, the linguistic value of LTS is x_i and x_j is the linguistic value of T at the time $t = k$, then we have linguistic logical relationship of the form $x_i \rightarrow x_j$.

Definition 6. [35] Suppose that we have more than one linguistic logical relationship with the same left-hand side denoted by $x_i \rightarrow x_{j_1}, x_i \rightarrow x_{j_2}, \dots, x_i \rightarrow x_{j_n}$, then we can group them into the linguistic logical relationship group denoted by $x_i \rightarrow x_{j_1}, x_{j_2}, \dots, x_{j_n}$.

In Def. 5, the logical relationship $x_i \rightarrow x_j$ indicates that the linguistic value x_j to be caused by the only linguistic value x_i , called the first-order linguistic logical relationship. If there is more than one element on the left-hand side of the relationship, then we have the high-order linguistic logical relationship.

Definition 7. In the linguistic time series T , if x_k , the linguistic value of \mathcal{X} at the time k , to be caused by $x_{k-1}, x_{k-2}, \dots, x_{k-\lambda}$ are the linguistic values of \mathcal{X} at the time $k - 1, k - 2, \dots, k - \lambda$, respectively, then the linguistic logical relationship is presented by $x_{k-\lambda}, \dots, x_{k-2}, x_{k-1} \rightarrow x_k$, so-called the λ -th-order linguistic logical relationship.

By Def. 7, we have first-order linguistic logical relationship when $\lambda = 1$ and we have high-order linguistic logical relationship when $\lambda > 1$.

In a similar way to Def. 6, the high-order linguistic logical relationship group is established based on the high-order linguistic logical relationships. Suppose that we have the high-order linguistic logical relationships:

$$x_{k-\lambda}, \dots, x_{k-2}, x_{k-1} \rightarrow x_{k1},$$

$$x_{k-\lambda}, \dots, x_{k-2}, x_{k-1} \rightarrow x_{k2},$$

$\dots,$

$$x_{k-\lambda}, \dots, x_{k-2}, x_{k-1} \rightarrow x_{km},$$

then we can group them into the high-order linguistic logical relationship group written in the form:

$$x_{k-\lambda}, \dots, x_{k-2}, x_{k-1} \rightarrow x_{k1}, x_{k2}, \dots, x_{km},$$

where λ is the order of the linguistic logical relationship group.

C. High-order LTS Forecasting Model

Using the high-order linguistic logical relationship groups, the high-order LTS forecasting model was established and briefly reviewed as in Fig. 1.

In the first step, the universe of discourse and the HA structure with syntactic, qualitative and quantitative semantics are determined. The second step computes the numerical semantics of the declared words of \mathcal{X}_L . In the third step, according to the historical data, the high-order linguistic logical relationship groups, used in the next step, were observed. Using the linguistic logical relationship groups, ones can calculate the forecasted values of the time series data. Some performance measures can be utilized to assert the effective of the forecasting model.

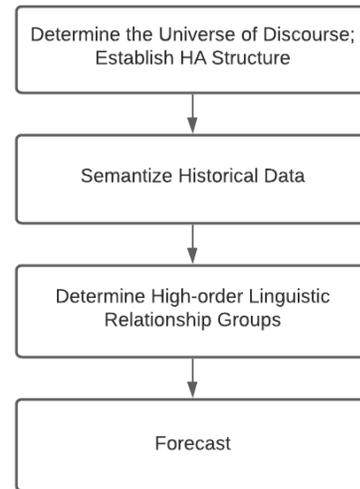


Fig. 1. The High-order Linguistic Time Series Forecasting Model.

III. SOME EXPERIMENTAL RESULTS AND DISCUSSIONS

This section presents the experimental results of the proposed high-order LTS forecasting models and shows their performance in comparison with their counterpart models. The comparative study is designed to compare the proposed high-order LTS forecasting model performance with those examined by Chen [5] and Singh [17], using the historical data of the enrollments of University of Alabama observed from 1971 to 1992, and Lahi crop production of India observed from 1981 to 2003.

The experiments are executed to show the following statements:

- The proposed high-order LTS forecasting models outperform their counterpart of fuzzy approaches with restricted linguistic words.
- In the proposed high-order LTS forecasting models, the forecasting linguistic variable \mathcal{X}_L is considered as the linguistic counterpart of the numerical forecasting one \mathcal{X}_N . So, we assume that there is not any restriction on its declared word set's cardinality. The accuracy of the proposed high-order LTS forecasting models can considerably increase when the declared word-set of \mathcal{X}_L grows.

A. Forecast the Enrollments of University of Alabama

1) Compare with the two counterparts examined by Chen and Singh: To show the efficiency of the proposed high-order LTS forecasting models, we apply them to the historical data of the enrollments of University of Alabama observed from 1971 to 1992, as shown the column “Actual Enrollment” in Table I. To be more convenient, the experiment follows the procedure as follows:

Step 1: Determine the syntactic, qualitative and quantitative semantics of the forecasting linguistic variable \mathcal{X}_L .

- The syntactic semantics are determined by declaring two primary words $c^- = small$, $c^+ = large$ (l) and the linguistic hedges. In this study, we just use two linguistic hedges $H^- = \{Little (L)\}$ and $H^+ = \{Very (V)\}$. For the model illustration and comparative study, the cardinality of the declared word-set of \mathcal{X}_L is set to 9 (2-specificity level). So, we have: $U_{\mathcal{X}_L} = X_{(2)} = \{\mathbf{0}, V_small, small, L_small, medium, L_large, large, V_large, \mathbf{1}\}$, where $\mathbf{0}$ stands for *Absolutely small* and $\mathbf{1}$ stands for *Absolutely large*.
- The qualitative semantics of \mathcal{X}_L is specified by the signs of hedges and the relative signs between them: $sign(L) = sign(L, L) = sign(L, V) = -1$ and $sign(V) = sign(V, V) = sign(V, L) = +1$.
- The quantitative semantics is defined by specifying the fuzziness parameter values of \mathcal{X}_L . In this study, they are specified by human’s experts and we have $m(c^-) = 0.437$, $\mu(L) = 0.511$.

Step 2: Compute the numerical semantics of the declared words of \mathcal{X}_L .

Transform the numerical semantics (SQM-values) of the declared words in the normalized domain $[0, 1]$ into the real numerical semantics in the universe of discourse $U_x = [13000, 20000]$ of \mathcal{X}_L by the following formula:

$$v_R(l_i) = U_{x_{min}} + (U_{x_{max}} - U_{x_{min}}) \times v(l_i), \quad (4)$$

where $l_i \in X_{(2)}$, $v(l_i)$ is the numerical semantic of the word l_i , $U_{x_{min}}$ is the lower bound and $U_{x_{max}}$ is the upper bound of U_x . Specifically, with the set $U_{\mathcal{X}_L}$, $m(c^-) = 0.437$ and $\mu(L) = 0.511$, we have the real numerical semantics of the declared words: $U_{\mathcal{X}_R} = \{13000, 13731, 14496, 15295, 16059, 17044, 18073, 19058, 20000\}$.

Step 3: Generate LTS from the given numerical time series.

Based on the real numerical semantics, the historical data of the enrollments are transformed into the linguistic words selected from the declared word-set of \mathcal{X}_L in such a way that each historical data value is assigned a linguistic word which has its real numerical semantics closest to that historical data value at that timestamp. For example, the enrollment value of 15460 is of the year of 1975 closest to the real numerical semantics of 15295 of the word *Little small* in $U_{\mathcal{X}_L}$. Therefore, that enrollment value is assigned the word *Little small*. By doing the same way for the enrollment data of other years, we obtain the LTS as shown in the column “Constructed LTS” in

Table I. Once the LTS is constructed, the high-order linguistic logical relationships (LLRs) and the high-order linguistic logical relationship groups (LLRGs) are created. Just for the illustration of the case of 2-specificity level of \mathcal{X}_L (9 declared words), the second-order and third-order LLRGs are shown in Table II and Table III, respectively.

TABLE I. THE LINGUISTIC TIME SERIES OF THE ENROLLMENTS OF THE UNIVERSITY OF ALABAMA FROM 1971 TO 1992 USING 9 WORDS

Year	Actual Enrollment	Constructed LTS
1971	13055	<i>Absolutely small</i>
1972	13563	<i>Very small</i>
1973	13867	<i>Very small</i>
1974	14696	<i>small</i>
1975	15460	<i>Little small</i>
1976	15311	<i>Little small</i>
1977	15603	<i>Little small</i>
1978	15861	<i>medium</i>
1979	16807	<i>Little large</i>
1980	16919	<i>Little large</i>
1981	16388	<i>medium</i>
1982	15433	<i>Little small</i>
1983	15497	<i>Little small</i>
1984	15145	<i>Little small</i>
1985	15163	<i>Little small</i>
1986	15984	<i>medium</i>
1987	16859	<i>Little large</i>
1988	18150	<i>large</i>
1989	18970	<i>Very large</i>
1990	19328	<i>Very large</i>
1991	19337	<i>Very large</i>
1992	18876	<i>Very large</i>

TABLE II. THE SECOND-ORDER LLRGs USING 9 WORDS

Group	2-order LLRG
1	$V_small, V_small \rightarrow V_small$
2	$V_small, V_small \rightarrow small$
3	$small, V_small \rightarrow L_small$
4	$L_small, small \rightarrow L_small$
5	$L_small, L_small \rightarrow L_small, medium$
6	$medium, L_small \rightarrow L_large$
7	$L_large, medium \rightarrow L_large, large$
8	$L_large, L_large \rightarrow medium$
9	$medium, L_large \rightarrow L_small$
10	$L_small, medium \rightarrow L_small$
11	$large, L_large \rightarrow V_large$
12	$V_large, large \rightarrow V_large$
13	$V_large, V_large \rightarrow V_large$

TABLE III. THE THIRD-ORDER LLRGs USING 9 WORDS

Group	3-order LLRG
1	$V_{small}, V_{small}, A_{small} \rightarrow small$
2	$small, V_{small}, V_{small} \rightarrow L_{small}$
3	$L_{small}, small, V_{small} \rightarrow L_{small}$
4	$L_{small}, L_{small}, small \rightarrow L_{small}$
5	$L_{small}, L_{small}, L_{small} \rightarrow medium, L_{small}$
6	$medium, L_{small}, L_{small} \rightarrow L_{large}$
7	$L_{large}, medium, L_{small} \rightarrow L_{large}, large$
8	$L_{large}, L_{large}, medium \rightarrow medium$
9	$medium, L_{large}, L_{large} \rightarrow L_{small}$
10	$L_{small}, medium, L_{large} \rightarrow L_{small}$
11	$L_{small}, L_{small}, medium \rightarrow L_{small}$
12	$large, L_{large}, medium \rightarrow V_{large}$
13	$V_{large}, large, L_{large} \rightarrow V_{large}$
14	$V_{large}, V_{large}, large \rightarrow V_{large}$
15	$V_{large}, V_{large}, V_{large} \rightarrow V_{large}$

Step 4: Calculate the crisp forecasted values. Two applied computation rules are as follows:

- 1) If an high-order LLRG is of the form: $l_{i1}, l_{i2}, \dots, l_{i\lambda} \rightarrow l_{j1}, l_{j2}, \dots, l_{jp}$ ($\lambda, p \geq 1$), where $l_{i1}, l_{i2}, \dots, l_{i\lambda}$ is the current state of a year, say k and $v_R(l_{j1}), v_R(l_{j2}), \dots, v_R(l_{jp})$ are the real numerical semantics of the words $l_{j1}, l_{j2}, \dots, l_{jp}$, respectively, then the numeric forecasted value of year $k + 1$ is $(v_R(l_{j1}) + v_R(l_{j2}) + \dots + v_R(l_{jp}))/p$.
- 2) If the current state of year k is $l_{i1}, l_{i2}, \dots, l_{i\lambda}$ and there does not exist any word in the right-hand side of any high-

order LLRG, then, the forecasted value of year $k + 1$ is $(v_R(l_{i1}) + v_R(l_{i2}) + \dots + v_R(l_{ip}))/\lambda$.

In this study, the mean square error (MSE) and average forecasting error (AFE) are used to evaluate the forecasting models, where N is the number of forecasted value, F_i and A_i are the forecasted value and the actual value, respectively; ϵ_i is forecasting error of year i :

$$MSE = \left(\frac{1}{N}\right) \sum_{i=1}^N (F_i - A_i)^2 \tag{5}$$

$$\epsilon_i \text{ (in percent)} = \frac{|F_i - A_i|}{A_i} \times 100 \tag{6}$$

$$AFE \text{ (in percent)} = \frac{\epsilon_i}{N} \tag{7}$$

The forecasted enrollments by applying the proposed high-order (from 2-order to 9-order) LTS forecasting models using 9 declared words (9 words selected from the word-set of X_L for describing the universe of discourse U_x) are shown in Table IV. The comparison of MSE values with three other high-order fuzzy time series forecasting models are shown in Table V. The MSE values of the proposed high-order LTS forecasting models are much better than all their counterparts (Chen [5], Hwang et al. [39] and Singh [17]) from the 4-order to 9-order. Evaluated based on AFE values, it is easy to see in Table VI that all proposed high-order LTS forecasting models have AFE values smaller than their counterparts (Song and Chissom [4], Hwang et al. [39] and Singh [17]).

TABLE IV. FORECASTED ENROLLMENTS BY PROPOSED HIGH-ORDER LTS FORECASTING MODELS USING 9 AND 17 DECLARED WORDS.

Year	Actu. Enrol.	Proposed (9 declared words)									Proposed (17 declared words)								
		2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order		
1971	13055																		
1972	13563																		
1973	13867	13731									13750								
1974	14696	14496	14496								14644	14644							
1975	15460	15295	15295	15295							15537	15537	15537						
1976	15311	15295	15295	15295	15295						15169	15169	15169	15169					
1977	15603	15677	15295	15295	15295	15295					15353	15537	15537	15537	15537				
1978	15861	15677	15677	16059	16059	16059	16059				15795	15795	15795	15795	15795	15795			
1979	16807	17044	17044	17044	17044	17044	17044	17044			16689	16689	16689	16689	16689	16689			
1980	16919	17558	17558	17558	17558	17044	17044	17044	17044		16689	16689	16689	16689	16689	16689	16689		
1981	16388	16059	16059	16059	16059	16059	16059	16059	16059		16164	16164	16164	16164	16164	16164	16164		
1982	15433	15295	15295	15295	15295	15295	15295	15295	15295		15537	15537	15537	15537	15537	15537	15537		
1983	15497	15295	15295	15295	15295	15295	15295	15295	15295		15537	15537	15537	15537	15537	15537	15537		
1984	15145	15677	15295	15295	15295	15295	15295	15295	15295		15169	15169	15169	15169	15169	15169	15169		
1985	15163	15677	15677	15295	15295	15295	15295	15295	15295		15353	15169	15169	15169	15169	15169	15169		
1986	15984	15677	15677	16059	16059	16059	16059	16059	16059		16164	16164	16164	16164	16164	16164	16164		
1987	16859	17044	17044	17044	17044	17044	17044	17044	17044		16689	16689	16689	16689	16689	16689	16689		
1988	18150	17558	17558	17558	17558	18073	18073	18073	18073		18053	18053	18053	18053	18053	18053	18053		
1989	18970	19058	19058	19058	19058	19058	19058	19058	19058		18855	18855	18855	18855	18855	18855	18855		
1990	19328	19058	19058	19058	19058	19058	19058	19058	19058		19327	19327	19327	19327	19327	19327	19327		
1991	19337	19058	19058	19058	19058	19058	19058	19058	19058		19327	19327	19327	19327	19327	19327	19327		
1992	18876	19058	19058	19058	19058	19058	19058	19058	19058		18855	18855	18855	18855	18855	18855	18855		
MSE		98783	94001	78665	81691	40704	37093	36943	35464	17996	13264	13851	14313	13951	14591	15322	15429		
AFE		1.60	1.58	1.37	1.39	1.11	1.05	1.03	1.0	0.69	0.56	0.57	0.57	0.55	0.56	0.57	0.56		

TABLE V. COMPARISON OF MSE VALUES

Methods/MSE	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order
Chen [5]	89,093	86,694	89,376	94,539	98,215	104,056	102,179	102,789
Hwang et al. [39]	333,171	299,634	315,489	278,919	296,950	316,720	301,228	306,485
Singh [17]	119,189	97,180	126,676	113,421	163,137	148,618	169,149	123,964
Proposed (9 declared words)	98,783	94,001	78,665	81,691	40,704	37,093	36,943	35,464

TABLE VI. COMPARISON OF AVERAGE ERROR (AFE) OF FORECASTS

Methods/MSE	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order
Song and Chissom [4]	3.15	3.89	4.37	4.41	4.49	4.35	4.45	4.23
Hwang et al. [39]	2.99	2.94	3.12	2.92	3.01	3.08	2.89	2.79
Singh [17]	1.80	1.56	1.74	1.68	2.07	1.89	1.98	1.65
Proposed (9 declared words)	1.60	1.58	1.37	1.39	1.11	1.05	1.03	1.0

3) The accuracy of the proposed high-order LTS forecasting models increases by allowing the current used word-set of the forecasting linguistic variable to grow.

A special feature of the LTS forecasting models is that it considers the forecasting linguistic variable \mathcal{X}_L as the linguistic counterpart of the numerical forecasting one. So, there is not any restriction on its declared word set's cardinality. This experimental study will show that the forecasting accuracy will considerably increase when the declared word-set of \mathcal{X}_L grows. In HA, the word-set of \mathcal{X}_L are generated automatically by acting the linguistic hedges on two generator words [24], [37]. The maximum length of words is limited by an integer k , so-called k -specificity level. For example, in *Step 1* of the proposed high-order forecasting model described above, we use two linguistic hedges L and V acting on two generator words, so when $k = 1, 2, 3, 4$ and 5 , we have $5, 9, 17, 33, 65$ generated words, respectively. To show the scalability of the word-set of \mathcal{X}_L , we assume that its cardinality grows from 9 to $17, 33$ and 65 words respectively corresponding to 3 -, 4 - and 5 -specificity levels. For each declared word-set cardinality, the fuzziness parameter values of \mathcal{X}_L should be adjusted suitably by human's experts as shown in *Table VII*. For example, in case of 17 declared words, the value of $m(c^-)$ is 0.527 and the values of $\mu(L)$ is 0.412 , etc.

Intuitively seen in *Table IV* and *Table VIII*, with the same order, the greater used linguistic words, the smaller corresponding MSE and AFE values we will have. Specifically, the 9 -order LTS forecasting model using 65 words has its MSE value of 283 and its AFE value of 0.07 , better than the respective ones of the other models with same order using $33, 17$ and 9 words in turn, respectively, 1023 and $1.4, 15429$ and $0.6, 35464$ and 1.0 . We also have the same analysis results for other order (from 2 to 8) using the same words. These comparison results state that the forecasting accuracy will considerably increase when the declared word-set of \mathcal{X}_L grows (from 9 words to 65 words) and show the effectiveness of the proposed high-order linguistic forecasting method.

B. Forecast the Lahi Crop Production of Indian

To prove that the proposed high-order LTS forecasting models are much more efficient than the ones proposed by Singh in [17], the proposed models are applied to forecast the Lahi crop production (productivity in kg per hectare) from the farm of G.B. Pant University, Pantnagar, India, observed from 1981 to 2003 [17]. With the observed production values are from 440 and 1067 , we can set the value interval of the universe of discourse of the variable \mathcal{X}_L to be $U_{\mathcal{X}} = [400, 1100]$ as in [17]. Two primary words are $c^- = low (l)$ and $c^+ = high (h)$ and two linguistic hedges are *Little (L) ∈ H⁻* and *Very (V) ∈ H⁺*. For illustration, the declared word-set of \mathcal{X}_L in case of 2 -specificity level is $U_{\mathcal{X},L} = X_{(2)} = \{0, Very\ low, low, Little\ low, medium, Little\ high, high, Very\ high, 1\}$. The fuzziness parameter values are chosen by human's experts as $m(c^-) = m(low) = 0.5$ and $\mu(L) = 0.51$ for all various orders and declared words. So, the real numerical semantics of the declared words are $U_{\mathcal{X},R} = \{400, 484, 571.5, 662.5, 750, 837.5, 928.5, 1016, 1100\}$. Based on the given data setting, transform the observed historical data of the Lahi crop production shown the column "Actual production" in *Table IX* to the linguistic time series, establish the high-order LLRs and the high-order LLRGs, finally calculate the crisp forecasted values. The crop production forecasted results of the various high-order (from 2 to 9) LTS forecasting models are shown *Table IX*.

The comparisons of MSE and AFE values between the proposed LTS forecasting models with various orders using nine words and the corresponding ones proposed by Singh in [8] are shown in *Table X*. Indeed, the comparison results prove that the proposed high-order LTS forecasting models are much better than the ones proposed by Singh on both evaluation methods of MSE and AFE.

TABLE VII. FUZZINESS PARAMETER VALUES

17 words		33 words		65 words	
$m(c^-)$	$\mu(L)$	$m(c^-)$	$\mu(L)$	$m(c^-)$	$\mu(L)$
0.527	0.412	0.65	0.35	0.65	0.35

TABLE VIII. FORECASTED ENROLLMENTS BY PROPOSED HIGH-ORDER LTS FORECASTING MODELS USING 33 AND 65 DECLARED WORDS

Year	Actu. Enrol.	Proposed (33 declared words)									Proposed (65 declared words)								
		2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order		
1971	13055																		
1972	13563																		
1973	13867	13812								13812									
1974	14696	14485	14485							14638	14638								
1975	15460	15520	15520	15520						15438	15438	15438							
1976	15311	15285	15285	15285	15285					15285	15285	15285	15285						
1977	15603	15520	15520	15520	15520	15520				15673	15673	15673	15673	15673					
1978	15861	15958	15958	15958	15958	15958	15958			15958	15958	15958	15958	15958	15958				
1979	16807	16750	16750	16750	16750	16750	16750	16750		16795	16795	16795	16795	16795	16795	16795			
1980	16919	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877		
1981	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388	16388		
1982	15433	15520	15520	15520	15520	15520	15520	15520	15520	15438	15438	15438	15438	15438	15438	15438	15438		
1983	15497	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520	15520		
1984	15145	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158		
1985	15163	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158	15158		
1986	15984	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958	15958		
1987	16859	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877	16877		
1988	18150	18176	18176	18176	18176	18176	18176	18176	18176	18152	18152	18152	18152	18152	18152	18152	18152		
1989	18970	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965	18965		
1990	19328	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327		
1991	19337	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327	19327		
1992	18876	18838	18838	18838	18838	18838	18838	18838	18838	18882	18882	18882	18882	18882	18882	18882	18882		
MSE		4234	4297	2062	1972	2053	1731	1182	1023	1284	1192	1078	1106	1133	882	273	283		
AFE		0.28	0.28	0.21	0.20	0.20	0.18	0.15	0.14	0.16	0.15	0.13	0.13	0.13	0.11	0.07	0.07		

TABLE IX. FORECASTED PRODUCTIONS BY PROPOSED HIGH-ORDER LTS FORECASTING MODELS USING 9 DECLARED WORDS

Year	Actual production	Linguistic time series	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order
1981	1025	<i>V_high</i>								
1982	512	<i>V_low</i>								
1983	1005	<i>V_high</i>	1016							
1984	852	<i>L_high</i>	837	837						
1985	440	<i>A_low</i>	400	400	400					
1986	502	<i>V_low</i>	484	484	484	484				
1987	775	<i>Medium</i>	750	750	750	750	750			
1988	465	<i>V_low</i>	484	484	484	484	484	484		
1989	795	<i>L_high</i>	837	837	837	837	837	837	837	
1990	970	<i>high</i>	929	929	929	929	929	929	929	929
1991	742	<i>medium</i>	750	750	750	750	750	750	750	750
1992	635	<i>L_low</i>	663	663	663	663	663	663	663	663
1993	994	<i>V_high</i>	1016	1016	1016	1016	1016	1016	1016	1016
1994	759	<i>medium</i>	750	750	750	750	750	750	750	750
1995	883	<i>high</i>	929	929	929	929	929	929	929	929
1996	599	<i>low</i>	572	572	572	572	572	572	572	572
1997	499	<i>V_low</i>	484	484	484	484	484	484	484	484
1998	590	<i>low</i>	572	572	572	572	572	572	572	572
1999	911	<i>high</i>	929	929	929	929	929	929	929	929
2000	862	<i>L_high</i>	837	837	837	837	837	837	837	837
2001	801	<i>L_high</i>	837	837	837	837	837	837	837	837
2002	1067	<i>A_high</i>	1100	1100	1100	1100	1100	1100	1100	1100
2003	917	<i>high</i>	929	929	929	929	929	929	929	929
MSE			713.6	743.3	770.5	724.4	748	755.7	782	711.9
AFE			3.3707	3.4846	3.5753	3.2689	3.2502	3.2518	3.1961	3.0471

TABLE X. COMPARISON OF MSE AND AFE VALUES

Evaluation method	Forecasting models	2-order	3-order	4-order	5-order	6-order	7-order	8-order	9-order
MSE	Singh [17]	1089.7	908.9	1191.5	1311.6	1597.0	1207.2	1259.9	1425.8
	Proposed	713.6	743.3	770.5	724.4	748	755.7	782	711.9
AFE	Singh [17]	4.1359	3.8137	4.3566	4.4677	4.8753	3.9747	4.1337	4.5543
	Proposed	3.3707	3.4846	3.5753	3.2689	3.2502	3.2518	3.1961	3.0471

IV. CONCLUSIONS

The aims of the study in this paper is to establish a formalized methodology for handling words directly to solve the time series forecasting problems in such a way that it can simulate the way human users making a forecasting process in terms of their words. In fact, when observing the numerical time series data, human users can quickly transform its variations into the word variations which form a linguistic time series. They can do reasoning in their mind in some ways based on the relationships of the words in linguistic time series to estimate the forecasting values.

This paper focuses on examining the high-order linguistic time series to establish an efficient linguistic forecasting method in such a way that a formal formalism is developed to ensure that the inherent semantics of words, both qualitative and quantitative, are defined in the full linguistic variable context and the declared word-set of the linguistic variable can grow to increase the forecasted results. The growing of the declared word-set means that the human domain knowledge is grown. It is true in practice because human knowledge grows day by day. So, when human experts want to increase the forecasted results, he can make the declared word-set grow whilst keeps all semantic types of the existing words unchanged.

The proposed forecasting method is applied to the forecasting problems of the enrollments of University of Alabama and Lahi crop production of India. The experimental results have shown that the proposed forecasting method is better than the ones proposed by Song and Chissom, Chen, Hwang and Singh and the growth of the declared word-set of the linguistic variable is significant in increasing the accuracy of forecasted results of the high-order linguistic forecasting method. Therefore, that proves the effectiveness of our proposed approach.

REFERENCES

[1] L. A. Zadeh, "Fuzzy Sets," *Inf. Control*, vol. 8, pp. 338–353, 1965, doi: 10.1016/S0019-9958(65)90241-X.

[2] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets Syst.*, vol. 54, pp. 269–277, 1993.

[3] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series - part 1," *Fuzzy Sets Syst.*, vol. 54, pp. 1–9, 1993.

[4] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series - part 2," *Fuzzy Sets Syst.*, vol. 62, pp. 1–8, 1994.

[5] S. M. Chen, "Forecasting enrollments based on high-order fuzzy time series," *Cybern. Syst.*, vol. 33, no. 1, pp. 1–16, 2002, doi: 10.1080/019697202753306479.

[6] S.-M. Chen and N.-Y. Chung, "Forecasting Enrollments Using High-Order Fuzzy Time Series and Genetic Algorithms," *Int. J. Intell. Syst.*, vol. 21, pp. 485–501, 2006.

[7] L.-W. Lee, S.-M. Chen, Y.-H. Leu, and L.-H. Wang, "Handling Forecasting Problems Based on Two-Factors High-Order Fuzzy Time Series," *IEEE Trans. Fuzzy Syst.*, vol. 14, p. 10, 2006.

[8] K. K. Gupta and S. Kumar, "A novel high-order fuzzy time series forecasting method based on probabilistic fuzzy sets," *Granul. Comput.*, no. 2016, 2019, doi: 10.1007/s41066-019-00168-4.

[9] P. Jiang, Q. Dong, P. Li, and L. Lian, "A novel high-order weighted fuzzy time series model and its application in nonlinear time series prediction," *Appl. Soft Comput.*, vol. 55, pp. 44–62, 2017, doi: 10.1016/j.asoc.2017.01.043.

[10] M. Bose and K. Mali, "A novel data partitioning and rule selection technique for modeling high-order fuzzy time series," *Appl. Soft Comput.*, vol. 63, pp. 87–96, 2018, doi: 10.1016/j.asoc.2017.11.011.

[11] E. Bas, C. Grosan, E. Egrioglu, and U. Yolcu, "High order fuzzy time series method based on pi-sigma neural network," *Eng. Appl. Artif. Intell.*, vol. 72, pp. 350–356, Jun. 2018, doi: 10.1016/j.engappai.2018.04.017.

[12] P. Singh and Y.-P. Huang, "A High-Order Neutrosophic-Neuro-Gradient Descent Algorithm-Based Expert System for Time Series Forecasting," *Int. J. Fuzzy Syst.*, 2019, doi: 10.1007/s40815-019-00690-2.

[13] S. Panigrahi and H. S. Behera, "A study on leading machine learning techniques for high order fuzzy time series forecasting," *Eng. Appl. Artif. Intell.*, vol. 87, no. November 2019, p. 103245, 2020, doi: 10.1016/j.engappai.2019.103245.

[14] R. M. Pattanayak, S. Panigrahi, and H. S. Behera, "High-Order Fuzzy Time Series Forecasting by Using Membership Values Along with Data and Support Vector Machine," *Arab. J. Sci. Eng.*, no. 1, 2020, doi: 10.1007/s13369-020-04721-1.

[15] R. Mohan, H. S. Behera, and S. Panigrahi, "A novel probabilistic intuitionistic fuzzy set based model for high order fuzzy time series forecasting Engineering Applications of Artificial Intelligence A novel probabilistic intuitionistic fuzzy set based model for high order fuzzy time series forecas," *Eng. Appl. Artif. Intell.*, vol. 99, no. March, p. 104136, 2021, doi: 10.1016/j.engappai.2020.104136.

[16] E. Egrioglu, C. H. Aladag, U. Yolcu, V. R. Uslu, and M. A. Basaran, "A new approach based on artificial neural networks for high order multivariate fuzzy time series," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10589–10594, 2009, doi: 10.1016/j.eswa.2009.02.057.

[17] S. R. Singh, "A computational method of forecasting based on high-order fuzzy time series," *Expert Syst. Appl.*, vol. 36, no. 7, pp. 10551–10559, 2009, doi: 10.1016/j.eswa.2009.02.061.

[18] S. M. Chen and C. D. Chen, "Handling forecasting problems based on high-order fuzzy logical relationships," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3857–3864, 2011, doi: 10.1016/j.eswa.2010.09.046.

[19] C. H. Aladag, U. Yolcu, and E. Egrioglu, "A high order fuzzy time series forecasting model based on adaptive expectation and artificial neural networks," *Math. Comput. Simul.*, vol. 81, no. 4, pp. 875–882, 2010, doi: 10.1016/j.matcom.2010.09.011.

[20] S. S. Gangwar and S. Kumar, "Partitions based computational method for high-order fuzzy time series forecasting," *Expert Syst. Appl.*, vol. 39, no. 15, pp. 12158–12164, 2012, doi: 10.1016/j.eswa.2012.04.039.

[21] M. Y. Chen, "A high-order fuzzy time series forecasting model for internet stock trading," *Futur. Gener. Comput. Syst.*, vol. 37, pp. 461–467, 2014, doi: 10.1016/j.future.2013.09.025.

[22] S. M. Chen and S. W. Chen, "Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups and the probabilities of trends of fuzzy logical relationships," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 405–417, 2015, doi: 10.1109/TCYB.2014.2326888.

[23] E. Egrioglu, C. H. Aladag, U. Yolcu, and A. Z. Dalar, "A Hybrid High Order Fuzzy Time Series Forecasting Approach Based on PSO and ANNs Methods," *Am. J. Intell. Syst.*, vol. 6, no. 1, p. 8, 2016.

[24] N. C. Ho and W. Wechler, "Hedge Algebras: An algebraic approach to structure of sets of linguistic truth values," *Fuzzy Sets Syst.*, vol. 35, pp. 281–293, 1990.

- [25] N. C. Ho, V. N. Lan, and L. X. Viet, "Optimal hedge-algebras-based controller: Design and application," *Fuzzy Sets Syst.*, vol. 159, p. 32, 2008.
- [26] C. H. Nguyen, D. A. Nguyen, N. L. Vu, and D. A. Nguyen, "Fuzzy Controller using Hedge Algebra based semantics of vague linguistic terms," in *Fuzzy Control Systems*, D. Vukadinovic, Ed. Nova Science Publishers, Inc., 2011.
- [27] D. Vukadinovic, M. Basic, C. H. Nguyen, N. L. Vu, and T. D. Nguyen, "Hedge-algebra-based voltage controller for a self-excited induction generator," *Control Eng. Pract.*, vol. 30, pp. 78–90, 2014.
- [28] H.-L. Bui, C.-H. Nguyen, V.-B. Bui, K.-N. Le, and H.-Q. Tran, "Vibration control of uncertain structures with actuator saturation using hedge-algebras-based fuzzy controller," *J. Vib. Control*, p. 19, 2015.
- [29] D. T. Tran, V. B. Bui, T. A. Le, and H. Le Bui, "Vibration control of a structure using sliding-mode hedge-algebras-based controller," *Soft Comput.*, vol. 23, no. 6, pp. 2047–2059, 2017.
- [30] H. Le Bui, T. A. Le, and V. B. Bui, "Explicit formula of hedge-algebras-based fuzzy controller and applications in structural vibration control," *Appl. Soft Comput.*, vol. 60, pp. 150–166, 2017.
- [31] H.-L. Bui, N.-L. Vu, C.-H. Nguyen, and C.-H. Nguyen, "General design method of hedge-algebras-based fuzzy controllers and an application for structural active control," *Appl. Intell.*, vol. 43, p. 25, 2015.
- [32] C. H. Nguyen, T. S. Tran, and D. P. Pham, "Modeling of a semantics core of linguistic terms based on an extension of hedge algebra semantics and its application," *Knowledge-Based Syst.*, vol. 67, p. 19, 2014.
- [33] C. H. Nguyen, V. T. Hoang, and V. L. Nguyen, "A discussion on interpretability of linguistic rule based systems and its application to solve regression problems," *Knowledge-Based Syst.*, vol. 88, pp. 107–133, 2015.
- [34] H. H. Ngo, C. H. Nguyen, and V. Q. Nguyen, "Multichannel image contrast enhancement based on linguistic rule-based intensifiers," *Appl. Soft Comput.*, vol. 76, pp. 744–763, 2019.
- [35] N. D. Hieu, N. C. Ho, and V. N. Lan, "Enrollment Forecasting Based on Linguistic Time Series," *J. Comput. Sci. Cybern.*, vol. 36, no. 2, pp. 119–137, 2020, doi: 10.15625/1813-9663/36/2/14396.
- [36] N. D. Hieu, N. C. Ho, and V. N. Lan, "An efficient fuzzy time series forecasting model based on quantifying semantics of words," in *IEEE RIVF International Conference*, 2020, pp. 1–6.
- [37] N. C. Ho and W. Wechler, "Extended hedge algebras and their application to fuzzy logic," *Fuzzy Sets Syst.*, vol. 52, pp. 259–281, 1992.
- [38] N. C. Ho and N. Van Long, "Fuzziness measure on complete hedge algebras and quantifying semantics of terms in linear hedge algebras," *Fuzzy Sets Syst.*, vol. 158, pp. 452–471, 2007.
- [39] J. R. Hwang, S. M. Chen, C. H. Lee, Handling forecasting problems using fuzzy time series. *Fuzzy Sets and Systems*, vol. 100, pp. 217–228, 1998.

Opinion Mining of Saudi Responses to COVID-19 Vaccines on Twitter

A Computational Linguistic Approach

Fahad M. Alliheibi^{1*}

Department of Arabic Language, Faculty of Arts
King Abdulaziz University, Jeddah, Saudi Arabia

Abdulfattah Omar²

Department of English, College of Science and Humanities
Prince Sattam Bin Abdulaziz University, Saudi Arabia
Department of English, Faculty of Arts
Port Said University, Egypt

Nasser Al-Horais³

Department of Arabic Language and Arts
College of Arabic Language and Social Studies
Qassim University, P.O. Box 6611, Buraidah 51425
Saudi Arabia

Abstract—In recent months, many governments have announced COVID-19 vaccination programs and plans to help end the crises the world has been facing since the emergence of the coronavirus pandemic. In Saudi Arabia, the Ministry of Health called for citizens and residents to take up the vaccine as an essential step to return life to normal. However, the take-up calls were made in the face of profound disagreements on social media platforms and online networks about the value and efficacy of the vaccines. Thus, this study seeks to explore the responses of Saudi citizens to the COVID-19 vaccines and their sentiments about being vaccinated using opinion mining methods to analyze data extracted from Twitter, the most widely used social media network in Saudi Arabia. A corpus of 37,467 tweets was built. Vector space classification (VSC) methods were used to group and categorize the selected tweets based on their linguistic content, classifying the attitudes and responses of the users into three defined categories: positive, negative, and neutral. The lexical semantic properties of the posts show a prevalence of negative responses. This indicates that health departments need to ensure citizens are equipped with accurate, evidence-based information and key facts about the COVID-19 vaccines to help them make appropriate decisions when it comes to being vaccinated. Although the study is limited to the analysis of attitudes of people to the COVID-19 vaccines in Saudi Arabia, it has clear implications for the application of opinion mining using computational linguistic methods in Arabic.

Keywords—Computational linguistics; COVID-19 vaccines; lexical semantics; opinion mining; vector space classification

I. INTRODUCTION

The COVID-19 pandemic, which emerged in late 2019, has had and continues to have grave health, economic, social, and political consequences for almost all countries around the world [1-4]. So far, more than three million people have died of this coronavirus. In response, many governments have enforced lockdown laws, placing restrictions on the movement of citizens globally. This has resulted in serious economic problems for many countries, including the Arab Gulf states, which have suffered due to the lower demand for oil. As

pointed out by Kesar, et al. [5], the coronavirus pandemic has resulted in a large-scale humanitarian crisis leading to misery and suffering for all of humanity, and pushing social and economic well-being to the brink of collapse.

To alleviate the dangers of the COVID-19 pandemic, several vaccines have been developed, including the AstraZeneca, Johnson & Johnson's Janssen, Moderna, Novavax, and Pfizer-BioNTech vaccines. Given the sophisticated development and long authorization process for vaccinations, the COVID-19 vaccines were given temporary authorization for emergency use. Nevertheless, the manufacturers and researchers stressed the safety of the vaccines and their effectiveness in reducing the risks of severe illness associated with catching the virus. They confirmed that the new vaccines had undergone rigorous testing to verify their safety before they were recommended for widespread use.

In light of these developments, many governments have announced COVID-19 vaccination delivery plans to help end the crises the world has been facing since the emergence of the coronavirus pandemic. In Saudi Arabia, the Ministry of Health called for citizens and residents to take up the vaccine as an essential step in returning life to normal. However, these appeals were met by skepticism from many Saudis who had various fears about the vaccines. Furthermore, many activists fueled concerns about vaccine safety on social media platforms, especially Twitter. In response, officials and professionals stressed the safety of the vaccines, and called for citizens and residents to get the vaccines for the good and safety of everyone.

Twitter in particular was the forum for controversy and heated debates about the effectiveness and usefulness of the COVID-19 vaccines in Saudi Arabia. Millions of users voiced their opinions concerning the uptake of the vaccines on the platform. Apart from reflecting the issue of the effectiveness and safety of the vaccines, this clearly indicated the growing influence of social media platforms and networks on public opinion in Saudi Arabia, as noted by Gunter, et al. [6]. The

*Corresponding Author

increasing numbers of users of social media platforms and networks in Saudi Arabia is having a direct effect on their ever greater role and effectiveness in shaping public opinion concerning a range of social, economic, and political issues [6, 7].

Thus, it is important for policymakers and those in government to understand people's attitudes towards certain issues as represented on social media. In this regard, they could exploit the potential of opinion mining, defined as a computational linguistics science that uses text analysis and natural language processing (NLP) to understand the drivers of public sentiment through extracting people's opinions and sentiments from the Web in general, and social media networks and platforms in particular [8, 9]. Despite the importance of opinion mining applications in helping governments and organizations to make appropriate decisions concerning critical issues, especially at times of crisis, their potential has not yet been fully explored in Arab nations [10]. In response to this lack of research, our study has sought to explore Saudi responses and sentiments regarding the COVID-19 vaccination take-up calls.

The study addresses two key research questions. First, what were individuals' responses and sentiments concerning the COVID-19 vaccines in Saudi Arabia as expressed on Twitter? Second, can the opinions, attitudes, and responses of individuals regarding given issues be analyzed conceptually through computational linguistic methods in Arabic?

The remainder of the article is structured as follows. Section 2 reviews the development of opinion mining and outlines its theoretical underpinnings and its relationship with other disciplines, including sentiment analysis, NLP, and computational linguistics. Section 3 sets out the data collection methods and procedures and the corpus design. Section 4 details the computational linguistic analysis based on a vector space classification (VSC) of the selected tweets. Section 5 concludes, summarizing the main findings and highlighting the implications of the study for further research.

II. PREVIOUS WORK

Opinion mining is a rising field of study that is gaining prominence largely due to the growing influence of social media platforms and networks in today's world. Indeed, there has been an extraordinary increase in the number of social media networks, providing easy access to millions of users around the world to express their opinions freely on different issues. Given the daily growth in numbers of users around the world, social media networks have become indispensable in contemporary life, and are now integral to the economic, social, and political marketing mechanisms that are employed in many fields.

With the increasing influence of social media networks, opinion mining was developed to analyze human behavior through computational tools to help manufacturers, marketing agencies, branders, government institutions, organizations, and policymakers make appropriate decisions concerning their products, brands, services, and policies [11]. This caters to the growing need for businesses and organizations to take account

of public opinion expressed on social media networks and platforms in their decision making [12].

In the literature, opinion mining is commonly related to sentiment analysis, and the two terms are often used interchangeably [13-15]. They are both concerned with the study and analysis of human behavior as expressed on social media networks and platforms using computer modeling approaches. As Liu and Zhang [12] point out, they have in common that they entail the "study of people's opinions, appraisals, attitudes, and emotions regarding entities, individuals, issues, events, and topics and their attributes." Our study, however, considers that they differ in terms of depth, arguing that opinion mining goes to the next level in understanding the conversational drivers underpinning the sentiment.

The literature reports extensive studies that have been carried out to measure public opinion on different issues and policies through computational-based opinion mining methods. There is growing interest in opinion mining applications to analyze political issues due to the already vast and constantly increasing creation of political content on the Web and its exchange, particularly on social media networks and platforms [16-18]. The underlying principle in such studies is that the analysis of public opinions expressed on social media networks and platforms, if well done, "can be quite useful for government agencies, as it can significantly assist them to understand the needs and problems of society, and the perceptions and feelings of the citizens, and to formulate effective public policies" [16].

Despite the prolific and increasing literature on opinion mining applications, there have been very few studies based on Arabic content to date. This can be attributed to several reasons. First, as noted by Badaro et al. [19], "there are fewer freely [available] resources for Arabic opinion mining in terms of large clean sentiment lexica and annotated collections of opinions" compared to English and other western languages. In other words, most opinion mining methods have been developed for English texts. It is difficult to use these methods with Arabic, which is completely different from English [20]. Second, the morphological system of Arabic poses serious challenges for NLP applications with Arabic data [21-24]. As pointed out by Ryding [25], Arabic is a highly inflectional language. Unlike English, words in Arabic have a unique morphological system, which typically has negative effects on the reliability and effectiveness of NLP processes and applications. Finally, the diglossic nature of Arabic, and its different vernacular and colloquial dialects make it difficult for the classification systems of opinion mining applications to work properly [26-28].

It has only been over the last five years that opinion mining in Arabic has started to receive considerable interest from the research community [29]. In a recent study, Althagafi, et al. [30] studied and analyzed the attitudes and responses of people to the shift in schools and universities towards online learning in 2020 due to the COVID-19 pandemic. The authors used Arabic tweets to gain an understanding of the sentiments and opinions of Twitter users regarding online learning to help government, educational institutions, and policymakers

recognize the emotional effects of online learning on students, teachers, parents, and individuals in the Saudi community.

In another study, Omar and Hamouda [10] analyzed the responses of citizens to a new real estate law in Egypt through computational tools. They aimed to explore Egyptians' sentiments, opinions, and attitudes regarding the new Real Estate Registration Law based on Facebook posts, comments, and replies. The study concluded that the citizens' views of the new law were predominantly negative. The authors highlighted the effectiveness of applying opinion mining to examine responses to regulations and policies through citizens' responses on social media. Accordingly, they suggested that opinion mining applications should be viewed as presenting valuable opportunities for policymakers and legislators to gauge people's sentiments, attitudes, and perceptions.

Despite the importance of such studies in terms of introducing researchers to opinion mining applications in Arabic, there was no focus on the manipulation of computational linguistic tools or lexical semantic analysis in the interpretation of the data and users' opinions.

III. METHODS AND PROCEDURES

This study used Twitter posts to identify the responses of citizens and residents to the COVID-19 vaccines in Saudi Arabia since social media networks and platforms have become an integral part of the lives of the Saudis today and Saudi Arabia is one of the leading countries in terms of Internet penetration. According to a report released by the World Bank, 95% of the total population of Saudi Arabia was using the Internet in 2019. In 2021, the percentage increased to 95.7%, with mobile connections up to 112.7%. Furthermore, the number of active social media users amounted to 27.80 million, individuals constituting around 79.3 % of the total population [31], as shown in Fig. 1.

In light of these statistics, it can be claimed that social media networks can usefully be employed to identify the responses of individuals in Saudi Arabia to specific issues. Of the platforms available, Twitter is one of the most reliable and suitable sites for the application of opinion mining and sentiment analysis [32, 33], and as of January 2021, Saudi Arabia has been ranked one of the top countries in terms of Twitter users [34], as shown in Fig. 2.

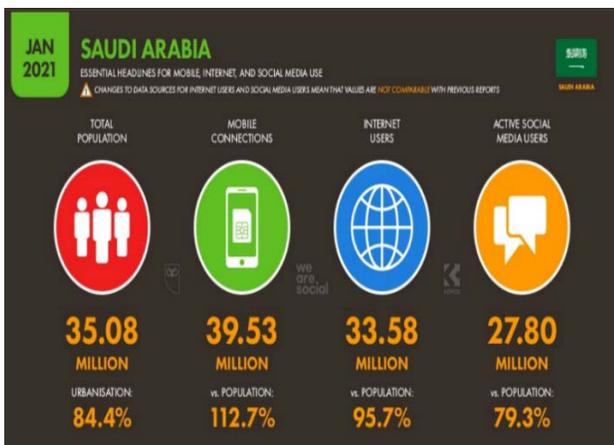


Fig. 1. The use of the Internet and Social Media in Saudi Arabia.

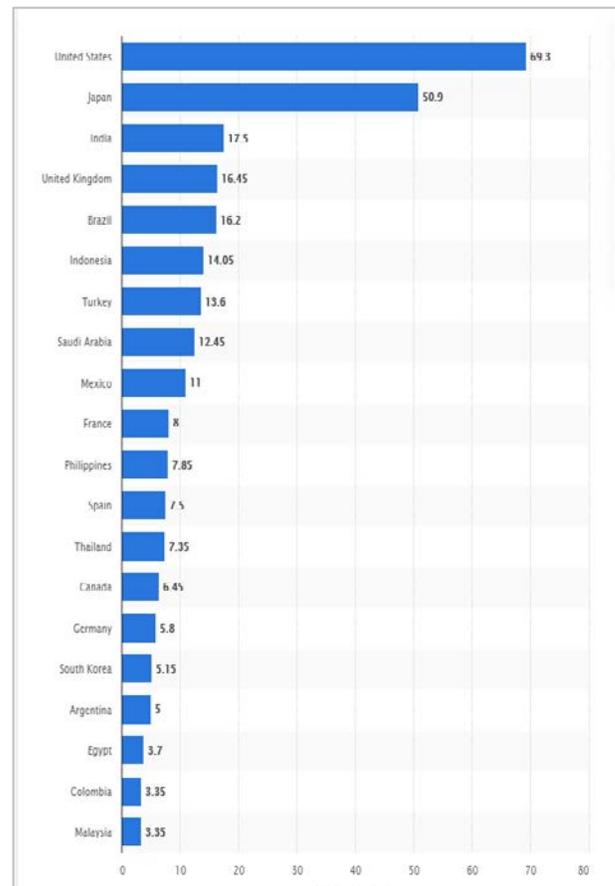


Fig. 2. Countries with the Most Twitter users in 2021.

Furthermore, Twitter was ranked topmost for information and posts on COVID-19 in Saudi Arabia. Finally, the Twitter account of the Saudi Ministry of Health has the highest number of followers, as shown in Fig. 3.

To extract information about users' polarity and responses concerning the issue of COVID-19 vaccines, four key terms were defined, as presented in Table I.

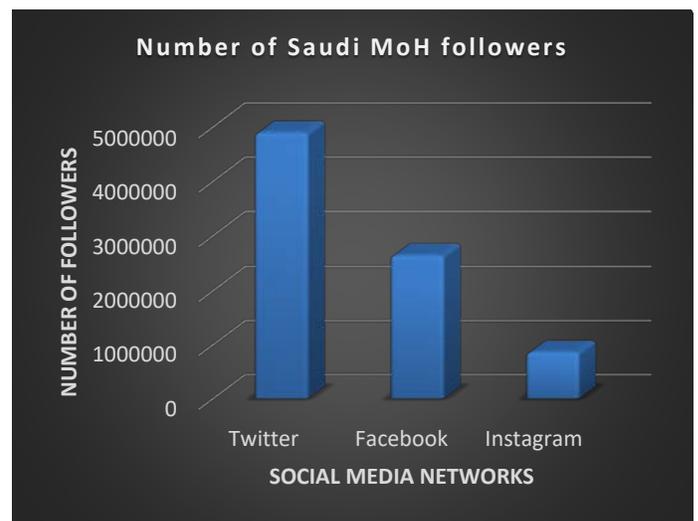


Fig. 3. Numbers of the Ministry of Health followers on Social Media Networks.

TABLE I. LIST OF THE DEFINED KEY TERMS

Key Terms	Transliteration	English Translation
لقاح كورونا	Liqah kooruna	Corona Virus vaccine*
تطعيم كورونا	Tat'eem kooruna	Corona virus vaccine*
التحصين	Altahsin	Vaccination
الحصول علي اللقاح	Alhusul eali alliqah	Vaccine uptake

* In Arabic, unlike English, coronavirus is expressed as two words.

Keywords were chosen rather than hashtags to ensure the data sample was as inclusive and representative as possible for generalizability and reliability purposes, since hashtags are generally used by more proficient social media users.

We built a corpus of 37,467 tweets from 23,748 users drawing on posts, comments on posts, and replies to posts on the topic of COVID-19 vaccines. The selected tweets covered the period from December 15, 2020, when the Ministry of Health called for the citizens and residents of Saudi Arabia to register for the COVID-19 vaccine, until May 25, 2021. Fig. 4 shows the first two citizens in Saudi Arabia who took the vaccine on December 17, 2020.

The selected tweets were cleaned and all non-alphabetical characters were removed. A stop-word list was also generated and all function words were removed. As a next step, a vector space matrix containing only the content or lexical words was designed and comprised 9,348 lexical types. There were fewer lexical types than tweets as many of the selected posts were repeated.

One critical problem with the matrix was the high dimensionality of the data, which commonly has adverse impacts on the reliability of classification [35]. High dimensionality results from a high volume of irrelevant information, which makes the classification process problematic and leads to poor, potentially unreliable, performance. As the number of vectors or attributes increases, the texts become increasingly homogeneous. For a data matrix in which the rows are the data items and the columns the variables, the matrix defines a manifold (the shape of the data in space) in \mathcal{N} -dimensional space. Let us assume a data matrix with 1,000 3-dimensional vectors. Plotting these vectors in 3-dimensional space will result in a cloud of points, as shown in Fig. 5.

Consequently, it becomes increasingly difficult to find meaningful and distinct groups based on such a dataset. The retention of unimportant variables in analysis gives invalid results [36]. It is thus important to retain only those variables that are most distinctive within a data matrix through dimensionality reduction, improving the reliability of classification performance [37]. To do this, two tests were carried out: variance and term frequency-inverse document frequency (TF-IDF). The rationale is that the classification is based on only the most distinctive lexical variables in the corpus.

First, all 9,348 lexical types were sorted in terms of variance based on the assumption that there must be variations in the characteristics of documents, distinguishing one from another, to enable classification. It follows that measuring variance within datasets can be useful in identifying the most

distinctive features within a matrix. The variance test indicated that the first 600 lexical types were the most significant variables in terms of variance, as shown in Fig. 6.



Fig. 4. The Initiation of the Vaccination Program in Saudi Arabia.

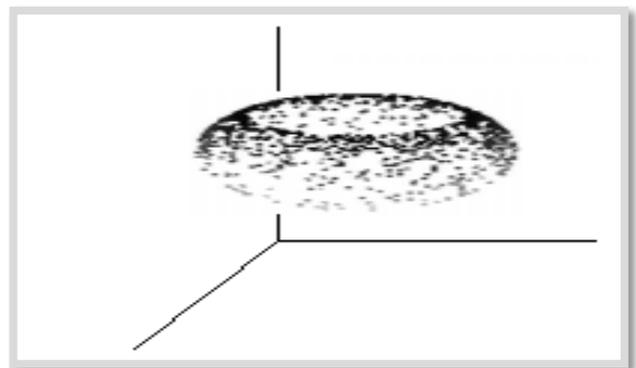


Fig. 5. A Manifold in 3-dimensional Space.

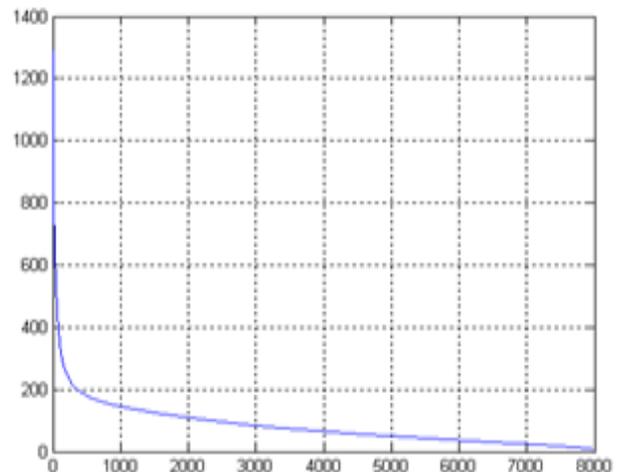


Fig. 6. Measuring Variables in Terms of Variance.

As such, only the lexical types presenting significant variation were retained and those with low or no variation were deleted. This resulted in a reduction in lexical types within the matrix to just 600 variables. As a final step in the data dimensionality reduction process, a TF-IDF analysis was carried out. The 600 lexical types were sorted in descending order in terms of TF-IDF using the following function:

$$tfidf(t_i) = tf(t_i) \log_2 \left[\frac{M}{df_1} \right]$$

where $tf(t_i)$ is the frequency of term t_i across all documents in the data matrix M . The result is shown in Fig. 7.

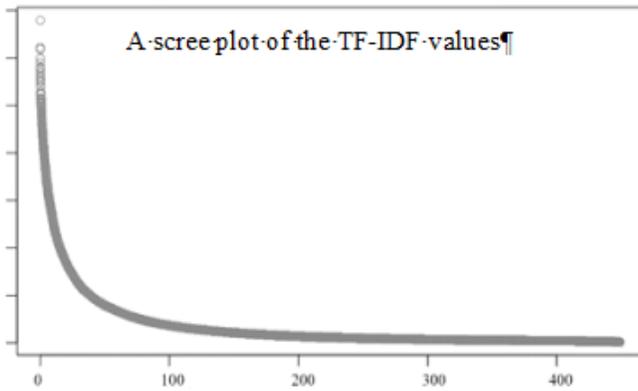


Fig. 7. Term Weighting using TF-IDF.

In this process, only the 60 lexical types with the highest TF-IDF were retained. These are assumed to be the most distinctive features within the data matrix.

IV. ANALYSIS AND DISCUSSION

To classify each opinion concerning the issue as positive, negative, or neutral, a computational lexical semantic analysis was conducted based on vector space classification (VSC). VSC is one of the most reliable classification systems in data mining applications, including information retrieval (IR), annotation and summarization, sentiment analysis, and opinion mining. VSC is based on measuring the relative distances between row vectors. The size of the angle between the lines that joins vectors to the origin of the space's coordinate system and the lengths of the lines jointly determine the distance between any two vectors in a space [38]. In VSC, for each selected variable, its semantics are used to obtain an interpretation of the domain of interest, the domain thus being measured based on the semantics [39-41].

The classification is formed using opinion words and phrases, these being the dominant indicators for opinion classification. Liu and Zhang [12] define opinion words as "words that are commonly used to express positive or negative sentiments. For example, beautiful, wonderful, good, and amazing are positive opinion words, and bad, poor, and terrible are negative opinion words. Although many opinion words are adjectives and adverbs, nouns (e.g., rubbish, junk, and crap) and verbs (e.g., hate and like) can also indicate opinions. Apart from individual words, there are also opinion phrases and idioms, e.g., cost someone an arm and a leg." The 37,467

tweets were classified into three categories or classes: positive, negative, and neutral. The results are shown in Fig. 8.

It is clear that most responses to the COVID-19 vaccine take-up calls are negative. The negative responses account for approximately 54%, positive responses 34%, and neutral responses 12%.

Positive responses were reflected in words such as (فعال) effective, (آمن) safe, (طبيعية) normality, (يحمي) protect, and (وقاية) protection. These focused on the concept of reducing the risks associated with COVID-19 infection, especially for the elderly, as shown in the examples given in Fig. 9.

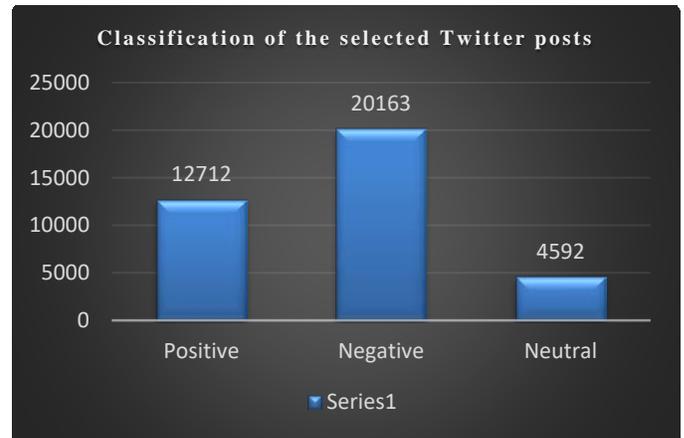


Fig. 8. Results of the Classification of the Selected Twitter Posts.



Fig. 9. Examples of Positive Responses to the COVID-19 Vaccines.

Negative responses, in turn, included words such as (أرفض) reject, (لا) No, (العقم) infertility, (جيش) Gates, (مؤامرة) conspiracy (الجلطات) clots, and (الحامض النووي) DNA. These responses focused on the ineffectiveness of the vaccines and that they had not been carefully evaluated in clinical trials. They also linked the COVID-19 vaccines to infertility and making changes within the DNA of individuals, as seen in the examples given in Fig. 10.

Furthermore, many posts claimed that COVID-19 vaccination was cover for a plan devised by Bill Gates to implant trackable microchips to control people. It is therefore not surprising that many hashtags rejecting vaccination were generated in response to the Saudi government's call for vaccine uptake. These included:

#لا_لربط_السفر_بالتطعيم (No vaccine certificates for travel).

#لا_للتطعيم_الاجباري (No for mandatory vaccination).

#نرفض_اللقاح_الاجباري (We reject mandatory vaccination).

#نرفض_التطعيم_الاجباري (We reject mandatory vaccination).

Based on these findings, it can be claimed that computational lexical semantic and opinion mining methods can usefully be employed to define people's attitudes and responses to given issues. It was clear that the classification of the Twitter posts based on their lexical semantic properties was meaningful. In particular, it was a straightforward matter to define polarity concerning the issue of COVID-19 vaccine uptake in Saudi Arabia.



Fig. 10. Examples of Negative Responses to the COVID-19 Vaccines.

V. CONCLUSION

This study analyzed the responses of individuals to COVID-19 vaccines in Saudi Arabia through their Twitter posts. To do so, computational lexical semantic methods were used for the classification of people's responses. The results clearly indicate that most people in Saudi Arabia hold negative attitudes concerning the COVID-19 vaccination take-up programs established by the government. The findings of the study illustrate the importance of understanding people's sentiments, opinions, and attitudes regarding government decisions and policies. The results argue for the benefits of government institutions and departments making use of the potential of data mining applications to identify trends that may have adverse impacts on policies and practices, as well as to help government institutions make appropriate decisions, and adopt reliable and workable policies and procedures. In view of the significance of social media networks and platforms in modern societies, opinion mining applications should be an integral part of all government processes. Although the study is limited to analysis of the attitudes of people to the COVID-19 vaccines in Saudi Arabia, it has clear implications in terms of the value of opinion mining applications using computational linguistic methods in Arabic.

ACKNOWLEDGMENT

The corresponding author would like to acknowledge the support provided by the Deanship of Scientific Research (DSR) in King Abdulaziz University, Jeddah, Saudi Arabia.

REFERENCES

- [1] Z. Allam, Surveying the Covid-19 Pandemic and Its Implications: Urban Health, Data Technology and Political Economy. Amsterdam: Elsevier Science, 2020.
- [2] M. Nicola et al., "The socio-economic implications of the coronavirus and COVID-19 pandemic: a review," International journal of surgery, 2020.
- [3] S. Gössling, D. Scott, and C. M. Hall, "Pandemics, tourism and global change: a rapid assessment of COVID-19," Journal of Sustainable Tourism, vol. 29, no. 1, pp. 1-20, 2020.
- [4] D. Bhuiyan and A. Bashar, "The covid-19 pandemic: conceptual model for the global economic impacts and recovery," The Covid-19 Pandemic: Conceptual Model for the Global Economic Impacts and Recovery (July 20, 2020), 2020.
- [5] S. Kesar, R. Abraham, R. Lahoti, P. Nath, and A. Basole, "Pandemic, informality, and vulnerability: Impact of COVID-19 on livelihoods in India," CSE Working Paper, vol. 2020, 2020.
- [6] B. Gunter, M. Elareshi, and K. Al-Jaber, Social media in the Arab world: Communication and public opinion in the Gulf states. Bloomsbury Publishing, 2016.
- [7] H. T. Alsudairy, The Role of the Social Media in Empowering Saudi Women's Expression. Newcastle, UK: Cambridge Scholars Publishing, 2020.
- [8] B. Pang and L. Lee, Opinion Mining and Sentiment Analysis. Delft, the Netherlands: Now Publishers, 2008.
- [9] B. Liu, Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. New York: Cambridge University Press, 2015.
- [10] A. Omar and W. Hamouda, "A Sentiment Analysis of Egypt's New Real Estate Registration Law on Facebook," International Journal of Advanced Computer Science and Applications, vol. 12, no. 4, pp. 656-663, 2021.
- [11] S. Choudhary, "Opinion Mining Techniques and Its Applications: A Review.," in Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019), vol. 121, P. Singh, W. Pawłowski, S. Tanwar, N. Kumar, J. Rodrigues, and

- M. Obaidat, Eds. (Lecture Notes in Networks and Systems, Singapore: Springer, 2020, pp. 549-559.
- [12] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in Mining text data New York; London: Springer, 2012, pp. 415-463.
- [13] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," IEEE Intelligent systems, vol. 28, no. 2, pp. 15-21, 2013.
- [14] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: tasks, approaches and applications," Knowledge-based systems, vol. 89, pp. 14-46, 2015.
- [15] P. K. Singh and M. S. Husain, "Methodological study of opinion mining and sentiment analysis techniques," International Journal on Soft Computing, vol. 5, no. 1, p. 11, 2014.
- [16] Y. Charalabidis, M. Maragoudakis, and E. Loukis, "Opinion Mining and Sentiment Analysis in Policy Formulation Initiatives: The EU-Community Approach," in Electronic Participation, vol. 9249, E. Tambouris et al., Eds. (Lecture Notes in Computer Science, Cham: Springer, 2015, pp. 147-160.
- [17] V. Bekkers, A. R. Edwards, R. Moody, and H. Beunders, "Caught by surprise? Micro-mobilization, new media and the management of strategic surprises," Public Management Review, vol. 13, no. 7, pp. 1003-1021, 2011.
- [18] E. Loukis and Y. Charalabidis, "Active and passive crowdsourcing in government," in Policy Practice and Digital Science: Integrating Complex Systems, Social Simulation and Public Administration in Policy Research, M. Janssen, M. Wimmer, and A. Deljoo, Eds. (Public Administration and Information Technology Series, Cham: Springer Verlag, 2014, pp. 261-289.
- [19] G. Badaro et al., "A Survey of Opinion Mining in Arabic: A Comprehensive System Perspective Covering Challenges and Advances in Tools, Resources, Models, Applications, and Visualizations," ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 18, no. 3, pp. 27:1-52, 2019.
- [20] M. N. Al-Kabi, A. H. Gigieh, I. M. Alsmadi, and H. A. Wahsheh, "Opinion Mining and Analysis for Arabic Language," International Journal of Advanced Computer Science and Applications, vol. 5, no. 5, pp. 181-195, 2014.
- [21] A. Omar and M. Aldossari, "Lexical Ambiguity in Arabic Information Retrieval: The Case of Six Web-Based Search Engines," International Journal of English Linguistics, vol. 10 no. 3, pp. 219-228, 2020.
- [22] A. Omar, B. Alghayesh, and M. Kassem, "Authorship Attribution Revisited: The Problem of Flash Fiction A morphological-based Linguistic Stylometry Approach," Arab World English Journal, vol. 10, no. 3, pp. 318-329, 2019.
- [23] A. Omar and W. Hamouda, "The Effectiveness of Stemming in the Stylometric Authorship Attribution in Arabic," International Journal of Advanced Computer Science and Applications, vol. 11 no. 1, pp. 116-122, 2020.
- [24] A. Omar, I. Shaalan, and W. Hamouda, "Ambiguity Resolution in Arabic Localization.," Applied Linguistics Research Journal, vol. 5 no. 1, pp. 1-6, 2021.
- [25] K. Ryding, "Arabic inflectional morphology," in Arabic: A Linguistic Introduction Cambridge: Cambridge University Press, 2014, pp. 89-106.
- [26] A. Farghaly and K. Shaalan, "Arabic Natural Language Processing: Challenges and Solutions," ACM Transactions on Asian Language Information Processing, vol. 8, no. 4, p. Article 14, 2009.
- [27] A. Omar, "An Evaluation of the Localization Quality of the Arabic Versions of Learning Management Systems," International Journal of Advanced Computer Science and Applications, vol. 12, no. 2, pp. 443-449, 2021.
- [28] A. Omar, "Authorship attribution of Morsi Gameel Aziz's lyrics: A clustering-based stylometry approach," Journal of Language and Linguistic Studies, vol. 17, no. 1, pp. 542-557, 2021.
- [29] N. Boudad, R. Faizi, R. Oulad Haj Thami, and R. Chiheb, "Sentiment analysis in Arabic: A review of the literature," Ain Shams Engineering Journal, vol. 9, no. 4, pp. 2479-2490, 2018/12/01/ 2018.
- [30] A. Althagafi, G. Althobaiti, H. Alhakami, and T. Alsabait, "Arabic Tweets Sentiment Analysis about Online Learning during COVID-19 in Saudi Arabia" International Journal of Advanced Computer Science and Applications(IJACSA)," vol. 12, no. 3, pp. 620-625, 2021.
- [31] S. Kemp, "Digital 2021: Saudi Arabia," Data Reportal 10 February 2021 2021, Available: <https://datareportal.com/reports/digital-2021-saudi-arabia>.
- [32] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, A Practical Guide to Sentiment Analysis. New York: Springer International Publishing, 2017.
- [33] G. Gupta and G. S. Bhatthal, Sentiment Analysis of English Tweets Using Data Mining: Data Mining, Sentiment Analysis. Munich, Germany: BookRix, 2018.
- [34] H. Tankovska, "Countries with the most Twitter users 2021," Statista, New York, NY Feb 9, 2021 2021.
- [35] A. Omar, "Feature Selection in Text Clustering Applications of Literary Texts: A Hybrid of Term Weighting Methods," International Journal of Advanced Computer Science and Applications, vol. 11 no. 2, pp. 99-107, 2020.
- [36] Y. Shao, Topics in Dimension Reduction. Minneapolis, Minnesota: University of Minnesota, 2007.
- [37] P. D. Waggoner, Modern Dimension Reduction. Cambridge: Cambridge University Press, 2021.
- [38] H. Moisl, Cluster Analysis for Corpus Linguistics. Berlin: De Gruyter, 2015.
- [39] A. Omar, "Identifying themes in fiction: A centroid-based lexical clustering approach," Journal of Language and Linguistic Studies, vol. 17, no. Special Issue 1, 580-594 2021.
- [40] A. Omar, "The Detective and Sensation Fiction of Wilkie Collins: A Computational Lexical-Semantic Analysis," Arab World English Journal (AWEJ), vol. 11 no. 1, pp. 195-211, 2020.
- [41] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval. Cambridge: Cambridge University Press, 2008.

AUTHORS' PROFILE

Fahad Alliheibi is a Full Professor of Arabic Language and Linguistics in the Department of Arabic Language, Faculty of Arts, King Abdulaziz University (KSA). Prof. Alliheibi received his PhD degree in Linguistics in 1999 from Durham University, UK. His research interests include Arabic Linguistics, Pragmatics, Text Linguistics, and Translation. ORCID: 0000-0002-2233-5165

Abdulfattah Omar is an Associate Professor of English Language and Linguistics in the Department of English, College of Science & Humanities, Prince Sattam Bin Abdulaziz University (KSA). Also, he is a standing lecturer of English Language and Linguistics in the Department of English, Faculty of Arts, Port Said University, Egypt. Dr. Omar received his PhD degree in computational linguistics in 2010 from Newcastle University, UK. His research interests include computational linguistics, digital humanities, discourse analysis, and translation studies. ORCID: 0000-0002-3618-1750

Nasser Al-Horais is a Full Professor of Arabic Language and Linguistics in the Department of Arabic Language and its Arts, College of Arabic Language & Social Studies, Qassim University (KSA). Prof. Al-Horais received his PhD degree in Linguistics in 2009 from Newcastle University, UK. His research interests include Arabic Syntax, Generative Linguistics (Minimalist Program), Discourse Analysis, and Comparative Syntax. ORCID: 0000-0002-2511-9791

Indoor Localization and Navigation based on Deep Learning using a Monocular Visual System

Rodrigo Eduardo Arevalo Ancona¹, Leonel Germán Corona Ramírez², Oscar Octavio Gutiérrez Frías³

UPIITA-IPN SEPI Section, Instituto Politécnico Nacional
Mexico City, Mexico

Abstract—Now-a-days, computer systems are important for artificial vision systems to analyze the acquired data to realize crucial tasks, such as localization and navigation. For successful navigation, the robot must interpret the acquired data and determine its position to decide how to move through the environment. This paper proposes an indoor mobile robot visual-localization and navigation approach for autonomous navigation. A convolutional neural network and background modeling are used to locate the system in the environment. Object detection is based on copy-move detection, an image forensic technique, extracting features from the image to identify similar regions. An adaptive threshold is proposed due to the illumination changes. The detected object is classified to evade it using a control deep neural network. A U-Net model is implemented to track the path trajectory. The experiment results were obtained from real data, proving the efficiency of the proposed algorithm. The adaptive threshold solves illumination variation issues for object detection.

Keywords—Visual localization; visual navigation; autonomous navigation; feature extractor; object detection

I. INTRODUCTION

Computer systems are important for autonomous mobile robots to perform crucial tasks using sensors to extract useful information from the environment [1]. Vision-based sensors are powerful tools, providing the capability to interact with the environment [2], [3]. Cameras are used as a sensing equipment for environment perception in artificial vision systems, extracting relevant features from the image, allowing autonomous navigation, 3D mapping, visual robot localization, and object recognition [4], [5], [6].

Visual perception describes the environment extracting key-points around the system. The extracted key-points can achieve full autonomy to a mobile robot [6], [7], [8].

The robot localization problem implies the absolute current position estimation of the system, but also includes the relative position with respect to the object [2], [9], [10]. Localization is mandatory for a successful indoor navigation, and situates the system and the goal on a map.

Many visual localization methods are based on deep learning extractors. In [11] Perez, Caballero and Merino proposed an extending Monte Carlo localization method, extracting information from the images to generate a visual recognition using a map-based localization approach. Fu *et al* [12] applied a 3D mapping with an RGB-D camera, using an improved ORB algorithm as a feature extractor and measure the similarity of the descriptors with Hamming distance. Chen and Cheng proposed a cloud-based visual SLAM in [13], using

the oriented Fast and ORB features to build the dense map and estimate the position. Chiang, Hsia and Hsu [14] proposed a stereo vision self-localization system applying a trigonometric function finding coarse distances from the robot to the landmark; therefore it is applied to a neural network to increase the precision of the distance measurement. Xu, Chou and Dong [15] presented a multi sensor based indoor algorithm for a localization system, integrating a convolutional neural network (CNN) based on image retrieval, to determine the location of the robot. Another example is presented by Xie, Wang, Li and Tang [16] use a CNN robot localization to recognize the scene, reducing the error of relocating the robot.

Object detection is necessary to predict the position of the obstacles to avoid them, and move through the environment using extracted landmarks from the images to detect objects and predict the next state of the system. Many approaches for object detection use image feature extractors based on edge features, color features and textures features. Edge features methods aim on the identification of key-points from brightness changes on images. Color features detectors compare nearest regions from key-points. Texture features used regions of interest to classify pixel intensity levels in a neighborhood, such copy-move detection algorithms that search similar areas in images [17], [18], [19], [20], [21]. Ahn and Chung in [22] combined the advantages from the multi-scale Harris corner and SIFT descriptors to make an object detection using depth range features for metric information. Zheng, Barahimi, Aoun and Amar [23] present a boosted convolutional neural network for object recognition, using a boosted blocks in a succession of convolutional layers. Pertusa, Gallego and Bernabeu [24] propose an application for smartphones that allows object recognition using CNN.

Robot navigation use images to determine the system position using key-points as representations of the environment and detect obstacles [4], [5], [6], [11], [25]. Indoor navigation incorporates some knowledge of the environment, in particular a space representation, such as a map. The environment map representation is divided into three groups: Map-based navigation, the system depends on a created model or topological map. Map building-based navigation, a topological representation is contracted while the robot navigates on the environment. Mapless, the system does not have a spatial representation of the environment [26].

Visual Navigation describes the environment using landmarks on the image to build the map, locate the robot and detect obstacles. Sineglazov and Ischenko [27] used a SURF and RANSAC algorithms to locate landmarks and a neural

network to determine the object coordinates. Yilmaz and Gupta focus on combine signals from visual and inertial sensors in [28] for an indoor environment, this approach is based on visual odometry and a building information model. Shao *et al* [29] fusion the visual and inertial navigation system and applied a Kalman Filter to decline the errors of visual navigation. Sineglazov and Ischenko [30] analyze the characteristic features of different methods to develop n algorithmic maintenance for intellectual visual navigation based on neural networks.

The present approach addresses indoor map-based localization and navigation problems. A deep convolutional neural network is used to locate the system on the map and the pose estimation is based on a background modeling for feature extraction. A copy-move detection algorithm is applied to detect obstacles and a neural network for its evasion. Furthermore, a U-Net model is used for path trajectory tracking. This paper is divided in: Section II presents the proposed method for visual localization and navigation; Section III consists on the experiment results and analysis; Section IV contains the conclusions of the proposed algorithm.

II. PROPOSED METHOD

The present paper provides a solution for indoor navigation problem, including a map-based robot localization on the environment, and planning the trajectory using visual information from the mobile robot. Fig. 1 presents a general diagram from the proposed algorithm.

A. Robot Localization

Robot visual localization interprets the image information to estimate the current position of the system. Artificial neural networks are computational models capable of adapting their behavior in response to the stimulus actions from the environment [31]. The proposed visual localization stage is based on a deep neural network (Fig. 2) as a feature extractor. The convolutional deep neural network works as a feature extractor and classifies the image to locate the system in the environment; afterward it measures the robot position using the extracted interest points. The architecture of the deep neural network consists on three blocks, the first block has two CNN blocks with 64 filters, a (3,3) kernel size, a stride (1,1), a relu activation function, and an average pooling; the second block has two CNN blocks with 128 filters, a (5,5) kernel size, a stride (2,2), a relu activation function, and an average pooling; the first block has two CNN blocks with 256 filters, a (5,5) kernel size, a stride (2,2), a relu activation function, and an average pooling, therefore is used a Dense block after the Flatten stage, and classify eight classes. An Adam optimizer and a 0.0005 learning rate are used to get more accurate results.

Illumination changes influence the image scene on indoor visual localization. The neural network training process uses an image processing (gamma transformation, logarithmic transformation, blurring, rotation and translation) to reduce the error on the classification and increase the accuracy of each area of the environment, such as, rooms or corridors. The localization stage is completed by estimating the position of the robot.

Subsequently, the background model is used as a reference image. The diagram of the background subtraction is in Fig. 3, the main idea is to extract points from the image to generate a background model and analyze the image frames using an adaptive threshold to detect object when the system is static. On the other hand, a copy move detection is used to detect objects while the system is on movement [32]. One of the most frequent image forgery attacks is duplicating a fragment and place it on an image. Copy-move detection is a common technique for image forensic for detection of the tamper zones with a duplicated region from the digital image and verifies its originality [33], [34], [35], [36], [37], [38].

Key-points on the background model are used for the identification of possible objects.

$$B = \frac{\sum I_r}{N} \tag{1}$$

B is background model, I is the r^{th} frame, and N is the number of frames. A subtraction between the current frame and the reference image is applied to estimate the position of the system. This process identifies pixel intensity variations to detect possible objects.

$$S = |f_s - B_s| \tag{2}$$

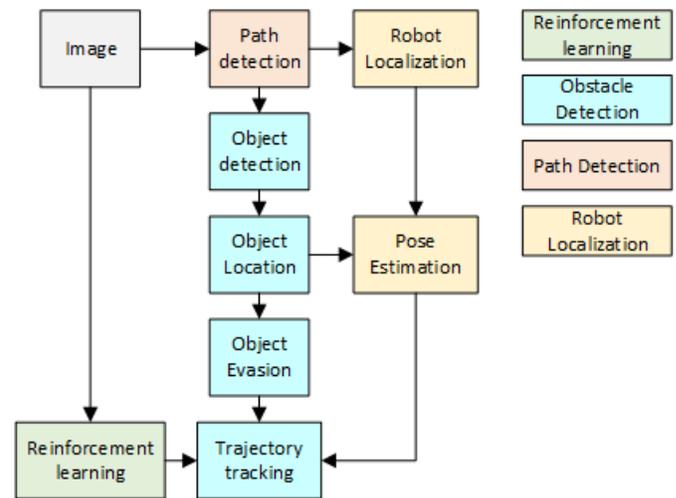


Fig. 1. Overview of the Proposed Method.

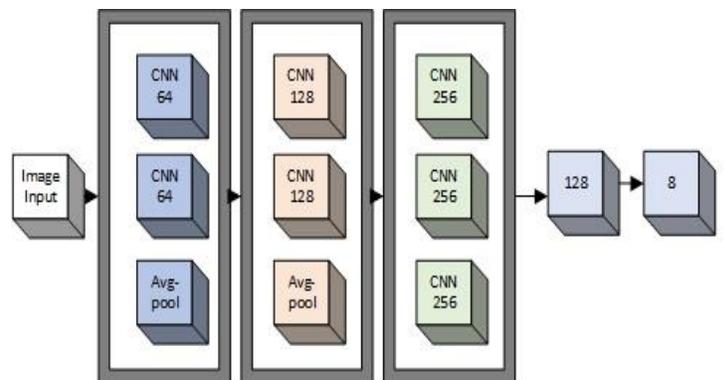


Fig. 2. Localization and Object Classification Deep Neural Network.

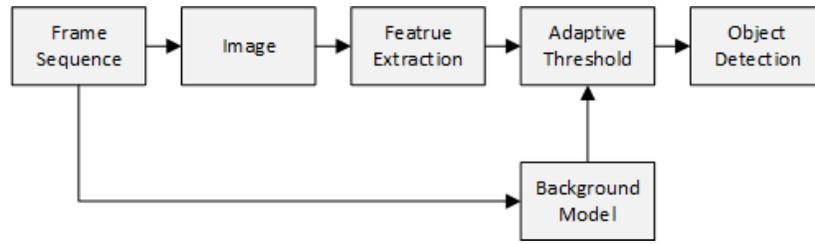


Fig. 3. Motion Detection Diagram.

where f_s is the current frame and B_s is the background. Therefore, it is extracted stable and invariant image features to get the position of the robot and obstacle detection, using an adaptive threshold based on copy move detection, such as in [35], [39]. The reference image S from the current image sequence and the last current image sequence are segmented the Discrete Cosine Transform is apply (DCT) [40].

$$C_{i,j}^K = \begin{pmatrix} C_{11}^k & \dots & C_{18}^k \\ \vdots & \ddots & \vdots \\ C_{81}^k & \dots & C_{88}^k \end{pmatrix} \quad (3)$$

The first stage of thresholding is obtained from C:

$$\lambda_1 = \sigma^2(\sqrt{\log(N)}) \quad (4)$$

$$\lambda_2 = \frac{\sum c_r}{r} \quad (5)$$

$$\lambda_3 = (\lambda_2 + \sigma^2) \sigma^2(\sqrt{\log(N)}) \quad (6)$$

where σ^2 is the variance, N are the number of frames, c_r are the values of each block C .

The Discrete Wavelet Transform (DWT) is applied to each block B and divides the image into signal components as level filters, approximations (low frequency), and details (high frequency) using the approximation sub-band as a noise filter and obtain characteristic features from the image.

$$D_{i,j}^K = \begin{pmatrix} d_{11}^k & \dots & d_{18}^k \\ \vdots & \ddots & \vdots \\ d_{81}^k & \dots & d_{88}^k \end{pmatrix} \quad (7)$$

The feature extraction allows the detection and recognition of the object, through the correlation between the images.

$$EV^k = (C^k \cdot D^k) \quad (8)$$

$$ev = (ev_{k2}, ev_{k3}, ev_{k4}, ev_{k5}, ev_{k6}, ev_{k7}, ev_{k8}) \quad (9)$$

The descriptors are computed using the block statics.

$$mean = \frac{\sum ev_1 - ev_2}{8} \quad (10)$$

ev_1 is the feature vector extracted from the current reference image and ev_2 is the feature vector extracted from the last reference image. The second stage from the adaptive threshold avoids the error in object detection and its localization. Each matrix D_{ij}^k of the current image are a shift to the right the column, shift down rows as in [39] and the descriptors are computed generating $\beta = \kappa = \varsigma = ME$.

$$T_1 = \frac{\beta + \kappa + \varsigma}{3} \quad (11)$$

Therefore, the thresholds are generated as in [39], then is proceed to compare the descriptors and obtain the interest points.

$$a = |mean| < \lambda_2 \text{ and } |mean| < T_1$$

$$b = \lambda_2 < T_1 \text{ and } mean < \lambda_3 \quad (12)$$

$$obj = \begin{cases} 0, a \text{ or } b \\ 1, other \end{cases} \quad (13)$$

The detection and localization of the object is owed to the pixel values changes on the image sequence. Therefore, a neural control network is applied for object classification and evasion. The neural control network allows the recognition of the detected objects, and the control tasks to evade the objects.

Once the interest points are detected, the robot position is estimated. The camera position is on the center and bottom of the image O as in Fig. 4.

The distance between the camera h and the object is given in centimeters:

$$h = mean(h_s, h_i) \quad (14)$$

where h_s is acquired by ultrasonic sensor, h_i is the calculated distance with the camera.

$$h_s = f_s - f_f \quad (15)$$

$$h_i = \sqrt{(\sin(b) * (\overline{OQ}))^2} \quad (16)$$

$$b = \frac{P * Q}{|P| * |Q|} \quad (17)$$

$$\overline{OQ} = \sqrt{(O(x) - Q(x))^2 + (O(y) - Q(y))^2} \quad (18)$$

f_s is the start time from the pulse and f_f is the final time from the pulse. The width from the detected obstacle is:

$$w = (\overline{OQ} * h) / 100 \quad (19)$$

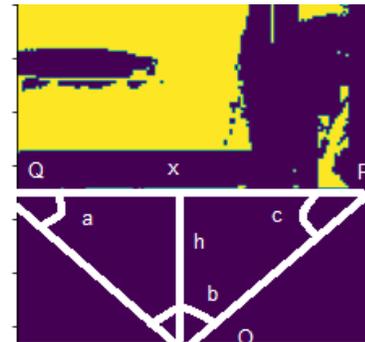


Fig. 4. Obstacle Detection and Pose Estimation.

where:

$$\overline{PQ} = \sqrt{(Q(x) - P(x))^2 + (Q(y) - P(y))^2} \quad (20)$$

B. Robot Navigation

The neural control network classifies the detected objects and the control tasks to evade them. A path planning is implemented a control system using the neural network approach described in using predefined controls to evaded the recognized object; considering a system with multiple inputs and multiple outputs.

$$v = F(\dot{s}, s) \quad (21)$$

The system considered involve both non-moving and moving rigid objects and define the robot position (p), rotation matrix (q), velocity (v), angular velocity (w).

$$p = [p_x, p_y]^T \quad (22)$$

$$q = [q_\theta, q_x, q_y]^T \quad (23)$$

$$v = [v_x, v_y]^T \quad (24)$$

$$w = [w_x, w_y]^T \quad (25)$$

Each control v defines a force-torque pair (F, τ) acting on the mass center B , therefore the motion equations can be defined as:

$$\dot{x}(t) = \begin{bmatrix} \dot{p} & v(t) \\ \dot{q} & 1/2((w(t) - q(t))) \\ \dot{v} & F/M \\ \dot{w} & R(t)I^{-1}R(t)^T \end{bmatrix} \quad (26)$$

The system used a set of predefined controls U evaluating the current system state.

$$U = \begin{bmatrix} U_1 & 0 & 1 \\ U_2 & 1 & 1 \\ U_3 & 0 & 0 \\ \vdots & \vdots & \vdots \\ U_4 & 1 & 0 \end{bmatrix} \quad (27)$$

Evaluating the states choose the control to evade the object and track the trajectory. The trajectory is predefined using reinforcement learning, using a cost function, based on the Euclidean distance, to assign priorities for the movements.

Then the trajectory is tracked by the cloning behavioral network. The system learns the trajectory and reproduces the user's control using the knowledge of a human pilot for training a convolutional neural network (CNN) that receives as input the current camera frame and outputs the next state. The idea is to reinforce the knowledge of the trajectory using a UNET model [41], [42].

A successful path tracking depends on the path and specific object recognition. To perform path recognition in an image is required a semantic segmentation. Path detection describes the systems movements in the environment [43]. The predicted class $P(class)$ represents the coincidence between the input image $O_{detected}$ and the predicted class $O_{predicted}$.

$$P(class) = \frac{A \cap B}{A \cup B} \quad (28)$$

where:

$$A \cap B = O_{detected} * O_{predicted}$$

$$A \cup B = O_{detected} + O_{predicted}$$

The existing probability of a specific class and the object detected (in this case the trajectory) are taken in consideration. During the training, the following loss function is used the categorical cross-entropy:

$$L(y, \bar{y}) = \sum \sum (y_{ij} * \log(\bar{y}_{ij})) \quad (29)$$

y is the training image and \bar{y}_{ij} is the label. The use of the categorical cross-entropy will compare the distribution of the predictions, where the probability of the true class is 1 and 0 for other class.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section the results are presented for the visual robot localization, object detection and trajectory tracking using a mobile robot with a Raspberry Pi 4 and a Raspberry V2 module (Fig. 5).

The proposed neural networks are trained under Windows 10 operating system using TensorFlow and Pytorch deep learning frameworks, the CPU model is 8GB Intel® Core i7-6700 @ 3.40 GHz, with a graphics card model NVIDIA GeForce GTX 960.

A. Robot Localization

The algorithm for robot localization uses extracted features from images as landmarks to determine where the robot is, by comparing this feature with the environment image.

The class coincidence in Table I has a good performance an 88% of accuracy, obtaining a high efficiency for robot localization. To increase the performance of the localization neural network more images for each class have to be added to the dataset.

The train dataset uses 1532 images from Room1, 1665 images from Room2, 1869 from Hall and 2052 from Room3, the validation dataset uses 1002 images from Room1, 1105 images from Room2, 1157 from Hall and 2500 from Room3, the test dataset uses 600 images, in addition, a pre-processing image process was applied to increase the classification accuracy (translation, rotation, and blurring). The neural network for localization presents in Fig. 6. An accuracy of 95%, reducing the classification error, then the position of the system is estimated by extracting key points from the background model and the object detection.

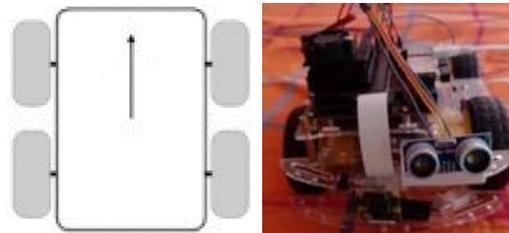


Fig. 5. Robot.

TABLE I. LOCALIZATION

Localization	Room 1	Room 2	Hall	Room 3
Coincidence	90 %	91 %	93 %	85 %

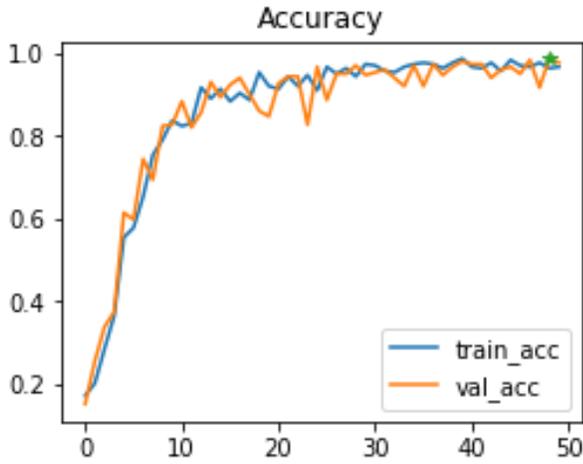


Fig. 6. Neural Network for Localization Accuracy.

The proposed approach tracks all the detected key points on the frame sequence, obtaining the robot position and locates the obstacle as in Fig. 7. The main advantage is the detection and points matching despite the illumination changes or the system position. In order to evaluate the proposed algorithm is used the accuracy and feature ratio sequence rs .

$$rs = \frac{np}{np_i} \quad (30)$$

where np are the detected points on the current image and np_i are the detected points on the previous image frame. In order to analyze the effectiveness of the object detection algorithm is compare in Table II with popular algorithms.

Clearly, the proposed method has a good performance tracking 97% of the detected features. In Fig. 8 is observed the behavior of the precision in a sequence, the proposed method has a 90%. The extracted features can be tracked without losing them on the frame sequence; the ratio sequence has a variation because the robot position changes.

The experiment was realizes using the robot on different conditions of illumination on the indoor environment. An advantage of the proposed algorithm is the feature detection and tracking on frame sequences, this advantage is useful for the detection and tracking of moving objects. In Fig. 8, shows that the extracted features can be tracked without losing information regardless of changes on the system position and illumination changes, being an advantage over the conventional methods. Once the obstacle is located a neural network classifies the object for its evasion. The effectiveness of the object classification is evaluated using the confusion matrix in Table III.

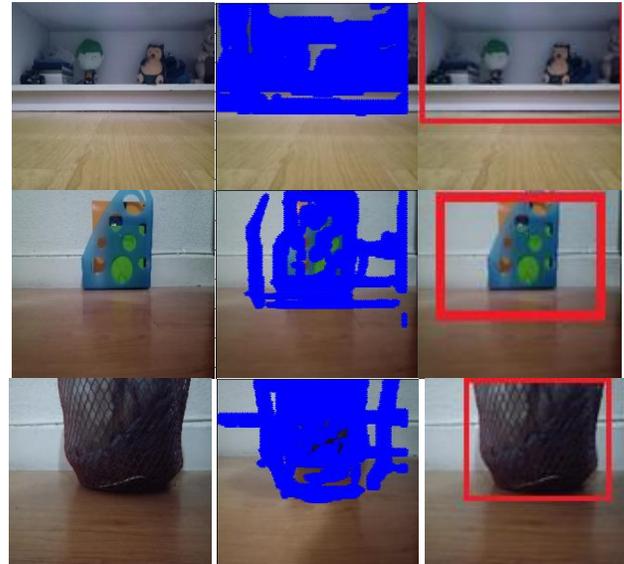


Fig. 7. Object Detection.

TABLE II. COMPARISON OF FEATURES EXTRACTORS

Method	Kaze [44], [45]	Harris [40], [46]	Eigen-Features [47]	Proposed Method
Accuracy	0.8322	0.6821	0.5051	0.9035
Ratio sequence	0.7418	0.544	0.3608	0.8011

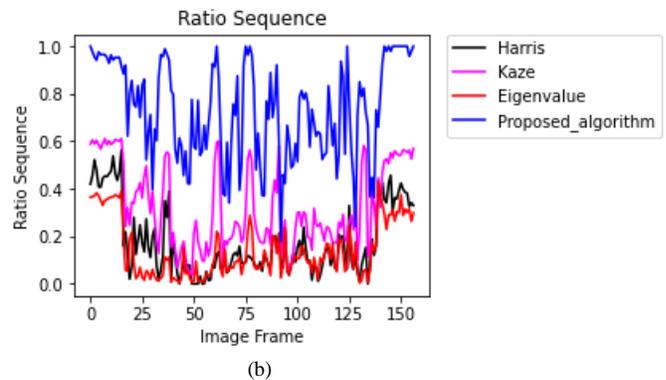
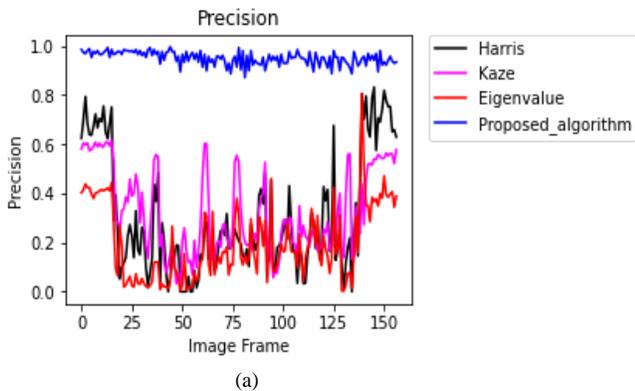


Fig. 8. Object Detection, (a) Precision, (b) Ratio Sequence.

TABLE III. CONFUSION MATRIX

	False	True	Total
Negative	93	7	100
Positive	6	474	480
total	99	481	580

The accuracy and the recall from the control neural network are described using statics values from the confusion matrix. Where T_p are the true positives, F_p are the false positives and F_n the false negatives.

$$precision = \frac{T_p}{T_p+F_p} \quad (31)$$

$$recall = \frac{T_p}{T_p+F_n} \quad (32)$$

The efficiency of the object classification neural network has high accuracy. Clearly, in Table IV, the proposed algorithm has a good performance with a 93% of accuracy and a 98% recall for object classification.

The trajectory is generated using reinforcement learning, and it is tracked by the U-Net model. In order to verify the accuracy performance of the U-Net model trained for the trajectory detection, the used data was acquired using the proposed system. The dataset contains real 1500 images with the path, in addition, the pre-processing image was applied to increase the accuracy (translation, rotation, and blurring). Furthermore, a segmentation on the image detects the path to decide the control actions to track a trajectory.

The accuracy from the U-Net model in Fig. 9 is 90% for path detection. The semantic segmentation recognizes the predicted class on the image, moreover is used to track the trajectory and evade a collision. The U-Net was trained using 1500 images and labels on different conditions on the environment to get a better performance.

To verify the applicability and accuracy of the algorithm in Fig. 10 is presented the image segmentation. The red region recognizes the path.

TABLE IV. STATISTICS VALUES

Precision	93 %
Recall	98 %

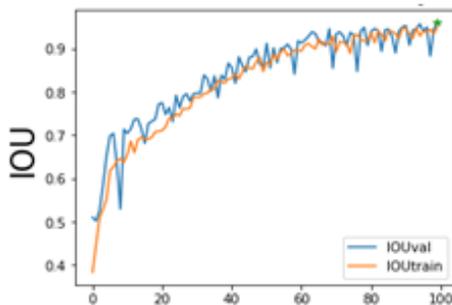


Fig. 9. U-Net Model Accuracy.

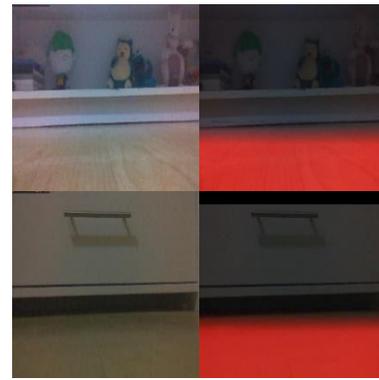


Fig. 10. Trajectory Tracking.

IV. CONCLUSIONS

This paper proposes an indoor navigation method, applying a self-localization solution and novel object detection. This approach has a high precision for feature tracking, this increases object detection and evasion.

A map-based is used in the present approach to self-locate the system on the map using a neural network. The accuracy for the localization neural network is >90%. In addition, the use of a deep learning neural network reduces the error in the system localization.

Furthermore, object detection is used to estimate the position of the system. The object detection has a >90% accuracy owing to the feature extractor and an adaptive threshold solving problems with illumination changes. The proposed solution tracks the object in the image frames; therefore, the system classifies the obstacle and reinforcement learning to evade it.

The path detection and track a trajectory a U-Net model is applied, providing the detection of the path, also the image semantic segmentation reduces the possibility of a collision. Semantic segmentation is an accurate process for the identification of the path.

To get an efficient object detection the system has to be on a dynamic environment for the detection of the objects, such as, an object has to move or the system is moving through the goal. The system navigates on unknown environments, but has to be trained to get the localization of system, because is used a supervised learning, also for the path detection is necessary to train the system with on the environment.

The proposed algorithm provides a novel solution for the navigation problem. In addition, the detection of similar zones on the frames increases the accuracy and precision for the object detection stage, also the detected features are used as landmarks to estimate the position of the system. This algorithm can be used to build a map representation of the environment. Landmarks can be set on the map to train the system with a reinforcement learning to set the locations.

The proposed algorithm has many applications, such as, car detection and pedestrian detection for autonomous driving, a surveillance robot, face tracking for recognition, car tracking for speed infractions, explorer robot, a background subtraction for a surveillance system.

ACKNOWLEDGMENT

This article was supported by the Investigation and Postgraduate Secretary of the Instituto Politécnico Nacional. Rodrigo Eduardo Arevalo-Ancona is a scholarship holder of the CONACYT.

REFERENCES

- [1] J. Rantanen, M. Mäkelä, L. Ruotsalainen and M. Jaakkola, "Motion Context Adaptive Fusion of Inertial and Visual Pedestrian Navigation," in International Conference on Indoor Positioning and Indoor Navigation (IPIN), Nantes, France, 2018.
- [2] S. Tarao, Y. Fujiwara, N. Tsuda and S. Takata, "Development of Autonomous Mobile Robot Platform Equipped with a Drive Unit Consisting of Low-End In-Wheel Motors," in 5th International Conference on Control and Robotics Engineering (ICCRE), Osaka, Japan, 2020.
- [3] N. Alfiany, G. Jati, N. Hamid, R. A. Ramadhani, M. W. D. Santika and W. Jatmiko, "Kinematics and Simulation Model of Autonomous Indonesian "Becak" Robot," in 2020 IEEE Region 10 Symposium (TENSymp), Dhaka, Bangladesh, 2020.
- [4] M. Pan, Y. Liu, J. Cao, Y. Li, C. Li and C.-H. Chen, "Visual Recognition Based on Deep Learning for Navigation Mark Classification," IEEE Access, vol. 8, pp. 32767 - 32775, 2020.
- [5] R. Druon, Y. Yoshiyasu, A. Kanezaki and A. Watt, "Visual Object Search by Learning Spatial Context," IEEE Robotics and Automation Letters, vol. 5, no. 2, pp. 1279 - 1286, 2020.
- [6] T. Fäulhammer, R. Ambruş, C. Burbridge, M. Zillich, J. Folkesson, N. Hawes, P. Jensfelt and M. Vincze, "Autonomous Learning of Object Models on a Mobile Robot," IEEE Robotics and Automation Letters, vol. 2, no. 1, pp. 26 - 33, 2017.
- [7] C. Häne, Heng, Lee, Fraundorfer, Frugale, Sattler and Pollefeys, "3D Visual perception for self driving cars using a multicamera system: Calibration, mapping, localization and obstacle detection," Image and Vision Computing, pp. 1-35, 2017.
- [8] Q. Hu, S. Paisitkriangkrai, C. Shen, A. v. d. Hengel and F. Porikli, "Fast Detection of Multiple Objects in Traffic Scenes With a Common Detection Framework," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 4, pp. 1002 - 1014, 2015.
- [9] Y. Yasuda, L. E. Martins and F. Cappabianco, "Autonomous Visual Navigation for Mobile Robots," ACM Computing Surveys, vol. 53, no. 1, pp. 1-34, 2020.
- [10] S. Huang and G. Dissanayake, "Robot Localization: An Introduction," in Wiley Encyclopedia of Electrical and Electronics Engineering, Wiley, 2016.
- [11] J. Perez, F. Caballero and L. Merino, "Enhanced Monte Carlo Localization with Visual Place Recognition for Robust Robot Localization," Journal of Intelligent and Robotic Systems, vol. 80, no. 3-4, pp. 641-656, 2015.
- [12] Q. Fu, H. Yu, L. Lai, J. Weng, X. Peng, W. Sun and M. Sun, "A Robust RGB-D SLAM System With Points and Lines for Low Texture Indoor Environments," IEEE Sensors Journal, vol. 19, pp. 9908-9921, 2019.
- [13] Y. Zheng, S. Chen and H. Cheng, "Real-Time Cloud Visual Simultaneous Localization and Mapping for Indoor Service Robots," IEEE Access, pp. 16816-16829, 2020.
- [14] J. Chiang, C. Hsia and H. Hsu, "A Stereo Vision Based Self Localization System," IEEE Sensors Journal, vol. 13, no. 5, 2013.
- [15] S. Xu, W. Chou and H. Dong, "A Robust Indoor Localization System Integrating Visual Localization Aided by CNN Based Image Retrieval with Monte Carlo Localization," MPDI Senosrs, no. 249, pp. 1-19, 2019.
- [16] T. Xie, K. Wang, R. Li and X. Tang, "Visual Robot Relocalization Based on Multi-Task CNN and Image-Similarity Strategy," MDPI Sensors, vol. 20, pp. 2-20, 2020.
- [17] M. Ikhlal, M. Hariadi and I. K. E. Pumama, "Modified Multi-scale Feature Extraction for Copy-Move Forgery Detection Based on CMD-SIFT," in 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM), Surabaya, Indonesia, 2019.
- [18] A. Mumcu and I. Savran, "Copy move forgery detection with using FAST key points and SIFT description vectors," in 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2018.
- [19] M. M. Al-Hammadi and S. Emmanuel, "Improving SURF Based Copy-Move Forgery Detection Using Super Resolution," in 2016 IEEE International Symposium on Multimedia (ISM), San Jose, CA, USA, 2017.
- [20] W. C. N. Kaura and S. Dhavale, "Analysis of SIFT and SURF features for copy-move image forgery detection," in International Conference on Innovations in Information, Embedded and Communication Systems (ICIECS), Coimbatore, India, 2017.
- [21] Y. Aydin, G. Muzaffer and G. Uluş, "Detection of copy move forgery technique based on SIFT and SURF," in 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2018.
- [22] S. Ahn and W. Chung, "Efficient SLAM algorithm with hybrid visual map in an indoor environment," in 2007 International Conference on Control, Automation and Systems, Seoul, Korea, 2007.
- [23] S. Barahimi, N. Aoun and C. Amar, "Boosted Convolutional Neural Network for object recognition at large scale," Neurocomputing, vol. 330, pp. 337-354, 2019.
- [24] A. Pertusa, A. Gallego and M. Bernabeu, "MirBot: A collaborative object recognition system for smartphones using convolutional neural networks," Neurocomputing, vol. 293, pp. 87-99, 2018.
- [25] M. Rysuke, M. Adachi, H. Ishida, T. Watanabe, K. Matsutani, H. Komatsuzaki, S. Sakata, R. Yokota and S. Kobayashi, "Visual Navigation Based on Semantic Segmentation Using Only a Monocular Camera as an External Senso," Journal of Robotics and Mechatronics, vol. 32, no. 6, pp. 1137-1155, 2020.
- [26] D. Souza and Kak, "Vision for Mobile Robot Navigation: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 2, pp. 237-268, 2002.
- [27] V. M. Sineglazov and V. S. Ischenko, "Intelligent Visual Navigation System of High Accuracy," in IEEE 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), Kiev, Ukraine, 2019.
- [28] A. Yilmaz and A. Gupta, "Indoor positioning using visual and inertial sensors," in 2016 IEEE SENSORS, 2016.
- [29] W. Shao, L. Dou, H. Zhao, B. Wang, H. Xi and W. Yao, "A Visual/Inertial Relative Navigation Method for UAV Formation," in 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020.
- [30] V. Sineglazov and V. Ischenko, "Intelligent system for visual navigation," in 4th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC), Kiev, Ukraine, 2016.
- [31] A. Alanis, N. Arana, C. Lopez and E. Guevara Reyes, "Integration of an inverse Optimal Neural Controller with Reinforced SLAM for Path Planning and Mapping in Dynamic Environments," Computación y Sistemas, vol. 19, no. 3, pp. 445-456, 2015.
- [32] Kermani and Asemani, "A Robust Adaptive Algorithm of Moving Object Detection for Video Surveillance," Journal on Image and Video Processing, vol. 1, no. 127, pp. 2-9, 2014.
- [33] Mushtaq and Mir, "Copy-Move Detection Using Gray Level Run Length Matrix Features," Optical and Wireless Technologies, vol. 472, pp. 411-420, 2018.
- [34] Kuznetsov and Myasnikov, "A Copy-Move Detection Algorithm Using Binary Gradient Contours," International Conference on Image Analysis and Recognition, vol. 9730, pp. 349-357, 2016.
- [35] Ghorbani, Firouzmand and Faraahi, "DWT-DCT (QCD) Based Copy-move Image Forgery," in 2011 18th International Conference on Systems, Signals and Image Processing, Sarajevo, Bosnia and Herzegovina, 2011.
- [36] J. N. Yuan Rao, "A Deep Learning Approach to Detection of Splicing and Copy move Forgeries in Images," IEEE International Workshop on Information Forensics and Security, 2016.
- [37] Y. Wu, W. Abd-Almageed and P. Natarajan, "Deep Matching and Validation Network An end to end to constrained Image Splicing Localization and Detection," IEEE Computer Vision and Pattern Recognition, pp. 1480-1489, 2017.

- [38] S. L. Ashwini Malviya, "Pixel based Image Forensic Technique for copy-move forgery," 7th International Conference on Communication Computing and Virtualization 2016, vol. 79, pp. 383-390, 2016.
- [39] X. Wang, X. Zhang, Z. Li and S. Wang, "A DWT-DCT Based Passive Forensics Method for Copy-move Attacks," in Third Conference on Multimedia Information Networking and Security, 2011.
- [40] Harris and Pike, "3D Positional integration from image sequences," Image and Vision Computing, vol. 6, no. 2, pp. 97-90, 1988.
- [41] O. Ronneberger, P. Fischer and T. Brox, "U-Net Convolutional Networks for Biomedical Image Segmentation," Medical Image Computing and Computer Assisted, pp. 234 - 241, 2015.
- [42] X. Hu, M. Naiel, A. wong, M.-. Lamm and P. Fieguth, "RUNET: A Robust UNET Architecture for Image Super Resolution," in IEEE Conference on Computer Vision, 2019.
- [43] S. LaValle, Planning Algorithms, Cambridge University Press, 2006.
- [44] P. F. Alcantarilla, "KAZE features," in European Conference on Computer Vision, Berlin, Berlin, Germany, 2012.
- [45] M. Pourfard, T. Hosseinian, R. Saeidi, S. A. Motamedi, M. J. Abdollahifard, R. Mansoori and R. Safabakhsh, "KAZE-SAR: SAR Image Registration Using KAZE Detector and Modified SURF Descriptor for Tackling Speckle Noise," IEEE Transactions on Geoscience and Remote Sensing, vol. press, pp. 1-12, 2021.
- [46] B. A. Jasani, S.-K. Lam, P. K. Meher and M. Wu, "Threshold-Guided Design and Optimization for Harris Corner Detector Architecture," IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 12, pp. 3516 - 3526, 2018.
- [47] J. Shi and C. Tomasi, "Good Features to Track," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 1994.

Firm Performance Prediction for Macroeconomic Diffusion Index using Machine Learning

Cu Nguyen Giap¹, Dinh Thi Ha³
Vu Quang Huy⁴, Do Thi Thu Hien⁵

Faculty of Economic Information System and e-Commerce
Thuongmai University
Hanoi, Vietnam

Dao The Son², Le Mai Trang⁶

Department of Economics
Thuongmai University
Hanoi, Vietnam

Abstract—Utilizing firm performance in the prediction of macroeconomic conditions is an interesting research trend with increasing momentum that supports to build nowcasting and early warning systems for macroeconomic management. Firm-level data is normally high volume, with which the traditional statistics-based prediction models are inefficient. This study, therefore, attempts to assess achievements of Machine Learning on firm performance prediction and proposes an emerging idea of applying it for macroeconomic prediction. Inspired by “micro-meso-macro” framework, this study compares different machine learning algorithms on each Vietnamese firm group categorized by the Vietnamese Industry Classification Standard. This approach figures out the most suitable classifier for each group that has specific characteristics itself. Then, selected classifiers are used to predict firms’ performance in the short term, where data was collected in wide range enterprise surveys conducted by the General Statistics Office of Vietnam. Experiments showed that Random Forest and J48 outperformed other ML algorithms. The prediction result presents the fluctuation of firms’ performance across industries, and it supports to build a diffusion index that is a potential early warning indicator for macroeconomic management.

Keywords—Firm performance prediction; machine learning algorithms; diffusion index

I. INTRODUCTION

Macroeconomic situation is always an important factor for all economic sectors, and it is trivial that macroeconomic forecasting is very important [1]. From the necessity of macroeconomic forecasting, there have been many studies on this issue, such as predicting GDP [2], forecasting inflation [3], unemployment [4], exchange rates [5]. These indicators are also predicted in different aspects, such as forecasting growth level [6], or degree of fluctuation [7].

Faced with new practical demands, due to the re-emergence of economic crises, economists are more interested in alerting abnormal situations instead of just giving out the predicted value of indicators. And along with the improvement in the availability of input data of prediction problem, new achievements in data processing methods, and computing power, a new group of studies with the concept of “nowcasting” and early warning system in macroeconomic arises. Aastveit, Gerdrup et al. used big data and machine learning techniques to forecast real-time GDP [8]. The large data was used to report and forecast macroeconomic situation [9], and Galbraith and Tkacz applied the GDP reporting

method with electronic payment data [10]. On the other side, Reinhart et al. [11] studied the development of early warning models of financial market risk assessment for emerging markets. Ciarlone and Trebeschi (2005) studied how to design an early warning system for debt crises [12], and Sevimet al. (2014) built an early warning system to forecast currency crises [13].

In most cases, the input data used for macroeconomic prediction are aggregate economic indicators, such as consumer price index (CPI) or gross domestic product (GDP). However, firm performance is an indispensable and crucial factor that needs to be considered when predicting macroeconomic conditions. In recent decades, there is a new trend of using firm-level data in predicting macroeconomic aspects [14]–[20]. The researchers have proved the importance of using firm-level information in macroeconomic prediction from both theoretical and practical points. From the theoretical point, a framework to research the effect of micro-foundations information in macro-economic aspect was proposed in [19], [20]. In practical, Joao et al have using firm performance to conduct new micro-aggregated factor that was proved to be useful in GDP prediction [14]. Productivity in firm-level data was also used to build a micro-aggregated factor that is similar to total factor productivity (TFP) and this micro-aggregated factor is facilitated in the prediction of other macro-economic indicator models.

Using firm-level data on macroeconomic prediction brings several advantages. a) micro-aggregated series presents the dynamics of the published aggregate factors reasonably. b) micro-foundation information can identify the factors underlying the differences in the efficiency of all manufacturing. c) firm-level information is measured in high frequency based on information communication technology [14]–[16], [18]. This approach also brings a great opportunity for improving qualities of prediction models thank to data’s granularity. However, it is also posing a challenge because firm-level data is normally high-volume data. With high volume data, traditional statistic-based prediction models are inefficient. In proposals of Bartelsman et al, and Brito et al, had to build a microaggregate factors before using them in prediction models. In opposite, machine learning algorithms become the highly potential methods for macroeconomic state prediction based on firm-level information directly.

However, it can be concluded that until now, there has been no study which systematically evaluates the performance

of ML algorithms on this issue. From this point of view, this article uses ML algorithms to predict firms' performance based on firm survey data and other additional information. This study differs from other studies in the way that firm performance prediction is used to support for macroeconomic perspective, rather than for microeconomic management. For this purpose, a huge data including information of a wide range of companies must be processed, and the firms' information used is public information only. Using secret firms' information for macroeconomic analysis is inappropriate. Besides, the companies of each industry group have exclusive/specific characteristics, therefore the suitable models of each group would not be the same. Inspired by "Micro-meso-macro" framework [19], therefore, this study aims to build the different predictive classifier for the firm's performance of different industry groups. This article also aims to build the warning of the macroeconomic situation based on firms' performance classification.

Firm's performance is measured by different indicators including financial indicators and market signal indicators. This article uses the two most popular indicators for the firm's performance, which are the return on asset (ROA) and the return on equity (ROE). Predicting other firm performance indicators [21] is going to be a research question for the next study.

Macroeconomic forecasting in Vietnam has been implemented by the government for a long time. However, it was not until 1984 that Vietnam had the GDP index for the first time, which is the most basic indicator of the macroeconomy. There are some forecasts of basic macroeconomic indicators such as GDP and inflation. However, there is no research applies the same approach with this study for Vietnamese macroeconomic forecasting.

In this article, the information of Vietnamese enterprises collected by the General Statistics Office of Vietnam in the 2010-2015 period has experimented. The result shows the potential application of selected machine learning algorithm to supply important information: predictive firms' performance for the overall economy, which can indicate macroeconomic condition. A proposal micro-aggregate factor, akin to diffusion index, built from firm performance classification prediction and its potential use are shown in this research.

In the remaining part of this paper, section 2 presents the related works. In section 3, the research methodology and preliminary are introduced. Section 4 presents the experiments and evaluation of the performance of algorithms on Vietnamese firm's data sets. Finally, Section 5 presents some conclusions, discussion, and research extension in the near future.

II. RELATED WORKS

Utilization of Firm Performance Prediction (FPP) for macroeconomic perspective based on Machine learning algorithms is still an ongoing research area. There are few publishes proposed the clear application of FPP for policymakers. However, machine learning algorithms have been widely applied for firms' performance prediction for other purposes [22]–[32], and these models will be also useful

for the government's approaches on macroeconomic management.

Different novel models for FPP were proposed on each application such as bankruptcy, financial rating, financial distress, business failure, and so on. These applications consider different aspects of FPP, therefore its use different FPs indicators and independent variables. In the beginning, the researchers and managers have strongly focused on FPP methods using the accounting factors, then market-based indicators have been concerned also. Recently, new technological approaches have used sentiment-based analysis to improve the firms' performance prediction [33]–[35].

In firm performance prediction, successive machine learning algorithms including decision tree, support vector machine, artificial neural network, and its modification are well-known, and the remarkable studies are briefly summarized herein.

Decision tree included boosting the algorithm, the random forest is a well-known machine learning algorithm for FPP [36]–[41]. Delen deeply researched on applying decision tree for firm performance evaluation [36]. He studied common decision tree models including CHAID, C5.0, QUEST, and C&RT for the large and rich feature dataset and proposed an application framework for FPP using a decision tree with financial factors. Bankruptcy prediction is also well studied using decision tree model. Zibanezhad and Foroghi [37] extracted 25 Financial ratios and used these ratios as independent variables for building a Classification And Regression Tree. This work adapted very well to continuous financial ratios data and showed the applicable of C&R tree for bankruptcy prediction problem. Recently, Jardin [38] proposed a novel application of decision tree for firm financial evolution prediction in mid-term, and bankruptcy. His work made an improved approach by using closed time-series data of six financial dimensions to make his model become sensitive to short-term changes. Many other works used decision tree for firm performance issues were introduced, Basti [40] and Yeo [41] were good examples.

Support vector machine (SVM) is a supervised learning algorithm developed by Vapnik [42] and is successfully applied for classification problems. In the realm of finance applications, SVM has been applied in a various problems such as selecting bankruptcy predictors [43], stock price index predicting by financial time series data [44]–[46], forecasting bankruptcy or failure abilities [31], [47]–[50]. In those aspects, SVM approaches were compared with many other methods and this algorithm showed its potential ability. Francis [46] compared SVM with back-propagation neural network (BPNN) based on criteria such as normalized mean square error (NMSE), mean absolute error (MAE), directional symmetry (DS) and weighted directional symmetry (WDS). SVM is also assessed about the predictive performance along with some traditional approaches including the Linear Discriminant Classifier (LDA), Multi-layer Perceptron (MLP), and the Learning Vector Quantization (LVQ) [43]. Min [47] evaluated the SVMs with MDA, logistics, and BPNN or Wu [48] compared a genetic algorithm based SVMs with DA, logit, probit, and ANNs. In these research, the results

indicated that SVM performed better than other approaches with higher percentage classified correctly, or higher precision and lower error rates. However, there are some studies recently proved that ANN models to be superior to SVM approaches with smaller estimated relative error costs [50].

Many proposal ANN models has been applied to finance problems especially in bankruptcy and financial distress [24], [29], [32], [51], since 1990's. Odom and Sharda were pioneers in applying ANNs to predict the failures of firms [52]. They developed an ANN model for bankruptcy prediction and compared proposal model to multivariate discriminant analysis (MDA) by empirical tests using financial data from various companies. Tam [53] applied backpropagation neural network (BPNN) for bankruptcy prediction of banks and compared with existing models include discriminant analysis (DA), factor-logistic, K- nearest neighbor (k-NN) and decision tree ID3. The results showed that ANNs offered better predictive accuracy than other models. Coats and Fant [54] proposed an ANN model as an alternative analysis method of the same ratios used by MDA. They showed the ANN approach was more effective than MDA. Bell and T.B [55] compared logistic regression and ANNs in predicting bank failures. Their results indicated that both methods having similar predictive accuracy. Besides, there were many research combining ANN with other soft computing techniques such as fuzzy sets [56], genetic algorithm [57], rough sets [58].

To build a framework that uses machine learning algorithms to predict Vietnamese FP for macroeconomic perspective, those algorithms are tested and compared with different common machine algorithms to identify the best model for each industry group.

III. RESEARCH METHODOLOGY

A. Classification Model Construction

This study aims to predict performance at firm-level data with the purpose of supporting forecasting at macroeconomic level. For this, a huge data including information of enterprises across the economy is processed. For privacy and practical issues, only publicly available information of firms' performance is used. Following the framework introduced by Dopfer [19], this study aims to build the different predictive classifier for the firm's performance of different industry groups following Vietnam Standard Industrial Classification. The results of classifiers are used to estimate a proposed index that is a trigger in an early warning system for macroeconomic management.

The study uses a research methodology which includes two critical stages, the first stage collects and preprocesses the large raw data set, and the second stage builds the typical model for each industry group. The specific steps of this research methodology are depicted in Figure 1.

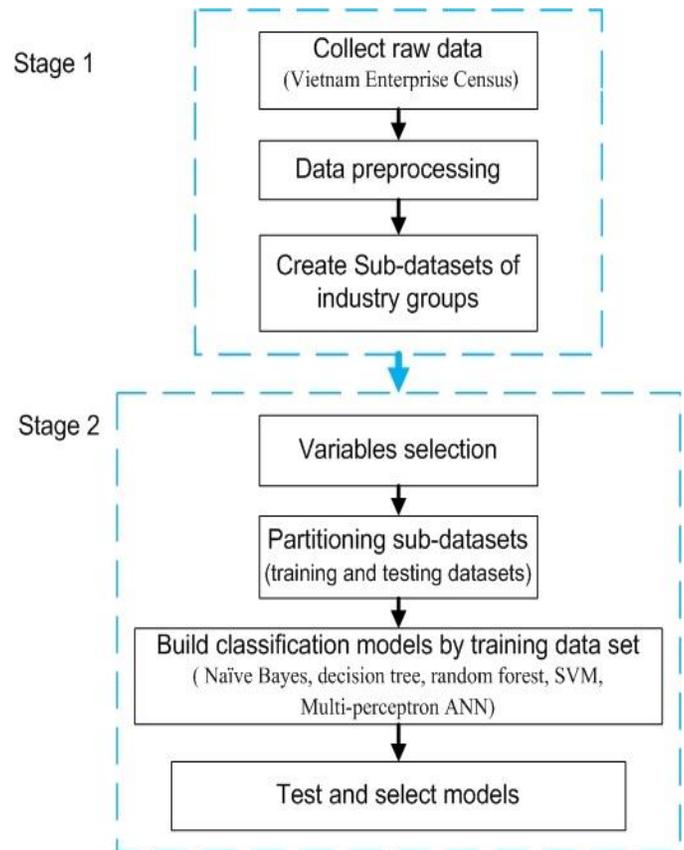


Fig. 1. Steps followed in Research Methodology.

The first stage has three steps: (1) raw data collection, (2) preprocessing, and (3) sub-setting raw data to each industry group. The output of this stage is valid datasets for stage 2 that builds suitable prediction models. In the first step, the raw dataset of firms' information is collected from GSO's annual economic census. In the second step, outliers, are removed. Missing values are either removed or replaced by the average value of variables it belongs to. Data noises are processed to provide an acceptable quality dataset to later stages. In the third step, the data set is partitioned by industry group code according to Vietnam Standard Industrial Classification, a slightly modified version of the Global Industry Classification Standard, in the third step. And then, the data sets belong to industry groups are validated before being passed on to the second stage.

In the second stage, each sub-dataset is split into training and testing datasets by a specific proportion in the first step. Then, the training dataset is used to train different ML approaches including Naïve Bayes, decision tree (J48), random forest, SVM, ANN. These models are tested on the testing dataset and the models' performances are evaluated by different measurements containing precision, recall, accuracy, ROC. The testing result is analyzed to make the insight of using ML approaches for FPP in macroeconomic perspective.

1) *Naïve bayes classifier*: Naïve Bayes is a common machine learning technique that is developed based on the Bayes's theorem and it is suited when the dimensionality of the inputs is high and assume that inputs are independent is satisfied. Naïve Bayes classifier takes input instance as a feature vector $x = \{x_1, \dots, x_n\}$ and classes dependent variable y by posterior probability $\text{Prob}(y_i|x)$ where y_i is a possible outcome of y . Naïve Bayes classifier is commonly trained by a supervised method such as Maximum-likelihood on a given training set [59]. Particularly:

Posterior probability $\text{Prob}(y_i|x)$

$$\text{prob}(y_i | x) = \frac{\text{prob}(y_i)\text{prob}(x_1, \dots, x_n \vee y_i)}{\text{prob}(x_1, \dots, x_n)}$$

In practice, $\text{prob}(y_i)$ and $\text{prob}(x_1, \dots, x_n)$ are calculated from a training set directly and $\text{prob}(x_1, \dots, x_n | y_i)$ is equivalent to the joint probability model $\text{prob}(x_1, \dots, x_n, y_i)$. Applying the chain rule and Naive independent assumption of inputs, there is:

$$\begin{aligned} \text{prob}(x_1, \dots, x_n, y_i) &= \text{prob}(x_1 | x_2, \dots, x_n, y_i)\text{prob}(x_2, \dots, x_n, y_i) \\ &= \text{prob}(x_1 | x_2, \dots, x_n, y_i)\text{prob}(x_2 | x_3, \dots, x_n, y_i)\text{prob}(x_3, \dots, x_n, y_i) \\ &= \text{prob}(x_1 | x_2, \dots, x_n, y_i)\text{prob}(x_2 | x_3, \dots, x_n, y_i) \dots \\ &\quad \text{prob}(x_{n-1}, x_n, y_i)\text{prob}(x_n, y_i)\text{prob}(y_i) \\ &= \text{prob}(x_1 | y_i)\text{prob}(x_2 | y_i) \dots \text{prob}(x_{n-1} | y_i)\text{prob}(x_n | y_i)\text{prob}(y_i) \\ &= \text{prob}(y_i) \prod_{j=1}^n \text{prob}(x_j | y_i) \end{aligned}$$

Because $\text{prob}(x_j | y_i)$ is proportional to $\text{prob}(x_1, \dots, x_n, y_i)$, henceforth, the classifier model of dependent variable y to a specific class \hat{y} is:

$$\hat{y} = \arg \max_{i \in \{1, \dots, k\}} \text{prob}(y_i) \prod_{j=1}^n \text{prob}(x_j | y_i)$$

2) *Decision tree*: Decision tree (DT) is a tree-based structure model that expresses the possible consequent states with its chance events and outputs. Decision tree is a non-parametric supervised machine learning algorithm that is used popularly for classification and regression problems. Learning a decision tree is based on partitioning the set of training examples into smaller and smaller subsets where each subset is as "pure" as possible. The purity for a particular subset is measured according to the number of training samples in that subset having the same class label. In practice, a DT structure is constructed directly from a training set by an iterative process that starts with a null root node and repeatedly split a node on an attribute-based on information gain. A DT is built by C4.5, J48, C5.0 algorithms, and they all follow the same

recursive process extends from Quinlan's earlier ID3 algorithm.

The most important task in decision tree construction process is finding the normalized information gain from splitting on an attribute. This is complete by following steps [60].

Calculate information conveyed by a distribution of the set of classes P in a current dataset, also called the Entropy of P , is:

$$I(P) = -(p_1 \log(p_1) + p_2 \log(p_2) + \dots + p_n \log(p_n))$$

Calculate the incremental information by partition a set T on a non-categorical attribute X into sets T_1, T_2, \dots, T_m

$$\text{Info}(X, T) = \sum_{i=1}^m \frac{|T_i|}{|T|} * I(T_i)$$

Calculate the quantity $\text{Gain}(X, T)$ defined as is the measure of the difference in entropy from before to after the set T is split on an attribute X $\{\displaystyle A\}$:

$$\text{Gain}(X, T) = I(T) - \text{Info}(X, T)$$

3) *Random forest*: Random forest is a resampling approach for classification and regression problems. Random forest builds a classifier by assembling individual simple classifiers trained on different sub-datasets generated by bootstrapping a training set [41]. Random forest classifier improves the quality of a classifier built on a single decision tree by solving overfit and bias problems. Random forest uses Bagging (bootstrap aggregating) algorithm, which uses multiple versions of the training set, each created by bootstrapping the original training data to train the models. Each of these bootstrap data sets is used to train different component classifiers that are simple decision trees commonly, and then a final classification decision is form by a voting process of each component classifier.

4) *Support vector machine*: Recent years, support vector machine (SVM) is a new succeeded supervised learning algorithm for classification problems including the firm's performance prediction problem [47]. SVM is a non-parametric algorithm aimed to find the optimal hyperplanes that separate classes on a training dataset. An SVM is trained by solving a large quadratic programming problem. For the proposed problem, SVM is trained by a sequential minimal optimization algorithm, in which a complex quadratic programming problem is broken into a series of smallest possible quadratic programming problems and these small quadratic programming problems are solved analytically.

For a binary classification problem with a dataset $\{(x_1, y_1), \dots, (x_n, y_n)\}$, where x_i is an input vector and y_i is a relative output label. A soft-margin support vector machine is trained by solving a quadratic programming problem:

$$\max_a \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j K(x_i, x_j) a_i a_j$$

Subject to:

$$0 \leq a_i \leq C, i = 1, 2, \dots, n$$

$$\sum_{i=1}^n y_i a_i = 0$$

Where C is a regularization of SVM model and $K(x, y)$ is a kernel function that can be linear, polynomial, or exponential kernel.

5) *Feed-Forward Artificial Neural Network (ANN)*: Feed-forward ANN is a great machine learning model that can represent a highly complex relationship between inputs and outputs. Feed-forward ANN has three layers of architecture includes input layer, hidden layers, an output layer, and the connections of nodes are represented by an acyclic directed graph. One node of ANN has multiple inputs, a weight vector w and one output with a bias, as depicted in Fig.2 [54].

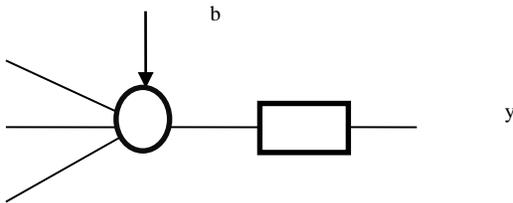


Fig. 2. A Neuron of Dynamic Structural ANN.

In this study, the ANN uses sigmoid activation function is defined below:

$$\varphi(net) = \frac{1}{1 + e^{-net}}$$

In which, for the node j

$$net_j = b + \sum_{i=1}^n x_i w_{ij}$$

In the input layer, each input node is linked with only one input and this node has connections to all hidden nodes. The hidden layer is a set of consequent layers or is one full connect layer only. In any case, a hidden node connects to all nodes in the next layer, and in the last hidden layer, a hidden node has direct connections to all nodes of the output layer. The number of nodes in the output layer is the same with the number of outputs.

Feed-forward ANN is commonly trained by Back – Propagation algorithm, in which the weights of nodes are changed with subject to minimize output error.

$$\varepsilon = \sum_{i=1}^n (t_i - y_i)^2$$

where: t_i is an instance output value in training set and y_i is its relative network output.

B. A proposal on a Diffusion Index as an early Warning System of Real Business Cycle

At this stage, our prediction model is good at forecasting firms' ROA classification. Using input data from the beginning of each year, firms are predicted how well they will perform at the end of that year and classified into each of the five categories from lowest ROA to highest ROA. Each firm then can be predicted to stay in the same class, move up, or move down. We propose to use a diffusion index-based computed indicator to describe the overall situation of firms across industries in the economy by looking at the number of firms that are performing better or worse. Diffusion indexes was formally mentioned in [61] and it has been applied regularly since then, such as for macroeconomic forecasting [62].

$$DI_{fp} = \frac{1 * N_{mh} + 0.5 * N_{ss} + 0 * N_{ml}}{N_{mh} + N_{ss} + N_{ml}} * 100$$

Where

- DI_{fp} is the diffusion index based on firm performance at an time point.
- N_{mh} is the number of firm moving to higher classes: note that we compare the end-of-year predicted value to the begin-of-year actual value.
- N_{ss} is the number of firm that stay in the same class.
- N_{ml} is the number of firm moving to lower classes.

The index can vary from 0 to 100. If the index is equal to 50, it can be consider as averagely all firms are performing in the same classes. If the index is higher than 50, more firms moving up them firms moving down and that's a signal of an improving macroeconomic performance. If the index is higher lower than 50, more firms moving down than firms moving up and that's a signal of a slowing down macroeconomic performance.

This framework is similar to the use of the Purchasing Manager Index that has been widely used as a leading indicators to forecast the economy [63]. However, the index proposed in this study has a significant advange of computing from actual firms' performance across the economy rather than from getting opinions from purchasing managers of selected enterprises.

Although the data used in this expriment study was the enterprise census implemented annually, the main purpose of showing the potential use of large data on firms' information is still valid. In practice and further investigation, the government with its ability to assess firms' information at higher frequency or sometime even real time, can apply the same framework. And given that this study tries to predict firms' performance in subset of industries will also allow the government to create diffusion indexes for each industry or sector, therefore having able to detect the macroeconomic performance with more details.

IV. EXPERIMENTAL RESULT

A. Data Description and Variables Selection

This study applies and compares different well-known machine learning algorithms on the data of Vietnamese enterprises collected by the General Statistics Office of Vietnam in the 2010-2015 period. This dataset was carried out with all firms in every type of business and every industry in Vietnam. It includes more than 500,000 companies with nearly 200 variables for each observation. These variables reflect different aspects of firms such as business tax ID, type of business, asset and capital structure, employee structure, the result of business, and others.

These are annual enterprise census conducted by the General Statistic Office of Vietnam to provide the government with information about the firms' performances. These surveys have some disadvantages for the application of machine learning algorithms because the questionnaire is designed with closed-ended questions for common statistical analysis. Despite several limitations to the performance of machine learning algorithms, it is still the best existing dataset with information of firm performance for macroeconomic perspective, and therefore suitable for this research.

As mentioned in the literature review, common theory of firm performance evaluation used many indicators, but in this emerging study, two indicators including ROE and ROA are used to fulfill the study targets. Variables in the raw dataset from the General Statistics Office of Vietnam which do not support ROE, ROA prediction are removed. On the other hand, some new variables are generated from original variables to use in prediction models. The final selected variables are described in Table 1, in which ROA, ROE are output indicators whereas other variables are input indicators. All these input parameters are values at the begin of each year and they are used for predicting firm performance indicators at the end the of year.

In addition, this study aims to assess not only the different impacts of indicators on the prediction model but also the difference between industry groups, therefore the dataset is divided into subsets corresponding with industries. This division is performed according to codes of each industry in Decision on Vietnam Standard Industrial Classification of Prime Minister, No. 27/2018/QĐ-TTg. Although there are 99 industries in Vietnam in this division, the study select test on industries that have the number of firms larger than 20,000 to guarantee learning quality of prediction models. These selected industries are described in Table 2.

TABLE I. INDICATORS OF RESEARCH MODEL

Index	Indicator	Index	Indicator
1	Type of business	16	Year-opening debt ratio [total debt/total capital]
2	Business size [1=SME, 2=big]	17	Export value
3	State-ownership status	18	Gross revenue (VND million)
4	Total assets (VND million)	19	Net revenue
5	Total equity (VND million)	20	Core-business Gross revenue
6	Total debt (VND million)	21	Ratio of core business revenue to gross revenue (%)
7	Total number of employees	22	Percentage of core-business Gross revenue
8	Total number of female employees	23	Profit before tax
9	Total number of core-business employees	24	Profit after tax
10	Number of change employees	25	Business tax
11	Percentage of female employees (%)	26	Total of earning (VND million)
12	Percentage of core-business employees (%)	27	Year performing the survey
13	Second-industries status	28	Return on Asset
14	Percentage of state shares (%)	29	Return on Equity
15	Owner's equity ratio (%)		

TABLE II. THE CODES OF SELECTED INDUSTRIES

No	Code	Description
1	01	Agriculture and related service activities
2	10	Manufacture of food products
3	14	Manufacture of wearing apparel
4	42	Construction of civil engineering structures
5	43	Specialized construction activities
6	47	Retail trade, (except motor vehicles, motorcycles
7	49	Land transport and transport via railways and via pipelines
8	55	Accommodation
9	56	Food and beverage service activities
10	68	Real estate activities
11	73	Advertising and market research

B. Experiment Result

Return on asset (ROA) and return on equity are two numerous variables in raw survey data set. However, the light changes of ROA, ROE are meaningful for managers of firms but not for the policymakers. Therefore, this study transforms ROA, ROE variables into two categorical variables by the specific interval border. The ROA is discretized by intervals $\{(, 0], (0, 5], (5, 20], (20, 30], (30,)\}$ and ROE is discretized by intervals $\{(, 2.5], (2.5, 7.5], (7.5, 15], (15, 20], (20,)\}$. These segments are encoded by class labels as 0, 1, 2, 3 and 4. Table 3 and Table 4 show the number of firms belongs to different classes of every industry group.

In general, the performances of five well-known machine learning algorithms, Naïve Bayes, decision tree (j48), random forest, SVM, MLP, for Vietnamese firms' performance prediction is shown in figure 3 and 4.

According to the proportion of correctly classified instances, as shown in figures 3 and 4, in both cases, two tree-based algorithms including J48 and random-forest have outperformed other algorithms. In the case of ROA prediction, J48 algorithm has the minimum proportion of correctly classified instances is 86.31% for No.56 industry and reach maximum proportion at 95.77% for No.42 industry, and the average proportion is 91.61%. The random forest algorithm is even better, it has the minimum proportion of correctly classified instances is 87.38% for No.56 industry and reach maximum proportion at 95.77% for No.42 industry, and the average proportion is 91.81%. Two algorithm SVM and MLP have close performance and they only work well in several industry groups including No.42, No.43, and No.49 with the proportions of correctly classified instances are higher than 80%. Naïve Bayes algorithm doesn't work in this case. It has proportions of correctly classified instances from 6.92% to 44.21%.

TABLE III. NUMBER OF FIRMS IN EACH CLASS AND INDUSTRY GROUP BY ROA CLASSIFICATION

Class\industry code	1	10	14	42	43	47	49	55	56	68	73
0	701	289	62	784	1,248	3,946	607	41	9	405	375
1	3,912	2,565	2,090	7,955	3,959	12,823	4,634	2,567	1,737	1,850	2,211
2	1,677	767	693	677	465	3,135	689	631	634	375	514
3	237	116	109	71	53	388	60	72	100	78	95
4	473	163	146	113	75	608	110	289	120	92	105
Total	7000	3900	3100	9600	5800	20900	6100	3600	2600	2800	3300

TABLE IV. NUMBER OF FIRMS IN EACH CLASS AND INDUSTRY GROUP BY ROE CLASSIFICATION

Class\Industry code	1	10	14	42	43	47	49	55	56	68	73
0	3,413	1,680	1,194	6,476	4,038	11,785	3,630	2,075	1,073	1,759	1,605
1	1,839	908	693	1,727	954	4,465	1,339	847	701	406	754
2	836	515	440	713	361	2,174	543	367	405	265	369
3	265	189	181	219	105	717	174	107	108	90	121
4	647	608	592	465	342	1,759	414	204	313	280	451
Total	7000	3900	3100	9600	5800	20900	6100	3600	2600	2800	3300

- Result of ROA prediction:

Proportion of Correctly Classified Instances

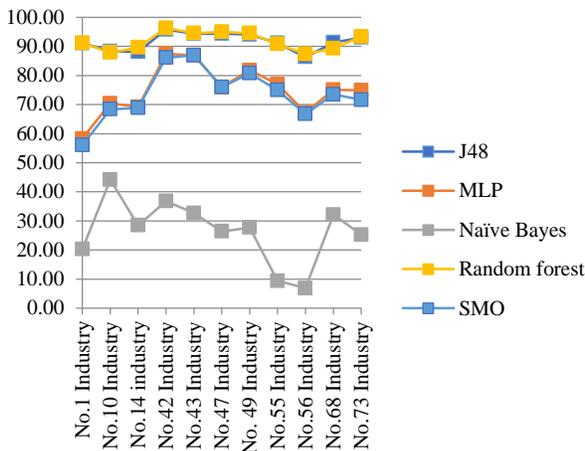


Fig. 3. The Proportion of Correctly Classified Instances of ROA.

In case of ROE prediction, J48 algorithm takes the lead with a minimum proportion of correctly classified instances is 86.31% for No.56 industry and reaches maximum proportion at 95.77% for No.42 industry, and the average proportion is 91.48%. The random forest algorithm follows with a minimum proportion of correctly classified instances is 80.05 % for No.10 industry and reaches maximum proportion at 91.48% for No.42 industry, and the average proportion is 86.37%. Three remain algorithms, Naïve Bayes, SVM, and MLP have the close performance in this case. SVM algorithm is a little better with average proportions of correctly classified instances is 55.82% while MLP reaches 54.00% and Naïve Bayes has 49.51%.

- Result of ROE prediction:

Proportion of Correctly Classified Instances

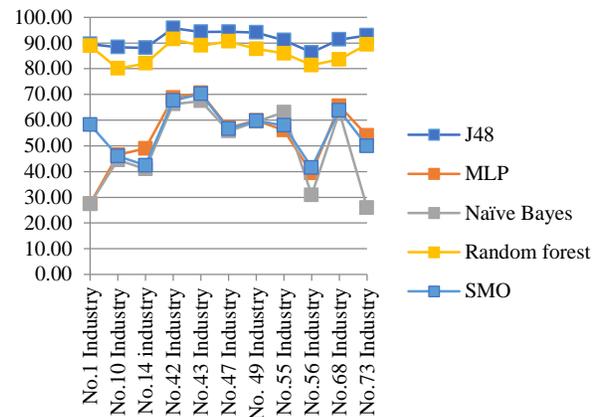


Fig. 4. The Proportion of Correctly Classified Instances of ROE.

1) Performance of ML algorithms for all industry groups:

To evaluate the performance of every selected algorithm in all industry, three measures are used including precision, recall, and ROC-area. The tables below show the performances of all algorithms, and the bold numbers show the best results of each measurement on each industry.

The tables 5-8 show that in both cases, regarding the ROC-area the random forest algorithm outperforms all other algorithms and J48 algorithm follows closely. In the case of ROA prediction, the random forest algorithm is better than J48 according to recall and precision (Table 5). However, J48 is better than random forest according to recall and precision in ROE prediction (Table 7). In this study, the random forest algorithm is the best for ROA prediction problem and J48 is chosen for ROE prediction problem.

- Result of ROA prediction:

TABLE V. THE PERFORMANCES OF NAÏVE BAYES, DECISION TREE AND RANDOM FOREST

Index	Code/ Division	Algorithms								
		Naïve Bayes			Decision tree			Random Forest		
		Precision	Recall	ROC-area	Precision	Recall	ROC-area	Precision	Recall	ROC-area
1	01	0.436	0.203	0.582	0.911	0.911	0.953	0.906	0.912	0.988
2	10	0.625	0.442	0.762	0.879	0.883	0.929	0.865	0.879	0.976
3	14	0.517	0.285	0.512	0.879	0.881	0.913	0.898	0.896	0.979
4	42	0.767	0.368	0.673	0.956	0.958	0.951	0.960	0.964	0.994
5	43	0.684	0.327	0.797	0.941	0.943	0.964	0.940	0.946	0.994
6	47	0.534	0.264	0.701	0.944	0.944	0.974	0.948	0.950	0.995
7	49	0.632	0.277	0.634	0.940	0.940	0.959	0.938	0.945	0.993
8	55	0.592	0.094	0.576	0.913	0.911	0.944	0.900	0.908	0.982
9	56	0.523	0.069	0.529	0.861	0.863	0.913	0.865	0.874	0.971
10	68	0.524	0.321	0.653	0.912	0.914	0.951	0.879	0.893	0.983
11	73	0.416	0.252	0.584	0.928	0.930	0.965	0.934	0.934	0.992

TABLE VI. THE PERFORMANCES OF SVM AND MLP

Index	Code/ Division	Algorithms					
		Support Vector Machine			Multi-perceptron		
		Precision	Recall	ROC-area	Precision	Recall	ROC-area
1	01	?	0.560	0.553	?	0.583	0.668
2	10	?	0.684	0.616	?	0.704	0.74
3	14	1	0.021	0.575	0.212	0.075	0.765
4	42	?	0.861	0.710	?	0.874	0.808
5	43	?	0.869	0.873	?	0.869	0.873
6	47	?	0.760	0.815	?	0.760	0.815
7	49	?	0.808	0.703	?	0.818	0.765
8	55	?	0.749	0.637	?	0.770	0.731
9	56	?	0.667	0.506	?	0.672	0.636
10	68	?	0.735	0.691	?	0.751	0.807
11	73	?	0.716	0.652	?	0.749	0.817

- Result of ROE prediction:

TABLE VII. THE PERFORMANCE OF NAÏVE BAYES, DECISION TREE AND RANDOM FOREST

Index	Code/ Division	Algorithms								
		Naïve Bayes			Decision tree			Random Forest		
		Precision	Recall	ROC-area	Precision	Recall	ROC-area	Precision	Recall	ROC-area
1	01	0.336	0.274	0.557	0.892	0.893	0.954	0.884	0.889	0.986
2	10	0.357	0.445	0.648	0.834	0.833	0.936	0.786	0.801	0.960
3	14	0.341	0.410	0.633	0.848	0.851	0.934	0.812	0.820	0.968
4	42	0.561	0.662	0.672	0.913	0.913	0.956	0.914	0.915	0.989
5	43	0.586	0.674	0.742	0.908	0.907	0.962	0.881	0.890	0.984
6	47	0.432	0.557	0.643	0.914	0.914	0.964	0.904	0.906	0.989
7	49	0.499	0.594	0.633	0.897	0.895	0.955	0.872	0.877	0.983
8	55	0.666	0.631	0.858	0.880	0.882	0.942	0.853	0.859	0.977
9	56	0.433	0.309	0.565	0.838	0.837	0.917	0.804	0.814	0.961
10	68	0.542	0.631	0.765	0.889	0.889	0.954	0.821	0.836	0.975
11	73	0.525	0.259	0.698	0.896	0.896	0.949	0.894	0.894	0.987

TABLE VIII. THE PERFORMANCES OF SVM AND MLP

Index	Code/ Division	Algorithms					
		Support Vector Machine			Multi-perceptron		
		Precision	Recall	ROC-area	Precision	Recall	ROC-area
1	01	?	0.583	0.668	?	0.478	0.587
2	10	?	0.460	0.572	?	0.465	0.656
3	14	0.462	0.267	0.615	0.478	0.490	0.736
4	42	?	0.676	0.524	?	0.688	0.708
5	43	?	0.702	0.534	?	0.706	0.742
6	47	0.449	0.566	0.508	?	0.570	0.686
7	49	?	0.598	0.508	?	0.598	0.628
8	55	?	0.579	0.510	?	0.561	0.611
9	56	?	0.415	0.508	?	0.394	0.603
10	68	?	0.638	0.561	?	0.655	0.773
11	73	?	0.499	0.522	?	0.540	0.710

2) Performance of best ML algorithms for each class:
Deeply analyze the efficiency of the best algorithms in both ROA and ROE prediction problems, the algorithms' performances on different classes are depicted in the following tables and figures.

- Result of ROA prediction is presented in Table 9 and Figure 5:

TABLE IX. THE PERFORMANCE OF RANDOM FOREST ALGORITHM

Class label	0	1	2	3	4
No.1 Industry	99.86	96.86	87.18	19.83	81.82
No.10 Industry	99.31	97.58	73.92	8.62	36.81
No.14 industry	96.77	96.79	84.70	27.52	53.42
No.42 Industry	99.36	99.64	70.31	16.90	50.44
No.43 Industry	100.00	98.56	65.59	7.55	33.33
No.47 Industry	99.80	97.99	86.86	42.27	75.16
No. 49 Industry	99.84	98.99	76.78	6.67	36.36
No.55 Industry	68.29	97.51	78.76	9.72	80.97
No.56 Industry	55.56	95.51	82.18	16.00	59.17
No.68 Industry	99.75	98.11	64.80	8.97	32.61
No.73 Industry	100.00	98.10	84.82	29.47	70.48
Average	92.59	97.78	77.81	17.59	55.51

Proportion of correctly classified instances per class

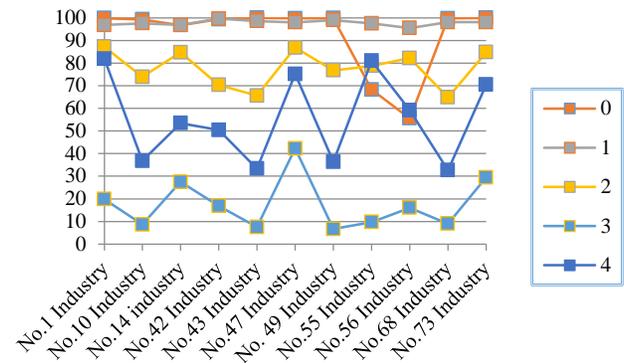


Fig. 5. Proportion of Correctly Classified Instances of the Best Algorithm per Class.

In the case of ROA prediction problem, the random forest algorithm is efficient for classes labeled 0,1,2 with the average proportions of correctly classified instances are 92.59%, 97.78%, and 77.81% relatively. However, it fails for a class labeled 3 with the average correction is only 17.59%, and it reached an average correct proportion at 55.51% for a class labeled 4.

- Result of ROE prediction is presented in Table 10 and Figure 6:

TABLE X. THE PERFORMANCES OF RANDOM FOREST ALGORITHM

Class label	0	1	2	3	4
No.1 Industry	96.63	87.98	76.44	48.68	87.48
No.10 Industry	94.29	79.07	11.46	45.50	82.40
No.14 industry	92.80	84.99	10.68	45.86	86.99
No.42 Industry	96.19	84.94	10.52	50.68	85.38
No.43 Industry	96.29	82.81	16.07	35.24	86.26
No.47 Industry	96.79	87.84	8.74	56.49	89.37
No. 49 Industry	95.87	85.21	12.89	47.70	82.13
No.55 Industry	95.81	84.42	13.62	28.04	79.41
No.56 Industry	93.01	82.60	13.58	43.52	82.11
No.68 Industry	97.38	79.56	14.34	41.11	86.43
No.73 Industry	95.58	88.33	10.03	57.02	89.14
Average	96.12	85.61	18.86	49.08	86.40

Proportion of correctly classified instances per class

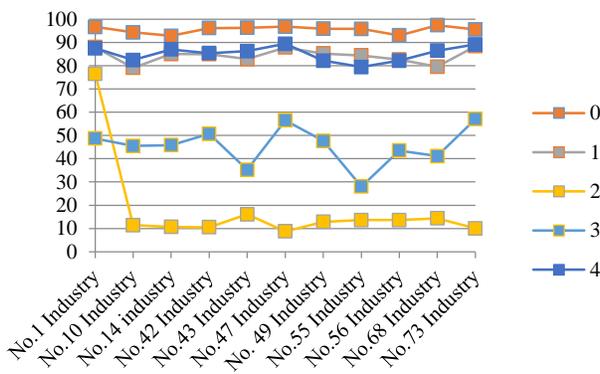


Fig. 6. Proportion of Correctly Classified Instances of the Best Algorithm per Class.

In case of ROE prediction problem, the J48 algorithm is efficient for classes labeled 0,1,4 with the average proportions of correctly classified instances are 96,12%, 85.61%, and 86.40% relatively. However, it fails for class labeled 2 with the average correction is only 18.86%, and it reached the average correct proportion at 49.08% for class labeled 3.

In both cases, the best algorithms are not successful in all classes, however, they succeed in the classification of largest categories including classes labeled 0, 1. Predict these classes are very important to building an early warning system for macroeconomic management as proposing in session 3.

V. CONCLUSION

Forecasting firms’ performance for macroeconomic perspective is still an ongoing research area however it is becoming more and more important given the development of

digital economy. This study reviewed the disjointed published studies on this research area and consolidated theoretically the application potential of firm performance prediction by ML techniques in macro-economic prediction problem. The study investigated the ability of utilizing micro-level information in macroeconomic monitoring and proposed a framework to process firm-level information to generate on demand information. This study also mentioned the major proven machine learning algorithms for firm’s performance prediction used in micro perspectives chronologically, and these algorithms were fundamentals to conduct a research framework that was tested on Vietnam’s economic data, keeping in mind that this data contains firm’s public information only. This research gave evidence to prove the enormous potential of proposed model for macroeconomic manager.

Particularly, five great machine learning algorithms were studied on data of Vietnamese companies belong to the different industry groups to identify the most suitable model for each industry groups. The applied research methodology had two stages, with the first stage preprocessed and divided the raw data set into sub-datasets belongs to different industry groups and validated these sub-datasets to satisfy the requirements of using machine learning algorithms. The second stage performed main processes including sub-data set partitioning, training, and testing processes for each ML algorithms. The testing result was evaluated by several measurements to ensure comparing comprehensiveness. Besides, this approach opened the ability to improve quality of final models by combination new data dimensional reduction techniques and machine learning algorithms together.

Experiments showed that in both cases, ROA and ROE prediction, regarding the ROC-area the random forest algorithm outperformed all other algorithms and J48 algorithm follows closely. In case of ROA prediction, random forest algorithm was better than J48 according to recall and precision also (table 5). However, J48 was better than random forest according to recall and precision in ROE prediction (table 7). In this study, random forest algorithm was the best for ROA prediction problem and J48 was chosen for ROE prediction problem.

For both ROA and ROE, the best algorithms was not successful in all classes, however, they succeed in classification of the largest categories including classes labeled 0, 1. In fact, predicting these low performance classes is very important to build an early warning system for macroeconomic management.

The proposed approach has high opportunity to use for macroeconomic management because the fast pace of modern economy requires the monitoring decision making in shorter and shorter period time. In fact, the e-government model has been developed and digital economy supplies a large and detail data at firm-level in high frequency and this high-volume data supports to create better machine learning model for firm’s performance prediction used in macroeconomic perspective. In expected case, automatic mechanism can be built, and it can generate early warnings for policy makers about economic state.

This study has some limitations on theoretical and experiment sides. The proposed approach utilizes the same selected variables for all machine learning algorithms, and this is not an ideal procedure. Theoretically, each ML algorithm might suit to different variable selection method therefore the ideal procedure should take this fact into account. On the other hand, the testing data set still contains exceptions and bizarre instances. Both limitations are going to be dealt with in the expanded stage of this study in near future.

ACKNOWLEDGMENT

This research is funded by the Ministry of Education and Training and Thuongmai University under grant number B2020-TMA-04.

REFERENCES

- [1] J. Gans, S. King, R. Stonecash, and N. G. Mankiw, *Principles of economics*. Cengage Learning, 2011.
- [2] C. N. Wang and V. T. Phan, "Enhancing the accurate of grey prediction for GDP growth rate in Vietnam," *Proc. - 2014 Int. Symp. Comput. Consum. Control. IS3C 2014*, pp. 1137–1139, 2014, doi: 10.1109/IS3C.2014.295.
- [3] J. H. Stock and M. W. Watson, "Forecasting in ation," *J. Monet. Econ.*, vol. 44, no. 2, pp. 293–335, 1999.
- [4] A. L. Montgomery, V. Zarnowitz, R. S. Tsay, and G. C. Tiao, "Forecasting the US unemployment rate," *J. Am. Stat. Assoc.*, vol. 93, no. 442, pp. 478–493, 1998.
- [5] M. T. Leung, A. Chen, and H. Daouk, "COM&OR_2000 Forecasting Exchange Rates Using General Regression Neural networks.pdf," vol. 27, 2000.
- [6] A. Banerjee, M. Marcellinoà, and I. Masten, "Leading Indicators for Euro-area Inflation and GDP Growth*," *Oxf. Bull. Econ. Stat.*, vol. 67, pp. 785–813, 2005.
- [7] H. S. Atesoglu, "Growth and fluctuations in the USA: a demand-oriented explanation," *Econ. demand led growth, challenging supply-side Vis. long-run*, pp. 55–63, 2002.
- [8] K. A. Aastveit, K. R. Gerdrup, A. S. Jore, and L. A. Thorsrud, "Nowcasting GDP in real time: A density combination approach," *J. Bus. Econ. Stat.*, vol. 32, no. 1, pp. 48–68, 2014, doi: 10.1080/07350015.2013.844155.
- [9] B. Bok, D. Caratelli, D. Giannone, A. M. Sbordone, and A. Tambalotti, "Macroeconomic Nowcasting and Forecasting with Big Data," *Annu. Rev. Econom.*, vol. 10, no. 1, pp. 615–643, 2018, doi: 10.1146/annurev-economics-080217-053214.
- [10] J. W. Galbraith and G. Tkacz, "Nowcasting GDP with electronic payments data," 2015.
- [11] C. Reinhart, M. Goldstein, and G. Kaminsky, "Assessing financial vulnerability, an early warning system for emerging markets: Introduction," 2000.
- [12] A. Ciarlone and G. Trebeschi, "Designing an early warning system for debt crises," *Emerg. Mark. Rev.*, vol. 6, no. 4, pp. 376–395, 2005.
- [13] C. Sevim, A. Oztekin, O. Bali, S. Gumus, and E. Guresen, "Developing an early warning system to predict currency crises," *Eur. J. Oper. Res.*, vol. 237, no. 3, pp. 1095–1104, 2014.
- [14] J. F. Brito, "Economic Growth as the Result of Firms Aggregate Performance: Evidence from the OECD Countries," *Econ. Manag. Res. Proj. An Int. J.*, vol. 3, no. 1, pp. 24–31, 2013.
- [15] K. Martinsen, F. Ravazzolo, and F. Wulfsberg, "Forecasting macroeconomic variables using disaggregate survey data," *Int. J. Forecast.*, vol. 30, no. 1, pp. 65–77, 2014, doi: <https://doi.org/10.1016/j.ijforecast.2013.02.003>.
- [16] P. Fornaro, "Predicting Finnish economic activity using firm-level data," *Int. J. Forecast.*, vol. 32, no. 1, pp. 10–19, 2016, doi: <https://doi.org/10.1016/j.ijforecast.2015.04.002>.
- [17] N. G. . van der Wijst, "The influence of CEO compensation on firm performance and its relation to economic growth," 2018.
- [18] E. J. Bartelsman and Z. Wolf, "FORECASTING AGGREGATE PRODUCTIVITY USING INFORMATION FROM FIRM-LEVEL DATA," *Rev. Econ. Stat.*, vol. 96, no. July, pp. 745–755, 2014, doi: 10.1162/REST_a_00395.
- [19] K. Dopfer, J. Foster, and J. Potts, "Micro-meso-macro," *J. Evol. Econ.*, vol. 14, no. 3, pp. 263–279, 2004, doi: 10.1007/s00191-004-0193-0.
- [20] S. T. Silva, A. A. . Teixeira, and M. R. Silva, "Economics of the firm and economic growth: a hybrid theoretical framework of analysis," *J. Organ. Transform. Soc. Chang.*, vol. 2, no. 3, pp. 255–274, 2005, doi: 10.1386/jots.2.3.255/1.
- [21] E. M. Al-Matari, A. K. Al-Swidi, and F. H. B. Fadzil, "The Measurements of Firm Performance's Dimensions," *Asian J. Financ. Account.*, vol. 6, no. 1, p. 24, 2014, doi: 10.5296/ajfa.v6i1.4761.
- [22] J. Darroch, "Knowledge management, innovation and firm performance," *J. Knowl. Manag.*, vol. 9, no. 3, pp. 101–115, 2005.
- [23] P. Goyal, Z. Rahman, and A. A. Kazmi, "Corporate sustainability performance and firm performance research," *Manag. Decis.*, 2013.
- [24] N. Omar, Z. 'Amirah Johari, and M. Smith, "Predicting fraudulent financial reporting using artificial neural network," *J. Financ. Crime*, vol. 24, no. 2, pp. 362–387, 2017.
- [25] C. Demartini, "Performance Management System. A Literature Review," in *Performance Management Systems*, Springer, 2014, pp. 55–88.
- [26] C. Kuzey, A. Uyar, and D. Delen, "The impact of multinationality on firm value: A comparative analysis of machine learning techniques," *Decis. Support Syst.*, vol. 59, no. 1, pp. 127–142, 2014, doi: 10.1016/j.dss.2013.11.001.
- [27] X. Jin, J. Wang, T. Chu, and J. Xia, "Knowledge source strategy and enterprise innovation performance: dynamic analysis based on machine learning," *Technol. Anal. Strateg. Manag.*, vol. 30, no. 1, pp. 71–83, 2018, doi: 10.1080/09537325.2017.1286011.
- [28] F.-H. Chen and H. Howard, "An alternative model for the analysis of detecting electronic industries earnings management using stepwise regression, random forest, and decision tree," *Soft Comput.*, vol. 20, no. 5, pp. 1945–1960, 2016.
- [29] F. Barboza, H. Kimura, and E. Altman, "Machine learning models and bankruptcy prediction," *Expert Syst. Appl.*, vol. 83, pp. 405–417, 2017, doi: 10.1016/j.eswa.2017.04.006.
- [30] P. Hajek and R. Henriques, "Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods," *Knowledge-Based Syst.*, vol. 128, pp. 139–152, 2017, doi: 10.1016/j.knosys.2017.05.001.
- [31] X. Y. Qiu, P. Srinivasan, and Y. Hu, "Supervised learning models to predict firm performance with annual reports: An empirical study," *J. Assoc. Inf. Sci. Technol.*, vol. 65, no. 2, pp. 400–413, 2014.
- [32] Y. Li, W. Jiang, L. Yang, and T. Wu, "On neural networks and learning systems for business computing," *Neurocomputing*, vol. 275, pp. 1150–1159, 2018, doi: 10.1016/j.neucom.2017.09.054.
- [33] Q. Cao, M. A. Thompson, and Y. Yu, "Sentiment analysis in decision sciences research: An illustration to IT governance," *Decis. Support Syst.*, vol. 54, no. 2, pp. 1010–1015, 2013, doi: 10.1016/j.dss.2012.10.026.
- [34] P. Hajek, V. Olej, and R. Myskova, "Forecasting corporate financial performance using sentiment in annual reports for stakeholders' decision-making," *Technol. Econ. Dev. Econ.*, vol. 20, no. 4, pp. 721–738, 2014, doi: 10.3846/20294913.2014.979456.
- [35] J. Li, H. Bu, and J. Wu, "Sentiment-aware stock market prediction: A deep learning method," *14th Int. Conf. Serv. Syst. Serv. Manag. ICSSSM 2017 - Proc.*, 2017, doi: 10.1109/ICSSSM.2017.7996306.
- [36] D. Delen, C. Kuzey, and A. Uyar, "Measuring firm performance using financial ratios: A decision tree approach," *Expert Syst. Appl.*, vol. 40, no. 10, pp. 3970–3983, 2013, doi: 10.1016/j.eswa.2013.01.012.
- [37] E. Zibanezhad, D. Foroghi, and A. Monadjemi, "Applying decision tree to predict bankruptcy," *Proc. - 2011 IEEE Int. Conf. Comput. Sci. Autom. Eng. CSAE 2011*, vol. 4, pp. 165–169, 2011, doi: 10.1109/CSAE.2011.5952826.

- [38] P. du Jardin, "Dynamics of firm financial evolution and bankruptcy prediction," *Expert Syst. Appl.*, vol. 75, pp. 25–43, 2017, doi: 10.1016/j.eswa.2017.01.016.
- [39] A. Gepp and K. Kumar, "Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques," *Procedia Comput. Sci.*, vol. 54, pp. 396–404, 2015, doi: 10.1016/j.procs.2015.06.046.
- [40] E. Basti, C. Kuzey, and D. Delen, "Analyzing initial public offerings' short-term performance using decision trees and SVMs," *Decis. Support Syst.*, vol. 73, pp. 15–27, 2015, doi: 10.1016/j.dss.2015.02.011.
- [41] B. Yeo and D. Grant, "Predicting service industry performance using decision tree analysis," *Int. J. Inf. Manage.*, vol. 38, no. 1, pp. 288–300, 2018, doi: 10.1016/j.ijinfomgt.2017.10.002.
- [42] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [43] A. Fan and M. Palaniswami, "Selecting bankruptcy predictors using a support vector machine approach," in *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, 2000*, vol. 6, pp. 354–359.
- [44] K. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, no. 1–2, pp. 307–319, 2003.
- [45] A. Kazem, E. Sharifi, F. K. Hussain, M. Saberi, and O. K. Hussain, "Support vector regression with chaos-based firefly algorithm for stock market price forecasting," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 947–958, 2013.
- [46] F. E. H. Tay and L. Cao, "Application of support vector machines in financial time series forecasting," *Omega*, vol. 29, no. 4, pp. 309–317, 2001.
- [47] J. H. Min and Y. C. Lee, "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters," *Expert Syst. Appl.*, vol. 28, no. 4, pp. 603–614, 2005, doi: 10.1016/j.eswa.2004.12.008.
- [48] C.-H. Wu, G.-H. Tzeng, Y.-J. Goo, and W.-C. Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," *Expert Syst. Appl.*, vol. 32, no. 2, pp. 397–408, 2007.
- [49] F. Lin, C.-C. Yeh, and M.-Y. Lee, "The use of hybrid manifold learning and support vector machines in the prediction of business failure," *Knowledge-Based Syst.*, vol. 24, no. 1, pp. 95–101, 2011.
- [50] F. Ecer, "Comparing the bank failure prediction performance of neural networks and support vector machines: The Turkish case," *Econ. Res. Istraživanja*, vol. 26, no. 3, pp. 81–98, 2013.
- [51] X. Wu et al., "ERMiner: Sequential rule mining using equivalence classes," *Indian J. Sci. Technol.*, vol. 7, no. 1, pp. 68–76, 2014, doi: 10.1109/ICDM.2004.10117.
- [52] M. D. Odom and R. Sharda, "A neural network model for bankruptcy prediction," in *1990 IJCNN International Joint Conference on neural networks, 1990*, pp. 163–168.
- [53] K. Y. Tam, "Neural network models and the prediction of bank bankruptcy," *Omega*, vol. 19, no. 5, pp. 429–445, 1991.
- [54] P. K. Coats and L. F. Fant, "Recognizing financial distress patterns using a neural network tool," *Financ. Manag.*, pp. 142–155, 1993.
- [55] T. B. Bell, "Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures," *Intell. Syst. Accounting, Financ. Manag.*, vol. 6, no. 3, pp. 249–264, 1997.
- [56] D. Vlachos, "Neuro-fuzzy modeling in bankruptcy prediction," *Yugosl. J. Oper. Res.*, vol. 13, no. 2, 2016.
- [57] A. Tsakonas, G. Dounias, M. Doumpos, and C. Zopounidis, "Bankruptcy prediction with neural logic networks by means of grammar-guided genetic programming," *Expert Syst. Appl.*, vol. 30, no. 3, pp. 449–461, 2006.
- [58] B. J. Zaini, S. M. Shamsuddin, and S. H. Jaaman, "Comparison between rough set theory and logistic regression for classifying firm's performance," *J. Qual. Meas. Anal. JQMA*, vol. 4, no. 1, pp. 141–153, 2008.
- [59] G. H. John and P. Langley, "Estimating continuous distributions in Bayesian classifiers," *arXiv Prepr. arXiv:1302.4964*, 2013.
- [60] J. R. Quinlan, "Simplifying decision trees," *Int. J. Man. Mach. Stud.*, vol. 27, no. 3, pp. 221–234, 1987.
- [61] J. H. Stock and M. W. Watson, "Diffusion indexes," *NBER Work. Pap.*, no. w6702, 1998.
- [62] J. H. Stock and M. W. Watson, "Macroeconomic forecasting using diffusion indexes," *J. Bus. Econ. Stat.*, vol. 20, no. 2, pp. 147–162, 2002.
- [63] E. S. Harris, *Tracking the economy with the purchasing managers index*. Federal Reserve Bank, 1991.

A Review on SDR, Spectrum Sensing, and CR-based IoT in Cognitive Radio Networks

Nadia Kassri¹, Abdeslam Ennouaary²

National Institute of Posts and Telecommunications, Rabat, Morocco

Slimane Bah³

Mohammadia School of Engineers
Rabat, Morocco

Hajar Baghdadi⁴

Faculty of Science and Technology
Settat, Morocco

Abstract—The inherent scarcity of frequency spectrum, along with the fixed spectrum allocation adopted policy, has led to a dire shortage of this indispensable resource. Furthermore, with the tremendous growth of wireless applications, this problem is intensified as the unlicensed frequency spectrum becomes overcrowded and unable to meet the requirement of emerging radio devices operating at higher data rates. Additionally, the already assigned spectrum is underutilized. That has prompted researchers to look for a way to address spectrum scarcity and enable efficient use of the available spectrum. In this context, Cognitive Radio (CR) technology has been proposed as a potential means to overcome this issue by introducing opportunistic usage to less congested portions of the licensed spectrum. In addition to outlining the fundamentals of Cognitive Radio, including Dynamic Spectrum Access (DSA) paradigms and CR functions, this paper has a three-fold objective: first, providing an overview of Software Defined Radio (SDR), in which the architecture, benefits, and ongoing challenges of SDR are presented; second, giving an extensive review of spectrum sensing, covering sensing types, narrowband and wideband sensing schemes with their pros and cons, Machine Learning-based sensing, and open issues that need to be further addressed in this field; third, exploring the use of Cognitive Radio in the Internet of Things (IoT) while highlighting the crucial contribution of CR in enabling IoT. This Review is elaborated in an informative fashion to help new researchers entering the area of Cognitive Radio Networks (CRN) to easily get involved.

Keywords—Cognitive radio; cognitive radio networks; software defined radio; spectrum sensing; machine learning; CR-based IoT

I. INTRODUCTION

The radiofrequency spectrum represents a scarce and finite resource that is used for transmitting information in the radio environment. This resource is used by several services, including radiocommunication, radio broadcasting, maritime radio, and satellite communications.

At the national level, the assignment of this spectrum to these services is managed and regulated by local authorities (governmental agencies) that are responsible for determining the appropriate frequency band, the geographical extent of the use of this band, the maximum transmission power, etc. One of the fundamental purposes of these agencies is to ensure a minimum interference level between the different radio technologies.

At the global level, the International Telecommunication Union (ITU) organizes, every three to four years, the world radiocommunication conferences (WRC)¹ so as to examine and revise the treaties governing the use of the radio frequency spectrum.

Nevertheless, the static spectrum allocation strategy adopted, where spectral frequency bands allocated to a wireless communication system can only be used by that system, has caused the shortage of frequencies. This shortage is confronted with a strong demand for spectrum resulting from the emergence and abundance of wireless technologies and the extremely rapid proliferation of radio applications developed in the scope of the Internet of Things (IoT).

In addition, while unlicensed bands like ISM (Industrial, Scientific, and Medical) bands can be freely used by all radios respecting a specific set of rules, such as a shared channel access mechanism and a maximum power per Hertz, they have become very crowded and can't accommodate more wireless applications [1].

On the other hand, the already assigned spectrum is inefficiently used in all domains, such as the time domain, the space domain, and the frequency domain, as shown in Fig. 1. This was confirmed by the Federal Communications Commission (FCC), which reported in 2002 that the radio spectrum, in most of the time, was from 15% up to 85% underutilized [2].

As a result, it has been found that while some bands are overcrowded, such as those bands used by cellular base stations, many other bands are not in use or are used only for short periods [2].

Cognitive Radio (CR) technology has been commonly regarded as an efficient solution to address the above-mentioned issues by enabling the opportunistic usage of the frequency bands that are not heavily occupied by licensed users [3,4].

¹ <https://www.ntia.doc.gov/category/wrc-19>

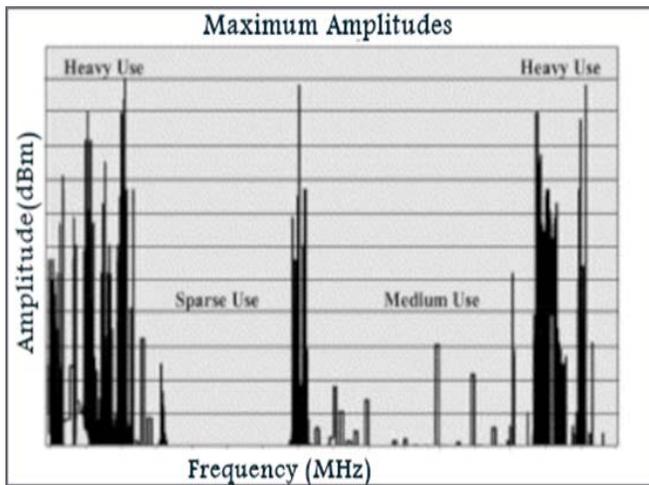


Fig. 1. Spectrum Utilization [3].

A Cognitive Radio represents a Software Defined Radio (SDR) system [5] capable of exploring the radio environment, learning, and deciding to use the unoccupied portions of the spectrum, called "spectrum holes" or "white spaces", and consequently of dynamically adjusting its operating parameters, based on the decisions made. These decisions must imperatively take into account the impact on the primary users (PUs), which have a license on the chosen spectrum and have the higher priority to access it, in order not to cause them harmful interference. Unlicensed users, which use Cognitive Radio to access the spectrum, are referred to as secondary users (SUs).

Furthermore, the transition from analog to digital television has released many frequencies, known as TV white spaces, which have too high propagation and penetration qualities [6]. For this reason, the FCC authorized in 2008 the use of these frequencies by non-licensed users following a list of rules like using the geo-location capability to obtain the available TV bands, from White Space DataBase (WSDB), before operating [7–9]. That has created a lot of opportunities for Cognitive radio users.

The remarkable contribution of this promising technology to the efficient use of spectrum, the minimization of interference, the reduction of cost, etc., has pushed its use in several areas, namely the Internet of Things (IoT).

The main objective of this paper is to provide a comprehensive survey of Cognitive radio with a particular focus given to:

- Software Defined Radio, as it is the building block of Cognitive Radio system;
- Spectrum sensing, as one of the most important pillars to set up a Cognitive Radio system;
- CR-Based IoT, given that Cognitive Radio is one of the important enabler technologies of IoT.

To the best of our knowledge, in the literature, there is no such work that gives a comprehensive survey of the three aforementioned aspects, highlights the relation between them, and exhibits the future research directions to handle their

challenges in a single paper. The existing surveys dealt only with a particular aspect.

The remainder of this paper is organized as follows: Section II gives the fundamentals of Cognitive Radio technology. Section III exhibits an overview of Software Defined Radio and its different challenges. Section IV provides a detailed view of spectrum sensing and its challenges. Section V highlights the use of Cognitive Radio in IoT. Finally, this article is concluded in Section VI.

II. COGNITIVE RADIO FUNDAMENTALS

Cognitive Radio technology was first introduced by Joseph Mitola in 1999 to depict an intelligent radio system capable of reconfiguring dynamically its radio parameters according to its operational environment and to the user's QoS requirements [4]. Afterward, Cognitive Radio has been defined in many ways by different entities and researchers. The Federal Communications Commission defined CR as follows: "Cognitive radio: A radio or system that senses its operational electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify system operation, such as maximize throughput, mitigate interference, facilitate interoperability, access secondary markets." [10].

Moreover, Simon Haykin defined CR as "an intelligent wireless communication system, capable of being aware of its environment, learning, and adaptively changing its operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time for providing reliable communication (anytime and anywhere) and efficient utilization of the radio spectrum" [11].

As mentioned in the above definitions, the main characteristics of Cognitive Radio are cognitive capability and reconfigurability [3,11].

Thanks to the cognitive capability feature, the CR user can sense and collect information related to its radio environment and choose the best channel to use. The collected information mainly involves transmission frequency, power, bandwidth, modulation, etc.

The reconfigurability feature allows CR users to adjust automatically its operating parameters (transmission frequency, modulation, power, etc.) based on the gathered information and without the need to change the hardware. Hence an efficient and effective Cognitive Radio system is imperatively built on Software Defined Radio platform. The latter will be briefly reviewed in the next section.

A. Dynamic Spectrum Access

By means of Dynamic Spectrum Access (DSA) techniques, Cognitive Radio enables secondary users to use the unoccupied portions of licensed spectrum which are known as spectrum holes or white spaces. When a primary user appears, the secondary user vacates the current band and moves to another spectrum hole or keeps transmitting on the same band and adjusting its transmission power level or modulation scheme in order not to cause any harmful interference to the licensed user or affect its QoS as illustrated in Fig. 2 [3].

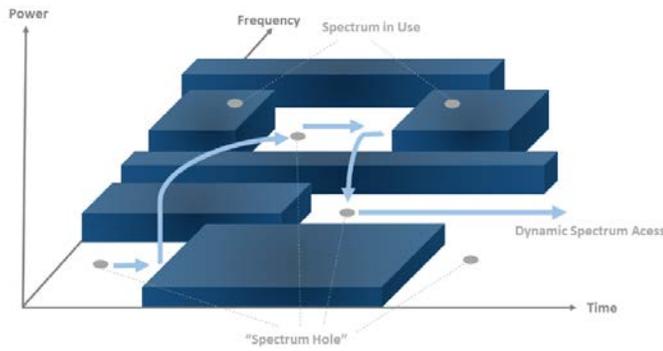


Fig. 2. Spectrum Hole Concept [3].

There are three DSA approaches that the SUs can use to dynamically access the available spectral opportunities: underlay, interweave, and overlay [12].

1) *Spectrum underlay*: In the underlay transmission mode, the SU can coexist with the PU as long as the SU operates below the noise floor of the PU. In other words, the coexistence between the SU and the PU may occur if only the interference caused by the SU at the PU receiver remains under a predefined threshold, known as the “interference temperature” [13,14]. To meet the interference threshold requirement, SU can use some techniques such as spreading its signal over a wide bandwidth below the noise floor of the PU or using multiple antennas to direct its signals away from the PU receiver [15].

Moreover, in the case of the underlay approach, the spectrum sensing process is not required and the SUs can transmit the data over the spectrum even if it is occupied all the time by the PUs. Therefore, this paradigm is more appropriate to use in situations when the spectrum usage status of PUs changes rapidly [12].

Nevertheless, meeting the interference constraint forces the SU to limit its transmission power and consequently to operate over a small coverage.

2) *Spectrum interweave*: As regards the interweave transmission mode, it stands for the opportunistic usage of the available spectrum holes which was the major motivation behind the introduction of the Cognitive Radio technology [16]. This paradigm mandates that SUs should have only access to the unoccupied spectrum resources. Thus, spectrum sensing is needed to identify the vacant spectrum bands that are not currently occupied by the PUs.

Furthermore, in this case, harmful interferences to PUs are avoided due to the fact that SUs do not transmit concurrently with the PUs. Whereas, the constant tracking of spectrum opportunities is a challenging task, especially in rapidly changing environments in terms of spectrum occupancy. Hence, this approach is more adequate for slowly changing environments [12,17].

Concerning power resources, they are only limited by the range of the identified spectrum holes.

It is also worth noting that this class of dynamic spectrum access is also referred to as opportunistic spectrum access [18].

3) *Spectrum overlay*: In respect to the overlay transmission mode, it is like the underlay approach in that both approaches allow SUs to simultaneously transmit with PUs. However, in the overlay mode, SUs can transmit at any power without a predefined interference threshold constraint.

Moreover, in this paradigm, the performance of the PU shouldn't be negatively affected by the presence of the SU. To meet this requirement, SUs use a variety of techniques that require prior knowledge about PUs' codebooks and messages [19]. For instance, SUs can divide their power into two parts: one is assigned to transmit their packets and the other is allocated to support the PUs' transmissions. In that way, the interference caused by the SU at the PU receiver can be compensated by the enhancement of the PU's signal-to-noise power ratio (SNR), by dint of the part of SU's power that is used to relay the PU packets.

Additionally, SUs can exploit knowledge about PUs' codebooks and messages to cancel the interference caused by PUs, using techniques such as dirty paper coding [19].

Finally, hybrid schemes that combine these transmission modes can be conceived to increase the overall throughput of the wireless networks [20].

B. Cognitive Radio Functions

The various tasks performed by a Cognitive Radio, including detecting spectrum holes, selecting the best available channel, determining the transmission parameters, sharing spectrum with other users, and moving to another frequency band when a licensed user appears are referred to as the cognitive cycle (See Fig. 3) [21].

In general, the cognitive cycle can be divided into three functional steps, namely spectrum sensing and analysis, spectrum management and handoff, and spectrum allocation and sharing [21]:

- **Spectrum sensing and analysis**: in this phase, CR monitors the radio environment, detects the spectrum holes, and estimates its different characteristics.
- **Spectrum management and handoff**: at this stage, CR chooses the best spectrum white space, determines the transmission parameters, and hops among different bands based on the channel characteristics and user requirements [3,21].
- **Spectrum allocation and sharing**: through these functions, CR can share and coordinate the spectrum access with other users. The coexistence with licensed users is restricted by their allowable interference level. Thus, secondary users should adjust their transmission parameters accordingly. As for sharing spectrum access with CR users, efficient spectrum access coordination is required so as to avoid collisions and interferences [21].

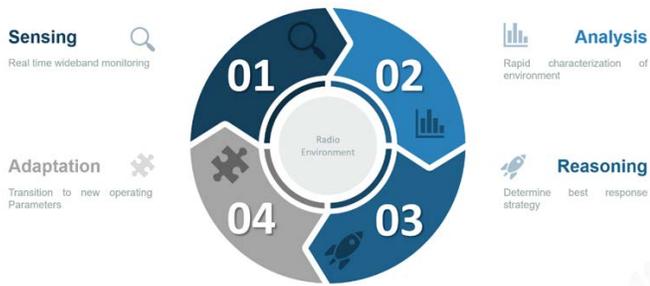


Fig. 3. Cognitive Cycle [21].

With respect to the layer they are in, the aforesaid functions can be classified into three groups namely physical layer (PHY) functions, the medium access control layer (MAC) functions, and network layer functions as shown in Fig. 4 [22]:

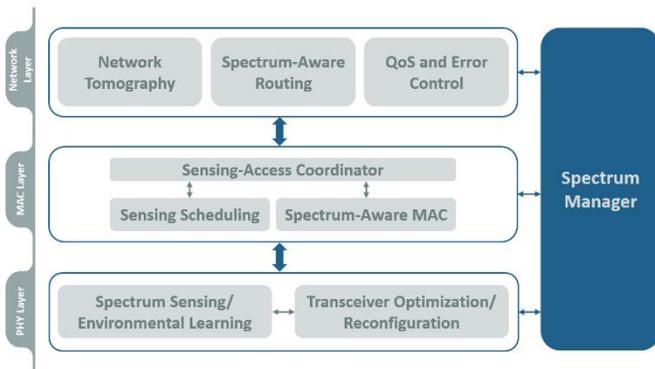


Fig. 4. Key Functions of the PHY, MAC, and Network Layers in a CR [22].

1) *Physical layer functions:* CR PHY layer is responsible for many tasks, including, but not limited to, identifying spectrum holes through spectrum sensing and acquiring advanced radio environment knowledge, such as the channel gain from the CR transmitter to the primary receiver and channel-state information (CSI), via environmental learning [22]. Based on the gathered information, cognitive spectrum access is performed by using transceiver optimization and reconfiguration.

2) *Medium access control layer functions:* CR MAC layer incorporates two key functions that allow to control and manage spectrum sensing operations as well as access to the identified spectrum opportunities. The first function is known as sensing scheduling, whereas the second is referred to as spectrum-aware access control. The operations of these functions are controlled by the sensing-access coordinator, on a time basis, by taking into consideration the compromise between the sensing requirement and the availability of the spectrum access opportunity.

In comparison with the conventional wireless communications MAC layer, the CR MAC layer is more complicated to implement because of the dynamic nature of the radio environment and the constant adaptation process that the CR should perform. More detailed information about the CR MAC layer can be found in [23,24].

3) *Network layer functions:* CR network layer provides three main functions: network tomography, quality of service (QoS) and error control, and spectrum-aware routing [22].

Network tomography refers to the operation by which the CR nodes sense the traffic patterns of the primary and the coexisting networks. The output of this operation provides important baseline data that allows a better understanding of the routing design and the network utilization at the packet level.

Quality of service (QoS) control and error control are of paramount importance to build a successful CR network. Statistical control can be used to address these tasks over CRNs that are characterized by opportunistic links.

Furthermore, In CR networks, data routing is a challenging problem due to varying link quality, frequent topology changes, and sporadic connectivity caused by the movement of PUs in the network [25]. In such a changing environment, the spectrum-aware routing function should be enabled in order to find optimal routes and paths while avoiding PUs. This function has two problems to handle [21]: the first problem relates to the fact that routing algorithms and protocols should be aware of the various network characteristics such as spectrum availability, PU activity, channel switching delay, and link qualities and take them into consideration while executing their different operations. The second problem deals with the setting up of interaction between routing algorithms and dynamic spectrum allocation routines so as to select routing paths with minimum interferences.

Moreover, the optimization of the energy consumption, the desired QoS, and the spectrum management should be considered while designing routing algorithms and protocols for CR networks. Routing protocols are further detailed in [26].

Finally, the spectrum manager serves as the means of establishing a connection between the three aforementioned layers and ensuring dynamic and efficient access to the available spectrum.

C. Cognitive Radio Networks

A Cognitive Radio Network (CRN) consists of a number of CR nodes with or without a secondary base station.

Based on the presence of infrastructure support, CRNs can be classified as either an Infrastructure-based network (centralized) or an Ad-Hoc network (distributed). Moreover, CRNs can be also deployed in another architecture known as Mesh architecture that combines Infrastructure-based and Ad-Hoc modes. A brief description of these architectures is given below.

1) *Infrastructure-based CR network:* It is a centralized architecture which contains a CR base station that is responsible for controlling and coordinating the transmission activities of the CR nodes. In this architecture, the CR base station retrieves the spectrum related information from all the SUs in the network and on the basis of the gathered information, it makes decisions on spectrum access and sharing for all CR nodes.

Compared to Ad-hoc network, infrastructure-based network provides many advantages including, the reliability of the sensing process, collision avoidance, and a high data right for SUs. However, an infrastructure is required to build such a network.

Some examples of centralized networks include:

- The IEEE 802.22 network: it is the first wireless regional area network standard that defines specifications for broadband wireless access using CR technology in TVWS bands [27];
- Spectrum Efficient Uni and Multi-cast Services Over Dynamic Radio Network in Vehicular Environments (Over DRiVE) [28];
- European Dynamic Radio for IP services in Vehicular Environment (DRiVE) [29];
- Wi-Fi (IEEE 802.11) [30].

2) *Ad-hoc mode CR network*: In Ad-hoc CRNs, there is no need for base stations or access points to coordinate the SUs. Thus, in the absence of a central controlling entity, SUs in such a distributed network, make independent decisions concerning spectrum access and transmission parameters. Additionally, SUs must use distributed DSA protocols so as to manage the spectrum access operation. Designing distributed DSA protocols for Ad-hoc CRNs is a challenging task because of the absence of central control entities and the completely distributed networking architecture [12]. These protocols should support a set of functions including, transparency for PUs, collision avoidance, accurate spectrum sensing, and efficient dynamic spectrum allocation [12].

III. SOFTWARE DEFINED RADIO: OVERVIEW

Given that SDR is considered as an enabling technology that handles the implementation of Cognitive Radio, this section will be dedicated to a brief overview of SDR.

A. Definition of SDR

Software Defined Radio [31,32] is referred to as a programmable radio transceiver where digital signal processing functions, such as modulation/demodulation, coding/decoding, error control, interleaving/deinterleaving, and scrambling/descrambling are implemented by means of software instead of hardware as it has been used in traditional radio communication systems. That enables the implementation of different waveform standards in a single platform and switching between them without any change in the hardware components. In some cases, there is just the need of a simple software upgrade to support other modes, bands, and functions.

Many hardware platforms are used to implement the software part of SDR, mainly, General Purpose Processor (GPP), Digital Signal Processor (DSP), Field Programmable Gate Arrays (FPGA), and Application Specific Integrated Circuit (ASIC) [33].

Each of the above-mentioned platforms has its own challenges, limitations, and strengths in terms of computational power, power consumption, implementation cost, flexibility and reconfigurability, and complexity of design (See Table I).

B. SDR Architecture

Joseph Mitola proposed the architecture of an ideal software radio transceiver that includes three components, namely an antenna, a Digital-to-Analog/Analog-to-Digital converter, and a processing unit as illustrated in Fig. 5. The processing unit allows performing, in software, all the digital signal processing functions, including modulation/demodulation, coding/decoding, and error control [3].

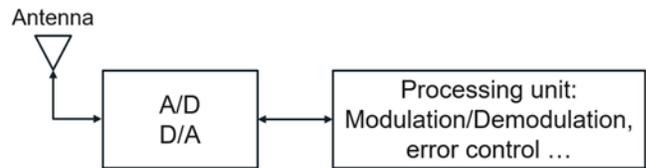


Fig. 5. Ideal Software Radio [16].

Many factors impede the real implementation of this proposed architecture such as the technology limitations, more particularly those related to ADC's performance, computing power, and power consumption of the processing unit.

Unlike the ideal software radio transceiver, the SDR transceiver is feasible. Indeed, the general architecture of SDRs contains more components that make its real implementation possible.

As shown in Fig. 6, the SDR transceiver includes mainly four parts which are an antenna, an analog RF front end, a digital RF front end, and the signal processing unit.

1) *Antenna*: One of the main tasks expected from an SDR platform is its ability to cover multiple frequency bands. Thus, SDR platforms often use Intelligent/Smart antennas so as to fulfill the aforementioned task.

A smart antenna consists of an antenna array combined with signal processing blocks that allow to smartly exploit the spatial diversity in order to select the appropriate frequency band and adapt with interference nulling, and mobile tracking [34,35].

An antenna for SDR should ideally integrate some features such as self-adaptation, self-alignment and self-healing [35]. These characteristics can be defined as follows:

- Self-adaptation: it is the capability of an antenna to adapt its parameters according to the selected band and the system requirements (gain...);
- Self-alignment: it is the capability of an antenna to control its radiation pattern;
- Self-healing: it is the capability of an antenna to avoid interferences.

2) *Analog RF front end:* The RF front end part [34] represents an analog circuitry where the following operations are performed:

In the transmission path, Digital-to-Analogue Converter (DAC) converts digital samples into an analog signal which represents the input of the RF Front End. After that, the analog signal is mixed with high frequency carriers, modulated to a preset RF frequency, and then transmitted.

In the receiving path, the RF signal captured by the antenna is fed to the RF Front end section through a matching circuitry which allows achieving an optimum signal power transfer. Then, in order to amplify very low-power signals while guaranteeing a minimum noise level, the RF signal passes through a Low Noise Amplifier (LNA) which is often mounted very close to the antenna. Afterward, the output of the LNA is mixed with a signal from the Local Oscillator (LO) so as to shift to a lower fixed frequency known as intermediate frequency (IF).

The frequency generated by the Local Oscillator is adjustable to ensure that the mixer produces a lower fixed intermediate frequency independent of the incoming RF signal.

The intermediate frequency presents several advantages: it improves the selectivity due to the fact of being fixed and it increases the performance of the processing unit and the global gain of the receiver owing to be lower than the incoming frequency.

3) *Digital front end:* The Digital Front end section [34] contains a set of blocks that are responsible for performing all of the succeeding functions:

In the transmission path, the digital baseband signal (near-zero frequency range) is shifted into the IF frequency by the Digital Up Converter (DUC) then it passes through the Digital Analog Converter (DAC) that converts it to the analog IF signal. The analog IF signal is later up-converted to RF signal.

In the receiving path, the analog IF signal is converted to digital IF samples by the Analog to Digital Converter (ADC). Next, the Digital Down Converter (DDC) changes the IF samples to a baseband signal which is then resampled and filtered before being processed by the signal processing unit.

These above-mentioned tasks can be divided into two principal functions which are Resampling and Channelization. Resampling or Sample Rate Conversion (SRC) is the process of converting samples from one sample rate to another. As for Channelization, it involves up/down conversion and channel filtering.

4) *Signal processing unit:* This block is regarded as the main part of SDR architecture which is designed to perform Digital Signal Processing functions such as modulation/demodulation and encoding/decoding.

As already mentioned in this paper many hardware platforms are used to implement this section, namely DSP, GPP, GPU, ASIC, and FPGA [36–39]. The real challenge is determining which of these platforms is best to achieve SDR goals and which design approach is appropriate to meet the requested quality of service.

Table I shows the strengths and weaknesses of some hardware platforms and highlights certain techniques used to improve their performance.

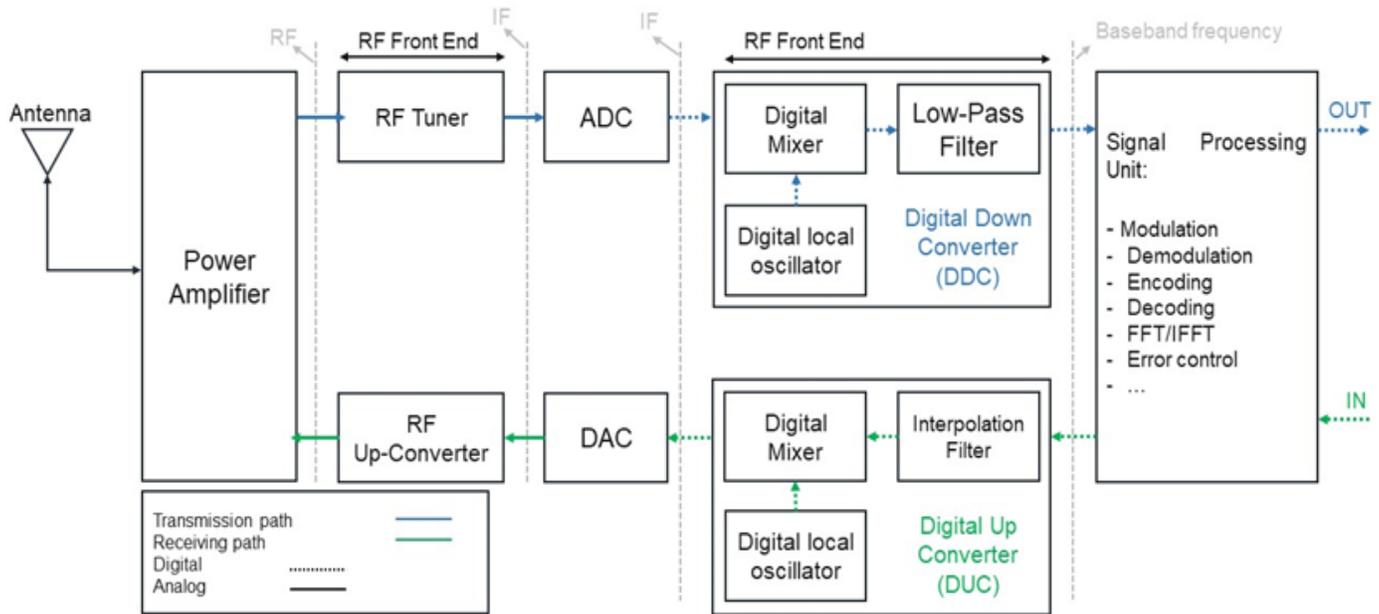


Fig. 6. SDR Transceiver.

TABLE I. COMPARISON OF SDR HARDWARE PLATFORMS

Hardware	Strengths	Weaknesses	Performance improvement
GPP [36–38]	Extremely flexible. Easily programmable. Reconfigurable. Use of high-level languages. Cost reduction. Functions portability. Extremely flexible. Easily programmable. Reconfigurable. Use of high-level languages (C++...). Cost reduction. Functions portability.	Real time applications. are not supported. Power inefficient.	Use of multi-core GPPs to enable the parallelism and perform more operations per clock cycle. Use of GPU for intensive computing task (Turbo code, Fast Fourier Transform...).
DSP [39]	Extremely flexible. Easily programmable. Power efficient. Faster than GPP. Use of high-level languages (C...).	Lack of the required processing speeds for wideband transmissions. Consuming more power than FPGA.	Running two DSPs in parallel.
ASIC [38]	Computationally powerful. Performing at higher speed than FPGA. Smaller in size. More power efficient than GPP.	Not reprogrammable. Very expensive (Each chip is designed for a specific application).	–
FPGA [38]	Computationally powerful. Capability to implement any design or function. Power efficient. Programmability. Seamless switching between. modes and functions. High speed performance. Cheaper than ASIC.	Difficulty to implement new modules. Prior Knowledge of the hardware architecture is required for an efficient module implementation. The implementation task takes a lot of time. Limited portability (HDL code). Consuming more power than ASIC. Taking more area than ASICs.	Use of High-Level Synthesis that enables the rapid implementation of new functions with no prior experience with hardware design.

Several metrics are used to compare the performance of hardware platforms, the most used are computing power, energy consumption, flexibility and reconfiguration, adaptability, cost, and complexity.

To exploit the advantages of each platform, researchers have proposed the co-design (hybrid) approach as a solution. This approach consists of regrouping design schemes that employ hardware techniques, such as FPGAs and ASICs, and those that use software solutions including GPPs, into one platform.

This approach presents some issues, the most known are the problem related to the shared access of the internal memory by different units (FPGAs, Processors ...) and the very expensive cost of the whole implementation.

One of the interesting readings that give a detailed review of the aforementioned platforms, can be found in [35].

C. SDR Advantages

Software Defined Radio technology has brought a lot of advantages to the world of wireless communications, such as reconfigurability and flexibility, interoperability, and cost reduction [34,37]:

- **Reconfigurability and Flexibility:** The main feature of SDR is its capacity to handle wireless standards newly developed by simply changing or upgrading the reconfigurable software instead of replacing the hardware platform or its analog components.

- **Interoperability:** One of the most important benefits of SDR is its ability to resolve interoperability problems between incompatible radios that work with different wireless standards.
- **Cost reduction:** Several expensive hardware components disappeared due to the fact that in SDR many digital signal processing functions are implemented by means of software.

In addition to the forenamed benefits, SDR allows to test and study several communication standards by using the same platform. Furthermore, SDR is regarded as the block building of a Cognitive Radio system and as already mentioned in this paper, SDR enables the reconfigurability of Cognitive Radio.

D. SDR Tools

There are different commercial and open-source tools that can be used for SDR development. Choosing the more appropriate development tool for a specific design methodology requires prior knowledge of the available tools' features. Examples of these tools include GNU Radio, Universal Software Radio Peripheral (USRP), MATLAB, LabVIEW, and CUDA.

GNU Radio and USRP are the most widely used tools to develop SDR systems [40]. A brief description of these two tools is given below:

GNU Radio: GNU Radio is a free & open-source software toolkit for building software radios [35]. It runs on host

computers and it provides many signal processing blocks such as filters, decoders, and demodulators which are required to implement software radios. These blocks are programmed in C++ and they are often connected using Python script that has the advantage of allowing the data flow to be at the maximum rate, without being interpreted [41]. In addition, this tool offers the possibility of easily programming and adding new blocks for supporting unavailable functions. GNU Radio can be used jointly with Universal Software Radio Peripheral systems or other alternatives to set up complete SDR platforms.

USRP: Universal Software Radio Peripheral is the most commonly used transceiver for SDR platforms, developed by Ettus² Research under GPL license. USRPs are available in several series and versions, differentiated by their hardware performances and their connection method to a host computer. Generally, USRP is a board that integrates the following components: ADC/ DAC, an FPGA board, an RF front end, and a PC host interface [42]. Based on the USRP series, a certain number of daughterboards can also be held. Daughterboards perform functions such as filtering and conversion between RF and baseband signals (up/down conversions) and may support applications operating up to 6 GHz due to being modular. Concerning the other signal processing functions, the majority of them are performed at the host machine and only some operations are processed by the FPGA board [42]. USRP platforms offer several advantages, including ease-of-use, affordability, and flexibility. However, bandwidth limitations of USRP components have an impact on system throughput. That makes USRP platforms only suitable for research experiments and rapid prototyping [35,42].

E. SDR Challenges

In this subsection, the challenges that still remain to be addressed for implementing efficient and practical SDR systems are presented.

1) *Security and attack issues:* Although SDR has brought many powerful advantages to the field of wireless communication, it has introduced new types of security threats and attack issues. Indeed, in addition to the known threats that exist in conventional wireless communication systems such as denial of service attacks, misconfiguration issues, listening and capturing data that can be used to perform malicious actions, there are several threats that are specific to SDR, including downloading and running malicious software.

In the absence of authentication and verification techniques, unauthorized software can easily be installed and activated on SDR terminals. The challenge of implementing protection techniques and security mechanisms, to prevent this issue, has been addressed by many researchers. For instance, in [43] a framework for establishing secure download for SDR is presented. This framework uses a public/private key scheme to verify the authenticity of the software. The digital signature is considered as a good solution to keep malicious and unauthorized software from being activated on SDR nodes, nevertheless, it greatly increases the complexity of the

framework due to the fact that for each combination of waveform and terminal, a digital signature should be created.

Furthermore, the data configuration of SDR components can be prone to extraction, alteration, or destruction. Those issues can be prevented by guaranteeing the integrity of the security administrative module (SAM) and implementing data integrity and protection techniques [44].

2) *Energy efficiency:* Addressing the power consumption issue is of primary importance especially when designing solutions that are intended for battery-powered devices and low power objects in an IoT network. Several reasons are behind the importance of addressing power consumption management, including limited size and battery, ensuring a longer lifetime of IoT objects, and enabling Green Computing [17,45]. In SDR, many factors contribute to the loss of energy efficiency, namely signal processing complexity and the increased hardware requirements. To alleviate this loss, the authors of [46] proposed a cooperative wireless network scheme that is based on resource sharing (Battery, processing unit, memory...).

3) *Antenna requirements for SDR:* One of the major challenges in SDR is to design wideband antennas that support different technologies and standards. Although smart and reconfigurable antennas are used to resolve this concern, several constraints still hamper the SDR implementation in portable handsets and other systems, such as bandwidth and gain limitations, which are imposed by the antenna size, and the complexity of a design that meets all the antenna requirements for SDR [47].

4) *Hybrid design:* Implementing an optimal SDR design, that meets the real-time requirements at low power and cost while maintaining flexibility and programmability, represents a great challenge. Researchers address this issue by using the hybrid approach which enables the use of hardware schemes (ASICs...) along with software schemes (GPP...) in the same platform in order to take advantage of their different benefits.

The hybrid design has also its own challenges in both physical and MAC layers, including partitioning and scheduling problems [35].

5) *ADC and DAC limitations:* The concept of an ideal software defined radio transceiver consists in placing the ADC/DAC as close as possible to the antenna. That requires a very high-speed ADC/DAC with sampling capability up to Giga Samples per second which is actually not feasible. To overcome this issue, an RF front end block is placed right after the antenna so as to shift the incoming frequency to an intermediate frequency that can be supported by the currently available ADC/DAC.

IV. SPECTRUM SENSING

Sensing is considered as the most important and critical phase in the Cognitive Radio cycle. It refers to the operation by which the CR users can be aware of the channel occupancy, the presence of the primary user, the quality of the radio channel,

² <https://www.ettus.com/>

and other parameters such as transmission power, bandwidth, modulation, etc. Based on this gathered information, the CR users can determine the vacant portion of the spectrum and choose the white spaces which meet its QoS.

This section gives a detailed view of sensing spectrum function in Cognitive Radio Networks.

A. Multi-dimensional Spectrum Sensing

Through the sensing function, Cognitive radio can identify the existing spectrum opportunities in its surrounding environment by using different sensing techniques. Spectrum opportunity is conventionally defined as “a band of frequencies that are not being used by the primary user of that band at a particular time in a particular geographic area” [48,49]. This definition takes into consideration only three dimensions, namely frequency, time, and space. However, there are other dimensions that can be exploited to create new opportunities such as code dimension and angle dimension. Developing sensing algorithms that take into account all those dimensions encounters more complex and challenging issues.

A radio environment in which all the aforementioned dimensions are exploited to share spectrum access among multiple users is known by several names including, hyperspace, electro space, and radio spectrum space. Table II shows parameters that are needed to be sensed for each dimension and highlights its main idea [49].

B. Types of Spectrum Sensing

Spectrum sensing approaches can be divided into several types on the basis of specific aspects, including the cooperation between secondary users, the bands of interest, the sensing execution time, and the number of sensed channels at a time (See Fig. 7) [50]:

Based on the bands of interest for the Cognitive Radio system, spectrum sensing falls into two groups: in-band sensing and out-of-band sensing. In-band sensing consists in sensing the channel that is already transmitting, in the aim of detecting primary user signals and avoiding harmful interferences [51,52]. As to out-of-band sensing, CR senses bands other than the band on which it is transmitting so as to discover new spectrum holes [51].

Based on the sensing execution time, two classes of spectrum sensing can be distinguished: reactive (on-demand) and proactive(periodic) sensing. The reactive sensing takes place when the secondary user intends to transmit or as a result of radio environmental changes. In proactive sensing, CR users sense the radio environment persistently so as to detect spectrum opportunities. Reactive sensing is more energy efficient but the time to identify an unoccupied channel may be longer than proactive sensing.

Among the types of spectrum sensing, we also find synchronous sensing and asynchronous sensing. In synchronous sensing, all CRs respect the same schedule to sense a frequency band. In that case, a high synchronization between CRs represents a challenging task.

TABLE II. MULTI-DIMENSIONAL SPECTRUM SENSING

Dimension	Parameters to sense	Idea and remarks
Frequency	Available frequency bands	The available frequency band is segmented into disjoint sub-bands. Identifying opportunities in the frequency domain consists in determining the unoccupied sub-bands. It is unlikely that all the bands can be used concurrently at the same time.
Time	Available time slots in a specific band	For a given frequency band, there will be times when it is unoccupied and available for opportunistic usage.
Geographic Space	PU’s location (latitude, longitude, an elevation) Distance of primary users	At a specific time, certain channels may be available for opportunistic usage in some geographical zones while being entirely occupied in other zones. The path loss in space enables secondary users to identify the presence or the absence of a primary user in a given local area by just looking at the interference level. If there is no interference, then the primary user is absent otherwise it is present. However, the risk of creating harmful interferences to a hidden primary user is possible. This issue is discussed later in this article.
Angle	PU’s beam directions (azimuth and elevation angle). PU’s location.	By leveraging advanced antenna technologies (e.g., Beamforming technology), identifying PU’s location, and determining PU’s beam directions, new spectrum holes in the angle domain might be available for opportunistic usage. Indeed, the secondary user can transmit in the same frequency band along with the primary user at the same time in the same location by choosing a different direction without causing interferences to the licensed user.
Code	PU’s spreading code, time hopping (TH), or frequency hopping (FH) sequences. Awareness of timing information is needed so that SU can synchronize its transmission with PU.	By being aware of code sequences that primary users are using at a given time, secondary users can transmit simultaneously along with primary users over the available spectrum by choosing different code sequences in such a way as to avoid creating interferences on primary users. It is important to note that simultaneous transmission is possible by using orthogonal codes.

In the case of asynchronous sensing, each CR has its own schedule to sense a frequency band. Distinguishing between SU signals and PU signals is a challenging problem.

In respect to the cooperation aspect, there are two classes of spectrum sensing: cooperative and non-cooperative (local) Sensing. In non-cooperative sensing, each CR settles for its own sensing data and independently decides on the channel state, i.e. the presence or the absence of the primary user. Decisions that are made through this sensing type are unreliable and error prone under bad channel conditions, multipath fading and shadowing effects, and hidden node issues [53].

The aforementioned issues can be avoided by using cooperative sensing [21,53]. Indeed, when cooperative sensing is enabled, CR users share their local observations and sensing data with others and exploit the shared sensing outcomes of other users to decide on the presence or absence of the primary user. Several works have been carried out on this topic and it has been shown that cooperative sensing contributes to the improvement of detection accuracy and reliability on the cost of increased latency and traffic overhead. It has also been proven that cooperation between secondary users can solve hidden primary user problem and decrease sensing time [53].

There are two types of cooperative sensing: centralized sensing and distributed sensing [3,54]. In centralized sensing [55], a central unit gathers local sensing information from CR users through a control channel, fuses them by means of one of the fusion decision rules [56], identifies spectrum holes by performing binary hypothesis testing algorithm such as Neyman-Pearson test or Bayesian test, and shares the result among other cognitive users or directly controls the cognitive radio traffic. In distributed sensing [54], cognitive nodes exchange their local observations among each other and make their own decisions on channel state. Distributed sensing doesn't require a backbone infrastructure or a centralized base station.

On the basis of interference detection, spectrum sensing is categorized as primary transmitter detection, primary receiver detection, and interference temperature [51]. In primary transmitter detection approaches, spectrum holes are identified by processing received signals at the PU receiver. This class includes several techniques such as energy detection and matched filter detection. In primary receiver detection, the status of primary channels is detected based on the local oscillator leakage power of the PU receiver [57]. This power is emitted by the PU receiver's RF front end while receiving the data from the PU transmitter [58]. As for interference temperature, secondary users transmit simultaneously with primary users as long as the interference caused by the SU at the PU receiver remains under a specified interference limit. In this case, the underlay DSA model is considered.

In regard to the requirement of the PU's information, spectrum sensing schemes can be categorized into two classes: blind and feature detection techniques [50]. The blind detection techniques serve to blindly determine the channel state without any prior knowledge about the primary user signals. This type

includes, among others, energy detection and Higher-Order-Statistics detection. The feature detection techniques allow performing signal classification to the detected signal. They are more advantageous than blind detection techniques in the sense that it is possible to distinguish between PU and SU signals and to characterize the different types of PU and SU signals. Matched filter and cyclostationarity detection techniques are examples of feature detection techniques.

Depending on the bandwidth (number of sensed channels at a time), spectrum sensing schemes are classified into two broad groups which are narrowband and wideband sensing techniques [59]: in the narrowband sensing, only one channel is analyzed at a time to detect available opportunities. The most widely and commonly known narrowband methods are energy detection, matched filter detection, cyclostationary feature detection, covariance-based detection, and waveform detection [59].

As for wideband sensing, multiple frequency bands are explored at a time so as to discover spectrum holes. In this case, the available spectrum is usually divided into narrower sub-bands which are sensed either concurrently or sequentially by using narrowband sensing techniques. Wideband sensing techniques contain two groups: Nyquist-based and Sub-Nyquist wideband sensing. The former type processes digital signals at the rate equal to or greater than the Nyquist rate, whereas the latter type processes the signals with a sampling rate lower than the Nyquist rate [59].

Wideband sensing allows to exploit the available spectrum more efficiently but it brings about new challenges related to its implementation.

On the basis of the latest mentioned classification of spectrum sensing schemes (Narrowband sensing/ Wideband sensing), the next subsection will review some of the most common spectrum sensing methods as well as the recent advanced spectrum sensing techniques.

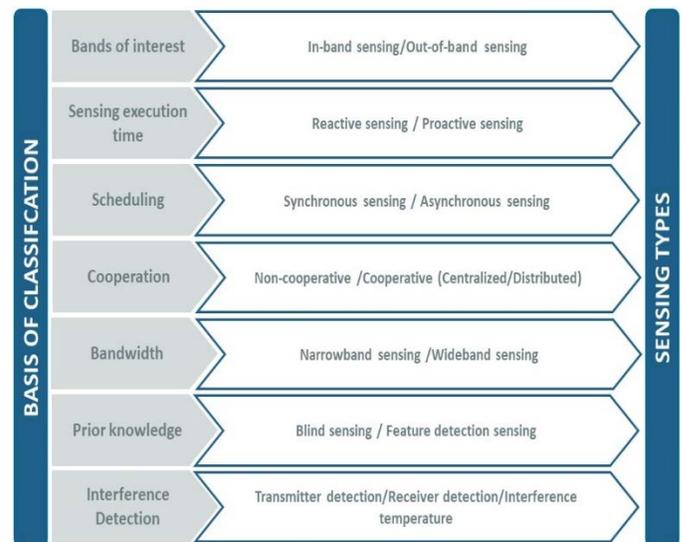


Fig. 7. Types of Spectrum Sensing.

C. Spectrum Sensing Techniques

Before reviewing in detail spectrum sensing techniques, more particularly narrowband sensing techniques, the system model for spectrum sensing is given below.

1) *System Model for spectrum sensing*: The fundamental objective of spectrum sensing is to make a choice between two hypotheses:

- H_0 : Channel is temporarily available for opportunistic usage i.e. no primary user signal is present;
- H_1 : Channel is occupied i.e. a primary user signal is present.

These hypotheses can be expressed as under:

$$H_0: Y(n) = N(n) \text{ when PU's signal is absent} \quad (1)$$

$$H_1: Y(n) = h.X(n) + N(n) \text{ when PU's signal is present} \quad (2)$$

Where $Y(n)$ represents the received signal by SU, $X(n)$ represents the transmitted PU signal, $N(n)$ denotes Additive White Gaussian Noise (AWGN) with mean zero and variance σ^2 . h represents the channel gain, $n = 1, 2, 3, \dots, N$ where n is the sample index, and N is the number of samples.

Generally, a test statistic of $Y(n)$ is compared against a decision threshold value λ in order to decide between the two hypotheses. If the test statistic is greater than the threshold, H_1 is assumed to be true and thus PU's signal is declared present. In contrast, if the test statistic is less than the threshold, H_0 is assumed to be true and then PU's signal is declared absent. Fig. 8 presents the general model of spectrum detection.



Fig. 8. Spectrum Sensing Model.

Sensing accuracy is typically evaluated using the Receiver Operating Characteristic (ROC) curves. These curves are the plot of the detection probability versus the false alarm probability or plot of the probability of miss detection against the probability of false alarm [60,61]. These probabilities are defined as:

Probability of detection (Pd): It is the probability that the SU declares the presence of PU signal when it is actually present.

$$Pd = \Pr(H_1/H_1) \quad (3)$$

A high Probability of detection allows avoiding interference impact on primary receivers. Hence a high Pd is eminently desirable.

Probability of false alarm (Pfa): It is the probability that the SU declares the presence of PU signal when it doesn't truly exist.

$$Pfa = \Pr(H_1/H_0) \quad (4)$$

A high Probability of false alarm decreases the efficiency spectral due to the loss of spectrum access opportunities. In

addition, the QoS may be negatively affected as a result of this loss. Thus, Pfa should be low to avoid the under-use of potential spectrum holes.

Probability of miss detection (Pm): It is the probability of missing a PU signal when it is present.

$$Pm = \Pr(H_0/H_1) \quad (5)$$

A high Probability of miss detection causes a harmful interference on primary users because, in the case of a miss detection, SUs may transmit simultaneously with PUs in the same band. So, to avoid interference to and from license holder users, Pm should be low.

2) *Narrowband sensing techniques*: In this subsection, we discuss several narrowband sensing techniques, focusing on their functions, mathematical models, strengths, and drawbacks.

a) *Energy detection scheme*: Energy detection [1,62–64] called also Radiometry or Periodogram, is the most widely used sensing scheme due to its low computational complexity, the simplicity of its implementation, and the absence of need for PU signal information.

In this sensing technique, the received signal energy is computed and compared with a threshold value. If the measured energy is higher than the threshold, the target frequency band is considered to be occupied by a PU; otherwise, the target frequency band is considered to be vacant. The block diagram of the energy detection method is illustrated in Fig. 9 and its decision metric can be written as:

$$T = \frac{1}{N} \sum_{n=1}^N |Y(n)|^2 \quad (6)$$

Despite its aforementioned advantages, energy detection has several drawbacks, including the degradation of detection accuracy as a result of noise uncertainty (NU). The latter refers to the noise power fluctuations with time caused by the effects of various factors such as thermal noise, filtering effects, radio-frequency circuits, and interference from other signals [65]. Authors of [65] analyzed the impact of NU on energy detection performance using different OFDM system designs. The results of their analyses show that, for all OFDM system designs, the probability of detection is higher for signals, with lower NU and higher SNR, which are transmitted at higher PU transmit power and detected with a higher number of samples.

Moreover, energy detection is unreliable at low SNR values [66], ineffective in detecting spread signals [60], and unable to distinguish between PU signals from other signals.

To enhance the performance and the accuracy of energy detection, several methods based on the use of dynamics thresholds are investigated in [67–70].

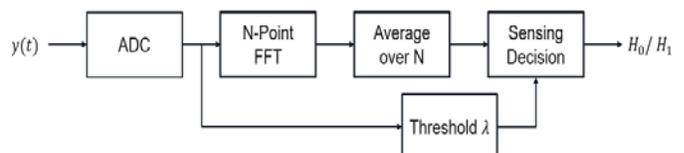


Fig. 9. Block Diagram of Energy Detection [1,62–64].

b) *Matched filter detection*: The matched filter detection technique is viewed as the best method for identifying the unutilized bands when the PU transmission characteristics are known a priori, e.g., bandwidth, operating frequency, the modulation type and order, packet format, and pulse shaping [70–72]. In this technique, the received signal is correlated with the known PU signal and the result is compared with a threshold to decide on the presence of the PU. The block diagram of matched filter detection is shown in Fig. 10 and its test statistic is given by:

$$T = \sum_{n=1}^N Y(n)X^*(n) \quad (7)$$

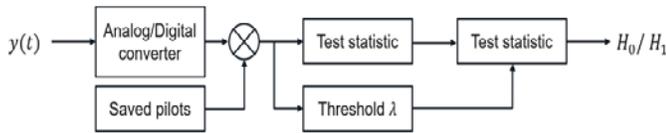


Fig. 10. Block Diagram of Matched Filter Detection [70–73].

The main advantage of this sensing scheme is that it requires only a small number of samples and thus less sensing time to achieve good detection performance due to its resilience against noise uncertainty [73]. Nevertheless, it has several drawbacks. First, there is a need for precise and accurate prior information about the primary user signal, which may not always be available [65]. Second, the performance of the matched filter technique drops when the prior information is incorrect. Third, to detect different types of PUs, a dedicated matched filter structure for each type must be used, which increases the complexity of the system.

c) *Cyclostationary Feature Detection*: Cyclostationary Feature Detection [74–76] depends on cyclostationary characteristics of the received signals. Indeed, in wireless communications, the transmitted signals are typically characterized by the periodicity of some of their statistics such as mean and autocorrelation [76,77]. This periodicity results from the fact that signals are modulated and coupled with cyclic prefixes, hopping sequences, pulse trains, sinusoidal wave carrier, and other features before being transmitted [76,77]. Furthermore, additive noise signals are stationary with no correlation. Therefore, Cyclostationary Feature Detection schemes are very robust to noise uncertainties and can distinguish between the signal and noise by calculating and analyzing the cyclic autocorrelation of the received signals.

The periodicity of the mean and the autocorrelation, of a cyclostationary signal $y(t)$, can be expressed mathematically by:

$$m_y(t) = E[y(t)] = m_y(t + T_0) \quad (8)$$

$$R_y(t, \tau) = R_y(t + T_0, \tau) \quad (9)$$

Where T_0 represents the period of the signal $y(t)$, τ represents the time offset, E denotes the expectation operator, and R_y is the autocorrelation function of $y(t)$ and it is given by:

$$R_y(\tau) = E[y(t + \tau)y^*(t - \tau)]e^{j2\pi\alpha\tau} \quad (10)$$

The block diagram of Cyclostationary Detection scheme is illustrated in Fig. 11, in which The Analog-to-Digital Converter (ADC) digitizes the received analog signal $y(t)$ into digital samples and then it is fed into the N-point FFT block which computes its Fast Fourier Transform. Next, these FFT values are correlated with themselves and then averaged over the number of samples. Finally, in the feature detection block, the sensing decision is obtained by detecting the features of the average outcome.

This scheme has also the advantage of being able to distinguish between various types of PU signals that are characterized by different transmit features [78].

However, this detector also presents some limitations such as the need for a large number of samples and high sampling rate which results in an increase in sensing time, power consumption, and complexity [76,78].

In order to reduce computation complexity while maintaining sufficient detection sensitivity, authors of [79] proposed an improved Cyclostationary Detector with SLC Diversity over Nakagami-m Fading channels, where the test statistic of conventional Cyclostationary detector is reliably simplified. For the same purpose, in [80], an improved Cyclostationary Feature Detection Algorithm is presented, in which authors proved that the cyclic spectrum is conjugate symmetry about the relevant axis, which decreases the computational complexity.

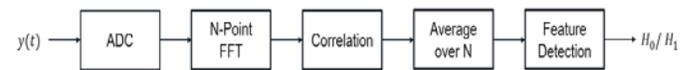


Fig. 11. Block Diagram of Cyclostationary Feature Detection [74–76].

d) *Covariance Based Detection*: Covariance-based detection methods exploit the correlation structure inherent in the received data and the noticeable differences between the statistical covariances of signal and noise to decide on the presence of primary signals in background noise without the need for prior knowledge about signal, channel, or noise power [81].

Different statistical tests can be extracted from the covariance matrix of the received signal and used to detect the presence of the signal like the ratio of the maximum eigenvalue to minimum eigenvalue in the case of Eigenvalue-Based Detection Method [82]. The steps of the latter can be summarized as follows: First, the sample covariance matrix of the received signal is built from the received signal samples. Then, the eigenvalues of this matrix are computed using techniques such as singular value decomposition (SVD). Finally, a decision on the presence of the signal is taken by comparing the ratio between the maximum eigenvalue and the minimum eigenvalue with a threshold. The block diagram of this method is shown in Fig. 12.

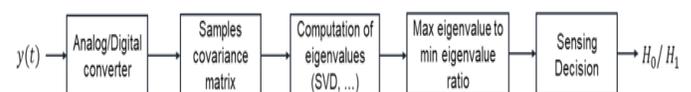


Fig. 12. Block Diagram of Covariance-Based Detection [82].

The sample covariance matrix of the received signal can be expressed as [82]:

$$R_y(N_s) = \frac{1}{N_s} \sum_{n=L-1}^{L-2+N_s} \hat{y}(n) \hat{y}^H(n) \quad (11)$$

Where N_s is the number of collected samples, L is a positive integer called "smoothing factor", $\hat{y}(n)$ is the received signal vector at sampling instance n , and $\hat{y}^H(n)$ is its conjugate-transpose.

The test statistic can be written as:

$$T = \frac{\lambda_{min}}{\lambda_{max}} \quad (12)$$

Where λ_{max} and λ_{min} are respectively the maximum and minimum eigenvalues of $R_y(N_s)$;

Using a predefined threshold λ , the decision can be made as follows: if $T > \lambda$, H_1 is considered to be true, otherwise, H_0 is considered to be true.

It is noteworthy that, despite the robustness of covariance techniques against noise uncertainty, they have a high computational complexity due to two factors, namely covariance matrix computation and eigenvalue decomposition [82].

e) Waveform-based detection: Waveform-based detection method, also known as coherent detection [49], can be considered as a simplified version of the matched filter scheme. Although, unlike the latter, this method doesn't need complete information about the PU signals. It simply exploits known patterns, that are used in wireless communication systems to support many functions such as synchronization, control, and equalization so as to execute coherent detection. These patterns involve but not limited to the following:

- Preambles: a preamble is a known sequence that is sent before each burst;
- Midambles: a midamble is a known sequence that is sent in the middle of a slot or burst;
- Pilot symbols: pilot symbols are an extra overhead added to the transmitted signal.

When such patterns are available, the received signal is correlated with a known copy of itself to perform signal detection. As such, the test statistic can be expressed as:

$$T = R[\sum_{n=1}^N y(n)x(n)^*] \quad (13)$$

where $R[\cdot]$ denotes the real part, $y(n)$ is the received signal by SU, $x(n)^*$ is the complex conjugate of the signal transmitted by PU, and $n = [1, 2, \dots, N]$ is the sample index.

In the absence of the primary user's signal, the given expression can be written as follows:

$$T = R[\sum_{n=1}^N n(n)x(n)^*] \quad (14)$$

Whereas in the presence of the primary user, the test static can be simplified as under:

$$T = \sum_{n=1}^N (|x[n]|)^2 + R[\sum_{n=1}^N n(n)x(n)^*] \quad (15)$$

Where $n(n)$ denotes Additive White Gaussian Noise (AWGN) with mean zero and variance σ^2 .

The value of the test statistic T is compared to a fixed threshold λ in order to decide on the presence of a primary user signal. H_1 is considered to be true if $T > \lambda$, and H_0 is considered to be true otherwise.

It is noteworthy that the waveform-based sensing scheme can only be applied to wireless systems having known signal patterns, such as Wi-Fi (IEEE 802.11b) [83] and WIMAX [84].

In [85], it is demonstrated that compared to energy detector technique, this method is more reliable and has less sensing time. In addition, it is demonstrated that the accuracy of this sensing scheme increases with the length of the known primary signal pattern.

In [86], the authors have opted for waveform-based sensing scheme to analyze the sensing performance of the simultaneous transmission-and-sensing (TS) mode, which has proven its effectiveness under imperfect self-interference signal (SIS), unlike energy detection that cannot differentiate between a PU signal and a residual SIS.

In spite of the aforementioned advantages, there are some downsides to this scheme, mainly the need for complete and accurate information concerning waveform patterns as well as high synchronization between PU and SU, which is not always possible to attain.

Concerning the implementation complexity, coherent detection is more complex than Energy detection method and has a lower complexity compared to match filter detection technique.

Table III provides a performance comparison of the previously described narrowband sensing methods based on a set of performance criteria, namely detection accuracy, complexity, need for prior information about the primary user, robustness against noise uncertainty, and sensing time required to achieve a good performance.

3) Wideband sensing techniques: Wideband spectrum sensing enables detecting spectral opportunities that lie within frequency bands greater than the coherence bandwidth of the channel. One application of this class of sensing is to exploit the available spectrum holes in the UHF (ultra-high-frequency) TV band ranging from 300 MHz to 3 GHz.

As have been discussed earlier, narrowband sensing schemes make a single binary decision for the whole frequency band. That makes these schemes incapable to perform directly wideband sensing, since the latter intends to determine the occupancy of several sub-channels that are involved in a given wideband spectrum at a time.

The two types of wideband sensing techniques (Nyquist / Sub-Nyquist) will be briefly reviewed in this subsection.

TABLE III. PERFORMANCE COMPARISON OF NARROWBAND SENSING TECHNIQUES

Criteria	Energy detection	Matched filter detection	Cyclostationary feature detection	Covariance-based detection	Waveform-based detection
Detection accuracy	Weak performance at low SNRs	Optimal performance at all SNRs	Very good performance at all SNRs	Moderate performance at all SNRs	Very good performance at all SNRs
Complexity	Low	High	High	High	Moderate
Need for prior information about primary user	No	Yes	Yes	No	Yes
Robustness against noise uncertainty	No	Yes	Yes	Yes	Yes
Sensing time required to achieve good performance	Low	Low	High	High	Low

a) *Nyquist wideband sensing*: Nyquist wideband sensing is carried out using a standard ADC operating at Nyquist rate to digitize the wideband signal and digital signal processing methods to detect spectral opportunities. It includes, among other, wavelet-based detection, multi-band joint detection, and filter bank detection:

Wavelet-based detection [87] determines spectrum holes over a wide frequency band, which is assumed to be composed of a multitude of sub-bands whose locations and power spectral densities are unknown, by using the wavelet transform. The latter serves to identify discontinuities that are located at the boundaries (edges) of each sub-band. The discontinuous changes on the edges correspond to irregularities in power spectral density. As such, information about the locations and intensities of the sub-bands can be retrieved through the wavelet transform and then their status (occupied or unoccupied) can be deduced.

Concerning the multiband joint detection algorithm [88], it senses the primary signal over multiple frequency channels by performing the following procedures: first, sampling the wideband signal using a high sampling rate analog to digital converter. Second, dividing the sampled data into parallel data streams by a serial-to-parallel converter. Third, computing the Fast Fourier Transform (FFT) of the digital signal. Fourth, dividing the wideband spectrum into multiple sub-bands that are occupied by narrowband signals. Finally, detecting spectrum opportunities over each sub-band by using a narrowband sensing scheme such as energy detection. Furthermore, this technique allows to jointly determine the optimal threshold, that maximizes the accuracy of sensing for all the sub-channels, by formulating an optimization problem.

As for filter bank detection [89], a filter bank, consisting of a set of bandpass filters, is used to separate the wideband signal into multiple sub-bands. The bandpass filters are realized through modulation (Poly-phase decomposition) of a prototype filter. The prototype filter is a lowpass filter that is used to implement the zeroth band of the filter bank. The obtained sub-bands are then sensed separately via a narrowband sensing scheme. In this type of sensing, the sub-bands can be down-converted and consequently re-sampled at a lower sampling rate. However, a large number of RF components will be added, resulting in a high implementation complexity [90].

While Nyquist wideband sensing allows to exploit efficiently the available spectrum, it has certain limitations, including the requirement of a high sampling rate (Wavelet-based detection and multiband joint detection) and the high implementation complexity (Filter bank).

b) *Sub-Nyquist wideband sensing*: Sub-Nyquist wideband sensing offers solutions to the challenges facing Nyquist-based sensing, more particularly high sampling rate and high implementation complexity concerns. Sub-Nyquist approaches aim to identify spectral opportunities by exploiting only a few measurements derived from processing wideband signals at sampling rates less than the Nyquist rate.

Among the most important Sub-Nyquist wideband sensing types is compressive-based sensing [91–93]. The latter refers to the process of recovering a sparse signal from a few measurements by following a three-step procedure (sparse representation, measurement, and sparse recovery). In the first step, the sparsity of the signal is derived by projecting the signal on an appropriate basis. In second, only a reduced number of measurements are retrieved by multiplying the sparse signal by a measurement matrix. In third, the signal is recovered by using the few collected measurements. Finally, the sensing operation is performed to identify the spectral opportunities.

It is noteworthy that the use of comprehensive sensing, in the context of wideband sensing, is enabled due to the sparsity feature that characterizes the wideband signal in the frequency domain.

Furthermore, the reliability of sub-Nyquist wideband sensing schemes depends largely on the accuracy of sparsity level estimation. Although, the latter suffers from uncertainty because of various factors, mainly the dynamic nature of PUs transmission activities as well as wireless channel impairments.

D. Machine Learning for Spectrum Sensing

As pointed out earlier, CR is an intelligent radio system that incorporates the three main components which any intelligent system should have, namely perception, learning, and reasoning [94]. Perception refers to the ability of the system to sense its radio environment and it can be realized by means of sensing measurements of the spectrum. Learning refers to the ability of the system to convert the gathered information into knowledge and it can be achieved through classification and

generalization algorithms. Finally, knowledge is exploited by the system to attain its objectives via reasoning ability [94]. Considering this intelligent design, machine learning based spectrum sensing presents itself as a good alternative to determine the channel occupancy in Cognitive Radio Networks by addressing two main issues which are classification and decision making [94].

Machine learning algorithms can be classified into two broad groups: supervised learning and unsupervised learning.

In supervised learning, a classifier learns from a training dataset to make predictions on unforeseen data. According to the training dataset, supervised machine learning problems can be further grouped into classification and regression problems. If the input variables are mapped to discrete output values (categories, labels, classes...), then it is a classification problem. If the input variables are mapped to continuous output values, then it is a regression problem. Some examples of supervised algorithms include support vector machines (SVM), Random forest, and naive Bayesian classifier (NBC).

As for unsupervised learning, its main function consists in finding patterns in a set of untagged data. K-means is one of the simplest and most widely used unsupervised machine learning algorithms.

In the context of Cognitive Radio Networks, several papers have dealt with the use of machine learning algorithms to predict the availability of frequency channels. For instance, in [95], the authors performed a comparative analysis of different supervised and unsupervised machine learning algorithms based on computational time and classification accuracy. They also proposed a new method that combines support vector machines with the firefly algorithm. The authors of [96] proposed a deep learning approach to learn channel activities and predict its availability in future time slots. The ability to predict the channel occupancy in the next time slots may increase the efficiency of selecting the more appropriate channel at the instant t (For example, choosing the channel having the highest probability of being unoccupied in the next time slots).

Generally, in the case of spectrum sensing, researchers adopt a two-phase machine learning approach [95,97]. In the first phase, an unsupervised learning technique, such as K-means algorithm, is applied to discover the transmission patterns of primary users. In the second phase, the discovered clusters (channel busy or free) are used to train supervised learning classifiers like support vector machine (SVM) and then to assign the new input data to the suitable cluster. It is worth mentioning that in several works, the sensing clusters are assumed to be known and thus the first phase is omitted. In this case, the researchers adopt a one-phase machine learning approach in which supervised learning techniques are trained with the already known clusters [98,99].

In building machine learning models for spectrum sensing, several features are used, including energy statistic, probability vector, and occupancy over time. Obviously, the accuracy of detecting the primary users may be affected by the selected features.

Finally, to evaluate the performance of a given model, a set of metrics can be used mainly, probability of detection, probability of false alarm, total error rate, sensing time, and accuracy.

E. Sensing Challenges

This subsection is devoted to discussing the various challenges related to the spectrum sensing process in Cognitive Radio Networks as well as highlighting some possible future research directions in this field.

1) *Hidden PU problem*: As the Carrier Sense Multiple Access (CSMA), which is characterized by the presence of hidden node problem, spectrum sensing suffers also from hidden primary user problem. The latter refers to the situation in which the secondary user misses the primary user presence due to many issues, including severe multipath fading and shadowing effects experienced by primary signals, during propagation from PU transmitter to SU receiver. As such, the secondary user may induce undesirable interference to primary user receivers. In several research papers, cooperative sensing has been proposed as a successful approach to handle this issue by exploiting spatial diversity [49,100].

2) *Challenges related to cooperative sensing*: Cooperative sensing has proven to be effective in improving the reliability and accuracy of detection, especially in the case of the presence of channel impairments. Moreover, it has also been shown to be useful in decreasing individual sensing durations and local processing requirements [21,53].

Despite its different advantages, cooperation among CR users also brought many challenges and invoked significant researches. One of these challenges is mitigating the additional signaling overhead and reducing delays that are induced by the cooperation process [101].

Another challenge is to find a tradeoff between the number of users participating in cooperation and additional processing requirements. Indeed, a large number of cooperative nodes can ensure a high probability of detection, even in the presence of channel uncertainties and detectors with less sensitivity, but it introduces a considerable amount of extra overhead which leads to an increase in the sensing time and processing. Hence, there exists a compromise.

Another challenge consists in developing asynchronous cooperative spectrum sensing algorithms [102,103]. In fact, CR nodes are located at different positions and may perform spectrum sensing at different times [104]. Therefore, CR nodes cannot report their local sensing results to the fusion center at the same time and thus some of the reported information may not be up to date [103]. Furthermore, there is always a time offset between the local observation and the final decision which may incur performance degradation. In addition, reporting channel uncertainties may impact negatively the sensing accuracy.

3) *Spectrum sensing duration*: Admittedly, a long sensing duration can guarantee a higher accuracy for spectrum sensing results, but it can compromise with interference avoidance, energy efficiency, and throughput.

Regarding interference avoidance, secondary users must be aware promptly of the presence of the primary user and vacate the frequency band rapidly when the incumbent user resumes its transmission in that band. Thus, a real trade-off arises between the quickness with which the secondary user must perform the aforementioned tasks and the required sensing duration to ensure more reliable results.

Concerning energy efficiency, the gain in accuracy achieved by a long sensing duration comes at the expense of energy. Hence, sensing duration and energy efficiency should be optimized.

As for throughput, it is inversely related to sensing time. Therefore, there is a compromise between longer sensing duration for higher performance and lower sensing duration for good throughput.

These trade-offs should be addressed jointly to find an optimum solution that ensure a higher performance under certain constraints.

4) *Noise uncertainty*: Several spectrum sensing techniques rely on a threshold value to decide on the presence of primary user signals. Hence, selecting the optimum threshold is crucial and important to ensure high accuracy of sensing results. To determine the best threshold value, many parameters should be taken into consideration such as noise power. The latter is uncertain and developing techniques that are robust to noise uncertainty is still a challenging task. Some research papers have addressed this issue using different approaches. For instance, in [105] a blind spectrum sensing technique based on goodness of fit testing of t - distribution has been proposed to cope with the noise uncertainty problem.

5) *Complexity and hardware requirements*: Several issues need to be addressed to bridge the gap between the theory and the hardware implementation realities of spectrum sensing schemes from both types, narrowband and wideband sensing. These include the complexity of narrowband sensing techniques, mainly covariance-based and cyclostationary feature detection schemes, which require large processing power that is unsuitable for portable devices. Hence, Additional efforts are required to reduce the complexity of these techniques and take advantage of their various benefits in real-life application. Besides, meeting the hardware requirements of wideband sensing schemes, like high-resolution ADCs and high-speed signal processors, while maintaining acceptable complexity and moderate computation power, is a challenging task.

6) *Sparsity level uncertainty*: In wideband compressive sensing, estimating the accurate sparsity level is of paramount importance, since it is a prerequisite for determining the optimal number of measurements. However, achieving this result is very difficult, especially in rapidly changing environments, due to random changes in spectrum activities and time-varying fading channels. That may lead to calculate the number of measurements using the worst-case sparsity level assumption, resulting in high energy consumption and inefficient use of sub-Nyquist sampling technologies.

Therefore, developing blind sub-Nyquist wideband sensing schemes, in which the estimation of sparsity level isn't needed, still a challenging task [90].

In the literature, only a few papers have dealt with the issue of sparsity level uncertainty. For instance, the authors of [106] proposed an algorithm for estimating the sparsity level of the channel over a learned dictionary using Machine Learning algorithms.

7) *Security*: As with any wireless network, Cognitive Radio Networks are vulnerable to various cybersecurity attacks that can be performed by selfish or malicious users and induce disruptive effects on network operation. Selfish users exploit network facilities for their interests, namely monopolizing the use of available spectrum opportunities and thus depriving legitimate Secondary users of their fair share of spectrum. As for malicious users, they abuse network facilities by exploiting the existing vulnerabilities and thereby hindering legitimate SUs from using the spectrum.

Some of the cybersecurity attacks in Cognitive Radio Networks include most active band (MAB), primary user emulation (PUE), and spectrum sensing data falsification (SSDF) attacks.

Most active band (MAB) attack aims at detecting and making the most active band, in a multi-band CR network, unavailable by targeting it via a denial of service (DoS) attack and consequently preventing the other users (PU and SU) from using it. The authors of [107] proposed a coordinated concealment strategy to counter this attack. In this strategy, a set of SUs cooperate and transmit useless data in a free band to make it the most active and therefore the attacker's target.

Primary user emulation (PUE) attack is an attack where a selfish or malicious node adjusts its air interface to emulate primary user characteristics and thus mislead secondary users concerning the availability of the spectrum holes [108]. During the sensing process, SUs, under this attack, may detect attacker signals as primary user signals and then refrain from using the spectrum. A range of countermeasures against PUE attack has been investigated in the literature [109]. These countermeasures can be categorized into four types: countermeasures based on cryptography, countermeasures based on fingerprint, countermeasures based on game theory, and hybrid countermeasures that combine the three other types [109].

Spectrum sensing data falsification (SSDF) attack, also known as Byzantine attack, specially targets cooperative spectrum sensing and consists in compromising the fusion center with false spectrum sensing results to deceive decision-making. In [110], a reputation-based approach, in which the fusion center recognizes the malicious attackers and eliminates them from the data fusion process, was proposed to counter SSDF attack.

While significant efforts have been carried out to detect and cope with cybersecurity issues, there are still several challenges that need to be addressed, including developing advanced techniques that do not need prior knowledge about PU location

, as this is not always available in real-world scenarios, developing detection and counter-measures methods based on cryptography that take into account resource constraint (e.g., power and bandwidth), and developing spectrum sensing techniques capable of differentiating between PU signals and malicious user signals.

V. CR-BASED IOT

The wide variety of IoT technologies have led to a rapidly increasing number of networked devices that are expected to reach 51.11 billion by 2023, according to research from Cisco states (See Fig. 13). Such devices involve daily tech gadgets such as smartphones and smart home devices, as well as industrial devices like smart machines and robots. These smart connected devices are capable of gathering, sharing, and analyzing information and creating actions accordingly. By 2023, global spending on IoT will reach 1.1 trillion U.S. dollar [111]. This puts IoT at the core of current and future social and economic transformation.

Furthermore, with the massive proliferation of these connected objects, there is an increased demand for spectrum resources. Cognitive Radio proves to be a promising technology to cater to current and future IoT devices in terms of spectrum access. In this section, the integration of the Internet of Things and Cognitive Radio is explored.

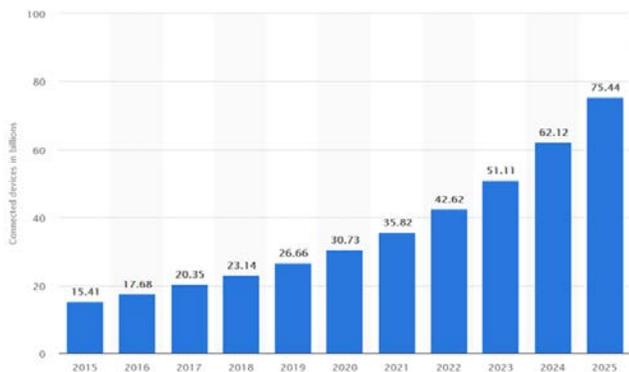


Fig. 13. IoT Connected Devices Installed base Worldwide from 2015 to 2025 (in billions)³.

A. Definition of IoT

The term "Internet of Things" was coined by Kevin Ashton to depict "a system in which the Internet is connected to the physical world by ubiquitous sensors"⁴.

Internet of Things (IoT) can be defined as a network of objects connected to the Internet and can communicate with each other using different communication technologies. These objects are equipped with sensors, that allow interaction with the environment, and communication modules [112–114]. Some of those sensors include temperature sensors, pressure sensors, proximity sensors, and humidity sensors.

Today, the IoT enables the real world and the virtual/or digital world interconnection. Thanks to IoT, things will

communicate with each other and develop their own intelligence. Televisions, cars, kitchen appliances, surveillance cameras, smartphones, utility meters, cardiac monitors, thermostats, and almost anything that we can imagine will be connected every day, everywhere and every time.

The idea of the IoT dates back to 1999 [115] and became a reality due to several technologies and protocols, including Machine to Machine (M2M), Wireless Sensor Networks (WSN), Radio frequency Identification (RFID), Internet Protocol version 6 (IPv6), IPv6 Low power Wireless Personal Area Networks (6LoWPAN), Routing Protocol for Low-Power and Lossy Networks (RPL), Constrained Application Protocol(CoAP), and Cognitive Radio that comes to empower the IoT revolution. This revolution gives birth to the proliferation of small devices such as sensors and actuators, with low consumption at lower cost, and a large number of platforms allowing users to develop their own applications. Thus, the number of connected « things » is increasing exponentially in the world (See Fig. 13).

B. Motivations behind the use of Cognitive Radio in the IoT

Cognitive Radio has a great potential to address several issues and challenges related to the deployment of IoT networks, namely, resource scarcity, interferences problem, limited communication range, purchasing license, storing and analyzing the data generated by IoT objects, heterogeneity, and reconfigurability and autonomy. Dealing with these issues presents the prime motivation behind the use of Cognitive Radio in IoT.

Regarding spectrum resources, it is difficult to allocate frequency bands to all the ever-increasing number of IoT objects. CR address this issue by enabling frequency reuse.

As for the interference problem, most IoT technologies such as RFID, IEEE 802.15.4 (ZigBee) operates in UHF and ISM frequency bands which are already saturated and can't accommodate more applications. Thus, sharing these bands between a large number of objects creates inevitable interferences. CR deal with this issue by enabling dynamic spectrum access to interference-free channels.

Concerning communication range, ISM unlicensed bands allow wireless technologies to operate only over a limited range. Acquiring a frequency spectrum, that ensures communications over long distances, requires the purchase of licenses which generates superfluous expenses. CR avoids the purchase of license and allows the opportunistic usage of unoccupied bands that enables long-distance communications.

In terms of storing and analyzing the data generated by IoT objects, the objects must find communication links to transmit the generated information to numerous servers (Cloud servers). CR is a proper solution to resolve this issue.

With respect to heterogeneity problems, IoT applications represent a wide range of design options (resources, deployment, connectivity, energy, communication modality, infrastructure, network size, network topology...). Thus, new communication paradigms should be designed to support this heterogeneity by providing environmental discovery, self-organization, and self-management capabilities [116]. CR is

³ <https://www.statista.com/statistics/668996/worldwide-expenditures-for-the-internet-of-things/>

⁴ <https://www.historyofinformation.com/detail.php?id=3411>

one of these paradigms that can be used to tackle heterogeneity issues.

Another reason is related to reconfigurability and autonomy. In fact, smart objects are expected to have the capacity to reconfigure themselves without any external action. Thus, objects need to be able to recognize and analyze their environment, detect neighboring objects, and then reconfigure themselves. CR proves to be a suitable solution to fit these circumstances.

Taking into account all of the above, it can be concluded that Cognitive Radio is genuinely a promising enabler technology for IoT.

C. IoT Applications and CR

The deployment of IoT products and services will be present in all sectors and industries, from the smart home to the smart city, education, health care, manufacturing, energy, utilities, commerce, transportation, monitoring, supply chain, and logistics as illustrated in Fig. 14. In general, the opportunities offered by IoT are unlimited and its full impact and potential will be realized in the near future as more and more devices connect to the Internet. At present, there are few works in the literature dealing with potential applications of CR technology for IoT, including military applications, cognitive radio-vehicular ad hoc networks (CR-VANET), emergency networks, smart grids, smart metering, and medical applications [117–123]. Table IV describes some IoT applications while demonstrating how CR can be used to address a number of their concerns.

D. Open Research Issues in CR-Based IoT

Several open research issues need to be addressed in order to take advantage of the full potential of Cognitive Radio for IoT networks. In this subsection, some of these issues will be discussed.

1) *Optimization of network resources:* Efficient spectrum utilization, in the context of a network composed of CR IoT nodes, requires the joint optimization of a set of parameters

such as transmission power, energy efficiency, transmission delay, and transmission data rate, which is a challenging task. The authors of [125] formulated an optimization problem to converge to an optimal solution, in a constrained environment, by taking into account three variables namely; transmission power, transmission rate, and transmission delay. To solve this problem, they applied a branch-and-cut polyhedral approach. The results show that an increase in network and packet size leads to an increase in transmission delay, power, rate, and interference.

Formulating and solving multi-objective optimization problems, considering a large number of factors, still need to be addressed.

2) *Energy efficiency:* One of the main challenges that need to be addressed in CR-based IoT networks is energy efficiency. IoT objects with cognitive radio capabilities consume more energy due to performing additional functions mainly spectrum sensing, which intensifies the power consumption issue, especially for energy-constrained nodes and battery-powered devices.

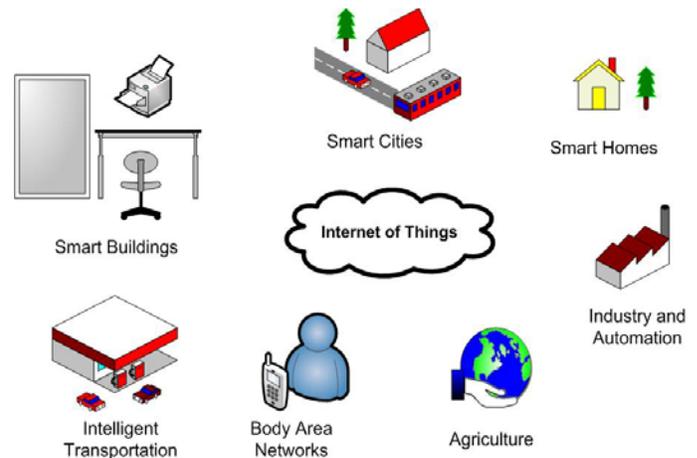


Fig. 14. Applications of IoT.

TABLE IV. CR CONTRIBUTION IN IoT APPLICATIONS

IoT application	Description and CR contribution
Smart Grids [120,121]	Smart Grid is an electricity distribution network that allows the flow of information between producers and consumers in order to control and regulate the flow of electricity in real time and enable more efficient management of the electricity grid. The major challenge of this technology is the transmission of data through long distances without investing in the purchase of licenses or the installation of cable trays. Cognitive radio offers a good solution to this problem.
Smart Homes [117]	Homes will be equipped with smart objects, allowing to carry out all daily task, such as smart lamps and smart fridges. These objects will integrate sensors endowed with cognitive capabilities that help to avoid interference in the ISM band.
Healthcare and Medical Applications [122,123]	Smart sensors are deployed to monitor critical data like blood pressure, glucose levels, and temperature. With Cognitive Radio this information can be transmitted to medical staff in real time, over long distances without worrying about spectrum availability.
Smart Cities [119]	The smart city is a paradigm of urban development which integrates information and communication technology (ICT) and IoT systems. In this paradigm, continuous connectivity should be maintained. CRNs can be a good solution to support this requirement.
Internet of Vehicles (IoV) [124]	Today, the trends toward less dependence on human beings have led us to the IoV paradigm in which vehicle control is achieved through the integration of communications and embedded systems. The IoV is supposed to be an autonomous travel decision-maker. Safe navigation may be possible in the future through vehicle-to-vehicle information exchange, vehicle-mounted sensors, and user intentions. The challenge in IoVs is the availability of spectrum for mobile vehicles. CRNs may be a good solution because of their long-range and interference-free spectrum sensing.

Among the solutions that have been introduced to deal with energy efficiency concerns are energy harvesting or scavenging and Cooperative Wireless Networks. Energy harvesting is defined as the process through which energy is extracted from external sources, such as wind and solar energies, and then stored so as to be used as a source of power [126]. Energy harvesting is considered as one of the most enabling technologies for Green Computing and real-world implementation of IoT. In [127] a differential game model has been proposed to resolve the resource allocation problem in cognitive WSN using energy harvesting. Optimal resource allocation strategies for all SUs are obtained by determining the loop Nash equilibrium and the feedback Nash equilibrium solutions of the proposed model. In [128], a CR-based energy harvesting approach has been used to extend the battery's lifetime for unmanned aerial vehicles (UAVs) under reliability and secrecy constraints. To reduce energy consumption, the energy harvested is maximized and the transmission energy consumption is minimized by solving a formulated optimization problem.

Currently, Cognitive Radio networks based on energy efficiency approaches are considered an important research direction.

3) *Security*: Addressing security issues in CR-Based IoT networks is a prime concern and a challenging task as most IoT objects are inherently heterogeneous and have their proper uniform security standards. These standards are not adequate for all heterogeneous networks. Several privacy and security aspects should be taken into account in designing IoT systems, including authentication, security assurance, and intrusion soft-ware [117]. Worthy efforts have been done to tackle security issues in some IoT applications with cognitive radio capabilities. For instance, the authors of [129] proposed a distributed cooperative spectrum sensing approach to deal with the security problem in CR-VANETs. This approach uses a weighted consensus-based spectrum sensing technique with trust assistance to ensure the reliability of spectrum sensing operation in a hostile CR-VANET. The validity of the proposed approach was approved by extensive simulations. In [128], the proposed CR-Based energy management scheme for UAVs has been designed in such a way that the secrecy and reliability of the system are ensured. This is illustrated by considering the scenario for the eavesdroppers.

Security remains an open issue to be addressed as cybersecurity attacks and threats are emerging and increasing in parallel with new technologies. Therefore, developing innovative countermeasures is demanded.

VI. CONCLUSION

Nowadays, wireless network devices are emerging at an unprecedented rate, which will further exacerbate spectrum shortages. Therefore, addressing the spectrum scarcity problem becomes more urgent. Cognitive Radio is a promising technology to overcome this issue for futuristic networks (IoT and 5G) by enabling an efficient, flexible, and opportunistic usage of the scarce frequency spectrum. This paper provides a review of Cognitive Radio with emphasis is put on SDR,

spectrum sensing, and CR-Based IoT. Concerning SDR, this article highlights its architecture, tools, advantages, and some of its ongoing challenges. As for spectrum sensing, it is extensively reviewed: spectrum opportunities considering multiple dimensions are discussed, classification of spectrum sensing based on several aspects is presented, Narrowband and Wideband sensing techniques are analyzed, ML-based spectrum sensing is studied, and challenges and open research directions are reviewed. Regarding CR-Based IoT, the focus was on explaining the motivations behind the use of CR in IoT networks and reviewing some of the open issues related to CR-Based IoT.

REFERENCES

- [1] Ranjan, A.; Anurag; Singh, B. Design and analysis of spectrum sensing in cognitive radio based on energy detection. In Proceedings of the 2016 International Conference on Signal and Information Processing, IConSIP 2016; Institute of Electrical and Electronics Engineers Inc., 2017.
- [2] Spectrum Policy Task Force | Federal Communications Commission Available online: <https://www.fcc.gov/document/spectrum-policy-task-force> (accessed on Feb 4, 2021).
- [3] Akyildiz, I.F.; Lee, W.Y.; Vuran, M.C.; Mohanty, S. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Comput. Networks* 2006, 50, 2127–2159, doi:10.1016/j.comnet.2006.05.001.
- [4] Mitola, J.; Maguire, G.Q. Cognitive radio: making software radios more personal. *IEEE Pers. Commun.* 1999, 6, 13–18, doi:10.1109/98.788210.
- [5] Jondral, F.K. Software-defined radio - Basics and evolution to cognitive radio. *Eurasip J. Wirel. Commun. Netw.* 2005, 2005, 275–283, doi:10.1155/WCN.2005.275.
- [6] Karatalay, O.; Erkuçuk, S.; Baykaş, T. Busy tone based coexistence algorithm for WRAN and WLAN systems in TV white space. *IET Commun.* 2018, 12, 1630–1637, doi:10.1049/iet-com.2017.0740.
- [7] Federal Communication Commission Available online: https://apps.fcc.gov/edocs_public/attachmatch/FCC-08-260A1.pdf%22 (accessed on Feb 4, 2021).
- [8] Federal Communications Commission. ET Docket 10-174: Second Memorandum Opinion and Order in the Matter of Unlicensed Operation in the TV Broadcast Bands. Active Regulation; FCC: Washington, DC, USA, 2012.
- [9] Federal Communications Commission. Third Memorandum Opinion and Order in the Matter of Unlicensed Operation in the TV Broadcast Bands (ET Docket No. 04-186) and Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band (ET Docket No. 04-186).
- [10] Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies | Federal Communications Commission;
- [11] Haykin, S. Cognitive radio: Brain-empowered wireless communications. *IEEE J. Sel. Areas Commun.* 2005, 23, 201–220, doi:10.1109/JSAC.2004.839380.
- [12] Ren, P.; Wang, Y.; Du, Q.; Xu, J. A survey on dynamic spectrum access protocols for distributed cognitive wireless networks. *Eurasip J. Wirel. Commun. Netw.* 2012, 2012, 1–21, doi:10.1186/1687-1499-2012-60.
- [13] Zhang, N.; Liang, H.; Cheng, N.; Tang, Y.; Mark, J.W.; Shen, X.S. Dynamic spectrum access in multi-channel cognitive radio networks. *IEEE J. Sel. Areas Commun.* 2014, 32, 2053–2064, doi:10.1109/JSAC.2014.141109.
- [14] Kumar, A.; Kumar, K. Multiple access schemes for Cognitive Radio networks: A survey. *Phys. Commun.* 2020, 38, 100953, doi:10.1016/j.phycom.2019.100953.
- [15] Awan, F.G.; Sheikh, N.M.; Hanif, M.F. Information Theory of Cognitive Radio System. In *Cognitive Radio Systems*; InTech, 2009.
- [16] Mitola(III), J. Cognitive Radio An Integrated Agent Architecture for Software Defined Radio, Doctor of Technology, Royal Inst. Technol. (KTH), Stockholm, Sweden, 2000.

- [17] Mittal, S. A survey of techniques for improving energy efficiency in embedded computing systems. *Int. J. Comput. Aided Eng. Technol.* 2014, 6, 440–459, doi:10.1504/IJCAET.2014.065419.
- [18] Huang, S.; Liu, X.; Ding, Z. Opportunistic Spectrum Access in Cognitive Radio Networks.; Institute of Electrical and Electronics Engineers (IEEE), 2008; pp. 1427–1435.
- [19] Goldsmith, A.; Jafar, S.A.; Maric, I.; Srinivasa, S. Breaking spectrum gridlock with cognitive radios: An information theoretic perspective. *Proc. IEEE* 2009, 97, 894–914, doi:10.1109/JPROC.2009.2015717.
- [20] Gupta, M.S.; Kumar, K. Progression on spectrum sensing for cognitive radio networks: A survey, classification, challenges and future research issues. *J. Netw. Comput. Appl.* 2019, 143, 47–76, doi:10.1016/j.jnca.2019.06.005.
- [21] Wang, B.; Liu, K.J.R. Advances in cognitive radio networks: A survey. *IEEE J. Sel. Top. Signal Process.* 2011, 5, 5–23, doi:10.1109/JSTSP.2010.2093210.
- [22] Liang, Y.C.; Chen, K.C.; Li, G.Y.; Mähönen, P. Cognitive radio networking and communications: An overview. *IEEE Trans. Veh. Technol.* 2011, 60, 3386–3407, doi:10.1109/TVT.2011.2158673.
- [23] De Domenico, A.; Calvanese Strinati, E.; Di Benedetto, M.G. A survey on MAC strategies for cognitive radio networks. *IEEE Commun. Surv. Tutorials* 2012, 14, 21–44, doi:10.1109/SURV.2011.111510.00108.
- [24] Mansoor, N.; Islam, A.K.M.M.; Zareei, M.; Vargas-Rosales, C. RARE: A Spectrum Aware Cross-Layer MAC Protocol for Cognitive Radio Ad-Hoc Networks. *IEEE Access* 2018, 6, 22210–22227, doi:10.1109/ACCESS.2018.2807781.
- [25] Che-Aron, Z.; Abdalla, A.H.; Abdullah, K.; Hassan, W.H.; Rahman, M.D.A. RACARP: A Robustness Aware routing protocol for Cognitive radio Ad Hoc Networks. *J. Theor. Appl. Inf. Technol.* 2015, 76, 246–257.
- [26] Elrhareg, H.; Ridouani, M.; Hayar, A. Routing protocols on cognitive radio networks: Survey. In *Proceedings of the 5th IEEE International Smart Cities Conference, ISC2 2019*; Institute of Electrical and Electronics Engineers Inc., 2019; pp. 296–302.
- [27] Stevenson, C.R.; Chouinard, G.; Lei, Z.; Hu, W.; Shellhammer, S.J.; Caldwell, W. IEEE 802.22: The First Cognitive Radio Wireless Regional Area Network Standard. *IEEE Commun. Mag.* 2009, 47, 130–138, doi:10.1109/MCOM.2009.4752688.
- [28] Tönjes, R.; Moessner, K.; Lohmar, T.; Wolf, M. OverDRIVE Spectrum Efficient Multicast Services to Vehicles. undefined 2002.
- [29] Xu, L.; Tönjes, R.; Paila, T.; Hansmann, W.; Frank, M.; Albrecht, M. DRiVE-ing to the internet: Dynamic Radio for IP Services in vehicular environments. *Conf. Local Comput. Networks* 2000, 281–289, doi:10.1109/LCN.2000.891040.
- [30] Pham, P.P. Comprehensive analysis of the IEEE 802.11. *Mob. Networks Appl.* 2005, 10, 691–703, doi:10.1007/s11036-005-3363-x.
- [31] Mitola, J. Software radios-survey, critical evaluation and future directions. In *Proceedings of the Proceedings - National Telesystems Conference, NTC 1992*; Institute of Electrical and Electronics Engineers Inc., 1992.
- [32] Robert, M.; Fette, B.A. The Software-Defined Radio as a Platform for Cognitive Radio. In *Cognitive Radio Technology*; Elsevier Inc., 2009; pp. 65–103 ISBN 9780123745354.
- [33] Haghight, A. A review on essentials and technical challenges of software defined radio. In *Proceedings of the Proceedings - IEEE Military Communications Conference MILCOM; 2002; Vol. 1*, pp. 377–382.
- [34] Sinha, D.; Verma, A.K.; Kumar, S. Software defined radio: Operation, challenges and possible solutions. *Proc. 10th Int. Conf. Intell. Syst. Control. ISCO 2016* 2016, doi:10.1109/ISCO.2016.7727079.
- [35] Akeela, R.; Dezfouli, B. Software-defined Radios: Architecture, state-of-the-art, and challenges. *Comput. Commun.* 2018, 128, 106–125, doi:10.1016/j.comcom.2018.07.012.
- [36] Ahn, C.; Kim, J.; Ju, J.; Choi, J.; Choi, B.; Choi, S. Implementation of an SDR platform using GPU and its application to a 2×2 MIMO WiMAX system. *Analog Integr. Circuits Signal Process.* 2011, 69, 107–117, doi:10.1007/s10470-011-9764-9.
- [37] Ulversoy, T. Software defined radio: Challenges and opportunities. *IEEE Commun. Surv. Tutorials* 2010, 12, 531–550, doi:10.1109/SURV.2010.032910.00019.
- [38] Vestias, M.; Neto, H. Trends of CPU, GPU and FPGA for high-performance computing. In *Proceedings of the Conference Digest - 24th International Conference on Field Programmable Logic and Applications, FPL 2014*; Institute of Electrical and Electronics Engineers Inc., 2014.
- [39] Frantz, G. Digital signal processor trends. *IEEE Micro* 2000, 20, 52–59, doi:10.1109/40.888703.
- [40] Tong, Z.; Arifianto, M.S.; Liao, C.F. Wireless transmission using universal software radio peripheral. In *Proceedings of the 2009 International Conference on Space Science and Communication, IconSpace - Proceedings; 2009*; pp. 19–23.
- [41] Zitouni, R. Software defined radio for cognitive wireless sensor networks: A reconfigurable IEEE 802.15.4 Standard, Paris Est, 2015.
- [42] Ettus, M.; Braun, M. The Universal Software Radio Peripheral (USRP) Family of Low-Cost SDRs. In *Opportunistic Spectrum Sharing and White Space Access: The Practical Reality*; John Wiley & Sons, Inc: Hoboken, NJ, 2015; pp. 3–23 ISBN 9781119057246.
- [43] Michael, L.B.; Mihaljevic, M.J.; Haruyama, S.; Kohno, R. A framework for secure download for software-defined radio. *IEEE Commun. Mag.* 2002, 40, 88–96, doi:10.1109/MCOM.2002.1018012.
- [44] Baldini, G.; Sturman, T.; Biswas, A.R.; Leschhorn, R.; Gódor, G.; Street, M. Security aspects in software defined radio and cognitive radio networks: A survey and a way ahead. *IEEE Commun. Surv. Tutorials* 2012, 14, 355–379.
- [45] Ostovar, A.; Zikria, Y. Bin; Kim, H.S.; Ali, R. Optimization of Resource Allocation Model with Energy-Efficient Cooperative Sensing in Green Cognitive Radio Networks. *IEEE Access* 2020, 8, 141594–141610, doi:10.1109/ACCESS.2020.3013034.
- [46] Kristensen, J.; Fitzek, F.; Koch, P.; Prasad, R. Conceptual Considerations for Reducing the Computational Complexity in Software Defined Radio using Cooperative Wireless Networks. undefined 2005.
- [47] Hall, P.S.; Gardner, P.; Faraone, A. Antenna requirements for software defined and cognitive radios. In *Proceedings of the Proceedings of the IEEE; Institute of Electrical and Electronics Engineers Inc., 2012; Vol. 100*, pp. 2262–2270.
- [48] P. Kolodzy, et al., “Next Generation Communications Kickoff Meeting,” *Proceedings of Defense Advanced Research Projects Agency, Arlington, 17 October 2001.*
- [49] Yücek, T.; Arslan, H. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Commun. Surv. Tutorials* 2009, 11, 116–130, doi:10.1109/SURV.2009.090109.
- [50] Gavrilovska, L.; Denkovski, D.; Rakovic, V.; Angelichinoski, M. Medium access control protocols in cognitive radio networks: Overview and general classification; 2014; Vol. 16; ISBN 9783319017181.
- [51] Akyildiz, I.F.; Lee, W.Y.; Chowdhury, K.R. CRAHNS: Cognitive radio ad hoc networks. *Ad Hoc Networks* 2009, 7, 810–836, doi:10.1016/j.adhoc.2009.01.001.
- [52] Kim, H.; Shin, K.G. In-band spectrum sensing in cognitive radio networks: Energy detection or feature detection? In *Proceedings of the Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM; ACM Press: New York, New York, USA, 2008*; pp. 14–25.
- [53] Akyildiz, I.F.; Lo, B.F.; Balakrishnan, R. Cooperative spectrum sensing in cognitive radio networks: A survey. *Phys. Commun.* 2011, 4, 40–62.
- [54] Li, Z.; Yu, F.R.; Huang, M. Distributed spectrum sensing in cognitive radio networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference, WCNC; 2009*.
- [55] Teguig, D.; Scheers, B.; Le Nir, V. Data fusion schemes for cooperative spectrum sensing in cognitive radio networks. In *Proceedings of the 2012 Military Communications and Information Systems Conference, MCC 2012; 2012*; pp. 104–110.
- [56] Kyperountas, S.; Correal, N.; Shi, Q. A Comparison of Fusion Rules for Cooperative Spectrum Sensing in Fading Channels. *Wireless.Vt.Edu* 2010.

- [57] Wild, B.; Ramchandran, K. Detecting primary receivers for cognitive radio applications. In Proceedings of the 2005 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005; 2005; pp. 124–130.
- [58] Rawat, D.; Yan, G. Spectrum sensing methods and dynamic spectrum sharing in cognitive radio networks: A survey. *Int. J. Res. Rev. Wirel. Sens. Networks* 2011, 1, 1–13.
- [59] Arjoune, Y.; Kaabouch, N. A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions. *Sensors (Switzerland)* 2019, 19, doi:10.3390/s19010126.
- [60] Garg, R.; Nitin, D. Current Trends and Research Challenges in Spectrum-Sensing for Cognitive Radios. *Int. J. Adv. Comput. Sci. Appl.* 2016, 7, 402–408, doi:10.14569/ijacsa.2016.070756.
- [61] Webster, J.G.; Kusaladharna, S.; Tellambura, C. An Overview of Cognitive Radio Networks. *Wiley Encycl. Electr. Electron. Eng.* 2017, 1–17, doi:10.1002/047134608x.w8355.
- [62] Ruan, L.; Li, Y.; Cheng, W.; Wu, Z. A robust threshold optimization approach for energy detection based spectrum sensing with noise uncertainty. In Proceedings of the Proceedings of the 2015 10th IEEE Conference on Industrial Electronics and Applications, ICIEA 2015; Institute of Electrical and Electronics Engineers Inc., 2015; pp. 161–165.
- [63] Margoosian, A.; Abouei, J.; Plataniotis, K.N. An Accurate Kernelized Energy Detection in Gaussian and non-Gaussian/Impulsive Noises. *IEEE Trans. Signal Process.* 2015, 63, 5621–5636, doi:10.1109/TSP.2015.2457400.
- [64] Atapattu, S.; Tellambura, C.; Jiang, H. Energy Detection for Spectrum Sensing in Cognitive Radio; SpringerBriefs in Computer Science; Springer New York: New York, NY, 2014; ISBN 978-1-4939-0493-8.
- [65] Lorincz, J.; Ramljak, I.; Begušić, D. A review of the noise uncertainty impact on energy detection with different OFDM system designs. *Comput. Commun.* 2019, 148, 185–207, doi:10.1016/j.comcom.2019.09.013.
- [66] Tandra, R.; Sahai, A. SNR walls for signal detection. *IEEE J. Sel. Top. Signal Process.* 2008, 2, 4–17, doi:10.1109/JSTSP.2007.914879.
- [67] Farag, H.M.; Ehab, M. An efficient dynamic thresholds energy detection technique for Cognitive Radio spectrum sensing. In Proceedings of the 2014 10th International Computer Engineering Conference: Today Information Society What's Next?, ICENCO 2014; Institute of Electrical and Electronics Engineers Inc., 2015; pp. 139–144.
- [68] Arjoune, Y.; Mrabet, Z. El; Ghazi, H. El; Tamtaoui, A. Spectrum sensing: Enhanced energy detection technique based on noise measurement. 2018 IEEE 8th Annu. Comput. Commun. Work. Conf. CCWC 2018 2018, 2018-Janua, 828–834, doi:10.1109/CCWC.2018.8301619.
- [69] Nallagonda, S.; Roy, S.D.; Kundu, S.; Ferrari, G.; Raheli, R. Censoring-Based Cooperative Spectrum Sensing with Improved Energy Detectors and Multiple Antennas in Fading Channels. *IEEE Trans. Aerosp. Electron. Syst.* 2018, 54, 537–553, doi:10.1109/TAES.2017.2732798.
- [70] Kockaya, K.; Develi, I. Spectrum sensing in cognitive radio networks: threshold optimization and analysis. *Eurasip J. Wirel. Commun. Netw.* 2020, 2020, 1–20, doi:10.1186/s13638-020-01870-7.
- [71] Lv, Q.; Gao, F. Matched filter based spectrum sensing and power level recognition with multiple antennas. In Proceedings of the 2015 IEEE China Summit and International Conference on Signal and Information Processing, ChinaSIP 2015 - Proceedings; Institute of Electrical and Electronics Engineers Inc., 2015; pp. 305–309.
- [72] Zhang, X.; Chai, R.; Gao, F. Matched filter based spectrum sensing and power level detection for cognitive radio network. In Proceedings of the 2014 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2014; Institute of Electrical and Electronics Engineers Inc., 2014; pp. 1267–1270.
- [73] Kapoor, S.; Rao, S.V.R.K.; Singh, G. Opportunistic spectrum sensing by employing matched filter in cognitive radio network. *Proc. - 2011 Int. Conf. Commun. Syst. Netw. Technol. CSNT 2011* 2011, 580–583, doi:10.1109/CSNT.2011.124.
- [74] Semba Yawada, P.; Wei, A.J. Cyclostationary Detection Based on Non-cooperative spectrum sensing in cognitive radio network. In Proceedings of the 6th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, IEEE-CYBER 2016; Institute of Electrical and Electronics Engineers Inc., 2016; pp. 184–187.
- [75] Kumar, A.; NandhaKumar, P. OFDM system with cyclostationary feature detection spectrum sensing. *ICT Express* 2019, 5, 21–25, doi:10.1016/j.icte.2018.01.007.
- [76] Alias, D.M.; Ragesh, G.K. Cognitive Radio networks: A survey. In Proceedings of the Proceedings of the 2016 IEEE International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2016; Presses Polytechniques Et Universitaires Romandes, 2016; pp. 1981–1986.
- [77] Darabkh, K.A.; Amro, O.M.; Bany Salameh, H.; Al-Zubi, R.T. A-Z overview of the in-band full-duplex cognitive radio networks. *Comput. Commun.* 2019, 145, 66–95, doi:10.1016/j.comcom.2019.06.007.
- [78] Lundén, J.; Koivunen, V.; Huttunen, A.; Poor, H.V. Spectrum sensing in cognitive radios based on multiple cyclic frequencies. In Proceedings of the Proceedings of the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom; 2007; pp. 37–43.
- [79] Zhu, Y.; Su, K.; Wang, D.; Zhu, M.; Li, Y. Analysis of improved cyclostationary detector with SLC diversity over Nakagami-m fading channels. In Proceedings of the 2013 8th International ICST Conference on Communications and Networking in China, CHINACOM 2013 - Proceedings; 2013; pp. 133–137.
- [80] Yan, T.; Xu, F.; Wei, N.; Yang, Z. An improved cyclostationary feature detection algorithm. In Proceedings of the Advances in Intelligent Systems and Computing; Springer Verlag, 2020; Vol. 905, pp. 544–555.
- [81] Zeng, Y.; Liang, Y.C. Spectrum-sensing algorithms for cognitive radio based on statistical covariances. *IEEE Trans. Veh. Technol.* 2009, 58, 1804–1815, doi:10.1109/TVT.2008.2005267.
- [82] Zeng, Y.; Liang, Y.C. Eigenvalue-based spectrum sensing algorithms for cognitive radio. *IEEE Trans. Commun.* 2009, 57, 1784–1793, doi:10.1109/TCOMM.2009.06.070402.
- [83] Geirhofer, S.; Tong, L.; Sadler, B.M. A measurement-based model for dynamic spectrum access in WLAN channels. In Proceedings of the Proceedings - IEEE Military Communications Conference MILCOM; Institute of Electrical and Electronics Engineers Inc., 2006.
- [84] Mishra, S.M.; Ten Brink, S.; Mahadevappa, R.; Brodersen, R.W. Cognitive technology for ultra-wideband/WiMax coexistence. In Proceedings of the 2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks; IEEE Computer Society, 2007; pp. 179–186.
- [85] Tang, H. Some physical layer issues of wide-band cognitive radio systems. In Proceedings of the 2005 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005; 2005; pp. 151–159.
- [86] Afifi, W.; Krunz, M. Adaptive transmission-reception-sensing strategy for cognitive radios with full-duplex capabilities. In Proceedings of the 2014 IEEE International Symposium on Dynamic Spectrum Access Networks, DYSPAN 2014; IEEE Computer Society, 2014; pp. 149–160.
- [87] Tian, Z.; Giannakis, G.B. A wavelet approach to wideband spectrum sensing for cognitive radios. In Proceedings of the 1st International Conference on Cognitive Radio Oriented Wireless Networks and Communications 2006, CROWNCOM; IEEE Computer Society, 2006.
- [88] Farooq, S.Z.; Ghafoor, A. Analysis of multiband sensing-time joint detection framework for cognitive radio systems. *IEEE Veh. Technol. Conf.* 2012, 0–4, doi:10.1109/VTCFall.2012.6399359.
- [89] Farhang-Boroujeny, B. Filter bank spectrum sensing for cognitive radios. *IEEE Trans. Signal Process.* 2008, 56, 1801–1811, doi:10.1109/TSP.2007.911490.
- [90] Sun, H.; Nallanathan, A.; Wang, C.X.; Chen, Y. Wideband spectrum sensing for cognitive radio networks: A survey. *IEEE Wirel. Commun.* 2013, 20, 74–81, doi:10.1109/MWC.2013.6507397.
- [91] Tian, Z.; Giannakis, G.B. Compressed sensing for wideband cognitive radios. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* 2007, 4, 1357–1360, doi:10.1109/ICASSP.2007.367330.
- [92] Li, Z.; Chang, B.; Wang, S.; Liu, A.; Zeng, F.; Luo, G. Dynamic Compressive Wide-Band Spectrum Sensing Based on Channel Energy Reconstruction in Cognitive Internet of Things. *IEEE Trans. Ind. Informatics* 2018, 14, 2598–2607, doi:10.1109/TII.2018.2797096.

- [93] Tian, Z.; Tafesse, Y.; Sadler, B.M. Cyclic feature detection with subnyquist sampling for wideband spectrum sensing. *IEEE J. Sel. Top. Signal Process.* 2012, 6, 58–69, doi:10.1109/JSTSP.2011.2181940.
- [94] Bkassiny, M.; Li, Y.; Jayaweera, S.K. A survey on machine-learning techniques in cognitive radios. *IEEE Commun. Surv. Tutorials* 2013, 15, 1136–1159, doi:10.1109/SURV.2012.100412.00017.
- [95] Azmat, F.; Chen, Y.; Stocks, N. Analysis of spectrum occupancy using machine learning algorithms. *IEEE Trans. Veh. Technol.* 2016, 65, 6853–6860, doi:10.1109/TVT.2015.2487047.
- [96] Shenfield, A.; Khan, Z.; Ahmadi, H. Deep Learning Meets Cognitive Radio: Predicting Future Steps. *IEEE Veh. Technol. Conf.* 2020, 2020-May, doi:10.1109/VTC2020-Spring48590.2020.9129042.
- [97] Arjoun, Y.; Kaabouch, N. On spectrum sensing, a machine learning method for cognitive radio systems. *IEEE Int. Conf. Electro Inf. Technol.* 2019, 333–338, doi:10.1109/EIT.2019.8834099.
- [98] Lu, Y.; Zhu, P.; Wang, D.; Fattouche, M. Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference, WCNC; Institute of Electrical and Electronics Engineers Inc.*, 2016; Vol. 2016-September.
- [99] Khalfi, B.; Zaid, A.; Hamdaoui, B. When machine learning meets compressive sampling for wideband spectrum sensing. In *Proceedings of the 2017 13th International Wireless Communications and Mobile Computing Conference, IWCMC 2017; Institute of Electrical and Electronics Engineers Inc.*, 2017; pp. 1120–1125.
- [100] Obite, F.; Yusof, K.M.; Din, J. A Mathematical Approach for Hidden Node Problem in Cognitive Radio Networks. *TELKOMNIKA* 2017, 15, 1693–6930, doi:10.12928/TELKOMNIKA.v15i3.6897.
- [101] Paisana, F.; Marchetti, N.; Dasilva, L.A. Radar, TV and cellular bands: Which spectrum access techniques for which bands? *IEEE Commun. Surv. Tutorials* 2014, 16, 1193–1220, doi:10.1109/SURV.2014.031914.00078.
- [102] Jiang, C.; Beaulieu, N.C.; Jiang, C. A novel asynchronous cooperative spectrum sensing scheme. In *Proceedings of the IEEE International Conference on Communications; Institute of Electrical and Electronics Engineers Inc.*, 2013; pp. 2606–2611.
- [103] Cao, K.T.; Wang, D.L. Sensor node-assisted asynchronous cooperative spectrum sensing for cognitive radio network. *J. China Univ. Posts Telecommun.* 2014, 21, 17–23, doi:10.1016/S1005-8885(14)60325-3.
- [104] Zhou, X.; Ma, J.; Li, G.Y.; Young, H.K.; Soong, A.C.K. Probability-based combination for cooperative spectrum sensing in cognitive radio networks. In *Proceedings of the IEEE International Conference on Communications; 2009*.
- [105] Shen, L.; Wang, H.; Zhang, W.; Zhao, Z. Blind spectrum sensing for cognitive radio channels with noise uncertainty. *IEEE Trans. Wirel. Commun.* 2011, 10, 1721–1724, doi:10.1109/TWC.2011.040511.101559.
- [106] Nazzal, M.; Aygül, M.A.; Arslan, H. Estimating sparsity level for enabling compressive sensing of wireless channels and spectra in 5G and beyond. *arXiv* 2020, 1–11.
- [107] Hu, N.; Yao, Y.D.; Mitola, J. Most active band (MAB) attack and countermeasures in a cognitive radio network. *IEEE Trans. Wirel. Commun.* 2012, 11, 898–902, doi:10.1109/TWC.2012.011812.110927.
- [108] Chen, R.; Park, J.M. Ensuring trustworthy spectrum sensing in cognitive radio networks. In *Proceedings of the 2006 1st IEEE Workshop on Networking Technologies for Software Defined Radio Networks, SDR; 2006; pp. 110–119*.
- [109] Salahdine, F.; Kaabouch, N. Security threats, detection, and countermeasures for physical layer in cognitive radio networks: A survey. *Phys. Commun.* 2020, 39, 101001, doi:10.1016/j.phycom.2020.101001.
- [110] Rawat, A.S.; Anand, P.; Chen, H.; Varshney, P.K. Countering Byzantine attacks in cognitive radio networks. In *Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings; 2010; pp. 3098–3101*.
- [111] Statista Internet of Things spending worldwide 2023 | Statista Available online: <https://www.statista.com/statistics/668996/worldwide-expenditures-for-the-internet-of-things/> (accessed on May 9, 2021).
- [112] Palattella, M.R. Internet of Things in the 5G Era: Enablers, Architecture and Business Models. *IEEE J. Sel. Areas Commun.* 2015, 1–9.
- [113] Want, R.; Schilit, B.N.; Jenson, S. Enabling the internet of things. *Computer (Long. Beach. Calif.)* 2015, 48, 28–35, doi:10.1109/MC.2015.12.
- [114] Al-Fuqaha, A.; Guizani, M.; Mohammadi, M.; Aledhari, M.; Ayyash, M. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Commun. Surv. Tutorials* 2015, 17, 2347–2376, doi:10.1109/COMST.2015.2444095.
- [115] Kevin, A. That ' Internet of Things ' Thing. *RFiD J.* 2010, 4986.
- [116] Rawat, P.; Singh, K.D.; Chaouchi, H.; Bonnin, J.M. Wireless sensor networks: A survey on recent developments and potential synergies. *J. Supercomput.* 2014, 68, 1–48, doi:10.1007/s11227-013-1021-9.
- [117] Awin, F.A.; Alginahi, Y.M.; Abdel-Raheem, E.; Tepe, K. Technical Issues on Cognitive Radio-Based Internet of Things Systems: A Survey. *IEEE Access* 2019, 7, 97887–97908, doi:10.1109/ACCESS.2019.2929915.
- [118] Singh, K.D.; Rawat, P.; Bonnin, J.M. Cognitive radio for vehicular ad hoc networks (CR-VANETs): Approaches and challenges. *Eurasip J. Wirel. Commun. Netw.* 2014, 2014, 1–22, doi:10.1186/1687-1499-2014-49.
- [119] Somov, A.; Dupont, C.; Giaffreda, R. Supporting smart-city mobility with cognitive internet of things. *2013 Futur. Netw. Mob. Summit, Futur.* 2013 2013, 1–10.
- [120] Cacciapuoti, A.S.; Caleffi, M.; Paura, L. On the probabilistic deployment of smart grid networks in TV white space. *Sensors (Switzerland)* 2016, 16, 671, doi:10.3390/s16050671.
- [121] Yang, Z.; Shi, Z.; Jin, C. SACRB-MAC: A high-capacity MAC protocol for cognitive radio sensor networks in smart grid. *Sensors (Switzerland)* 2016, 16, 464, doi:10.3390/s16040464.
- [122] Ahad, A.; Tahir, M.; Sheikh, M.A.; Ahmed, K.I.; Mughees, A.; Numani, A. Technologies trend towards 5g network for smart health-care using iot: A review. *Sensors (Switzerland)* 2020, 20, 1–22, doi:10.3390/s20144047.
- [123] Le, T.T.T.; Moh, S. Energy-efficient protocol of link scheduling in cognitive radio body area networks for medical and healthcare applications. *Sensors (Switzerland)* 2020, 20, 16–18, doi:10.3390/s20051355.
- [124] Wu, Q.; Ding, G.; Xu, Y.; Feng, S.; Du, Z.; Wang, J.; Long, K. Cognitive internet of things: A new paradigm beyond connection. *IEEE Internet Things J.* 2014, 1, 129–143, doi:10.1109/JIOT.2014.2311513.
- [125] Muwonge, B.S.; Pei, T.; Otim, J.S.; Mayambala, F. A joint power, delay and rate optimization model for secondary users in cognitive radio sensor networks. *Sensors (Switzerland)* 2020, 20, 1–18, doi:10.3390/s20174907.
- [126] Tentzeris, M.M.; Georgiadis, A.; Roselli, L. Energy harvesting and scavenging. *Proc. IEEE* 2014, 102, 1644–1648.
- [127] Xu, H.; Gao, H.; Zhou, C.; Duan, R.; Zhou, X. Resource allocation in cognitive radio wireless sensor networks with energy harvesting. *Sensors (Switzerland)* 2019, 19, 5115, doi:10.3390/s19235115.
- [128] Khalid, W.; Yu, H.; Noh, S. Residual energy analysis in cognitive radios with energy harvesting uav under reliability and secrecy constraints. *Sensors (Switzerland)* 2020, 20, 2998, doi:10.3390/s20102998.
- [129] Wei, Z.; Yu, F.R.; Boukerche, A. Cooperative spectrum sensing with trust assistance for cognitive radio vehicular ad hoc networks. In *Proceedings of the DIVANet 2015 - Proceedings of the 5th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications; Association for Computing Machinery, Inc.* 2015; pp. 27–33.

Artificial Intelligence based Recommendation System for Analyzing Social Business Reviews

Asma Alanazi¹, Marwan Alseid²
Department of Computer Science
Fahad Bin Sultan University
Tabuk City, Kingdom of Saudi Arabia

Abstract—Recently, analysing reviews presented by clients to products that are provided by e-commerce companies, such as Amazon, to produce efficient recommendations has been receiving a lot of attention. However, ensuring and generating effective recommendations on time is a challenge. This research paper proposes an artificial intelligence-based system. The proposed system uses the Incremental Learning - based Method (ILbM) to learn a neural network classifier. The ILbM uses the bagging technique in the process of training the classifier. To ensure a high degree of performance, the ILbM is implemented on the Hadoop since it allows the execution in parallel. Compared to a similar system, the proposed system shows better results in terms of accuracy (97.5%), precision (95.7%), recall (91.5%), and time of response (36 seconds).

Keywords—ILbM; reviews; classifier; text analysing; training bagging; MapReduce; big data

I. INTRODUCTION

As technology continues to evolve in most areas of life, data is being generated daily and growing rapidly. Therefore, we are in the big data era in which data sets are too large and complex.

A. Big Data and Its Properties

To consider data as big data it must contain some characteristics including; volume, variety and velocity [1]. A summarization of these characteristics is illustrated in Fig. 1.

B. Big Data and Its Relationship with Business Intelligence and Analytics

Business intelligence and analytics (BI&A) and the field of big data analytics have become increasingly discussed and studied in both the academic and the business communities over the past two decades. Industry studies have highlighted this significant development [2]. As a strong proof of this, Amazon Company is considered as an important source for big data related to transaction details or related to the customers. In addition, Amazon Company has employed Artificial Intelligence (AI) to analyze the data for increasing gain purpose [3]. The big data generated by the systems and used in Amazon (or any other online shopping companies) forms the base of applying BI&A to improve the work of companies in terms of artificial intelligence.

C. Problem Statement

Since financial transactions are considered sensitive in companies, employing artificial intelligence to enhance the

amount of gain is critical. The reason behind this is that when using wrong (or inaccurate) BI&A on the data base, this leads to poor or wrong decisions' making, and consequently leads to financial losses. This problem is strongly highlighted when it comes to analyse and mine text. Fig. 2 illustrates the problem of having poor BI&A.

As shown in Fig. 2, the failure in the implementation step requires returning to the early steps. This in turn reflects poor accuracy, which is the main issue in this domain.

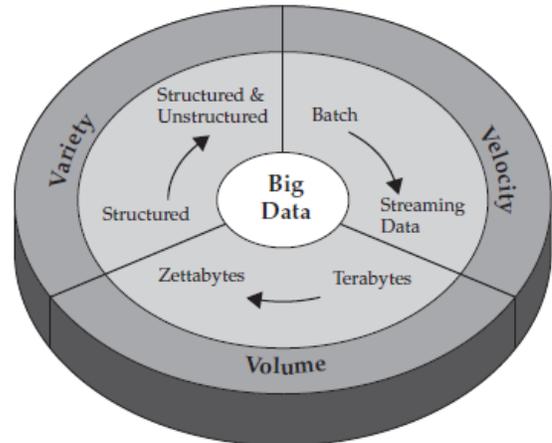


Fig. 1. IBM Characterizes of Big Data by its Volume, Velocity, and Variety [1].

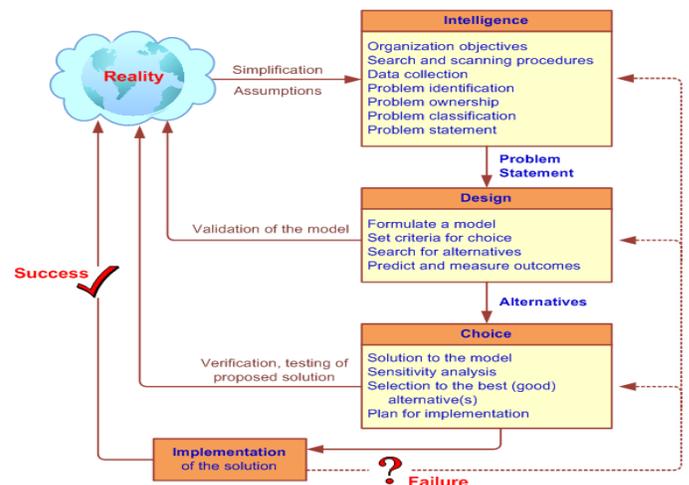


Fig. 2. General Steps in BI&A in Terms of Artificial Intelligence [4].

Besides the accuracy issue, providing decisions in the correct time is also critical. In other words, making decisions (such as providing good recommendations about a certain product for clients) taking into account the time dimension is a sensitive issue. That is because a delay in making the decision and providing it to the customer leads to change his/her mind and may direct him/her to another competitor [5]. For this reason, high performance is pressing in business domain when depending on artificial intelligence. On the other hand, dealing with big data in the business domain requires solving the performance issue [6].

D. Research Questions

Accuracy and performance issues explained above lead to setting and answering the following research questions:

- 1) How can artificial intelligence-based business systems provide satisfied decisions to customers in terms of accuracy?
- 2) How can artificial intelligence-based business systems provide satisfactory decisions on time, taking into consideration dealing with and analyzing big data effectively?

E. Contribution

In general, the contributions of this paper can be presented as follows:

- In responding to the first research question, this paper proposes a novel artificial intelligence system to provide accurate decisions. The system uses a novel algorithm called Incremental Learning based Method (ILbM). The key idea of the ILbM algorithm depends on incremental learning using the bagging method.
- For answering the second research questions, this paper employs Hadoop to perform the training phase in parallel [53]. This ensures fast time decision making in regard about providing recommendations based on using MapReduce technique.
- Extensive experiments are conducted to show the effectiveness of the proposed methods and compare them with similar approaches in term of Performance-based metric discussion and AI-based metrics discussion.

F. Paper Organization

The rest of this work is organized as follows. Section II reviews the related research works. Section III provides the proposed artificial intelligence system. In Section IV, the used metrics are presented for evaluation purposes. Section V presents the experiments and analyses and discusses the results in light of a comparison with similar approaches. Finally, the conclusion of the paper is discussed in Section VI.

II. RELATED WORK

This section provides a brief background regarding the domain we stand on, and it is followed by a review of approaches that have previously been proposed in analysing and predicting big data in social networks research field under the AI umbrella.

A. Background about Big Data

The development of information technology has led to the generation of huge data from various sources especially the social media and IoT sensors sources that is growing exponentially every day. Data created in social media networks vary from texts, images, videos, maps, and sounds. These types of data are divided into structured and unstructured data where the relationship between friends is structured data that can be placed in a specific structure, and text is considered the unstructured data [7]. Big data analysis is used primarily to handle large-scale data generated very quickly with a variety of sources and formats. It's very convenient to face challenges that relational databases cannot. Techniques and methods for analysing this type of data have been studied in many publications, such as what was studied in [8].

Analysing big data leads to what is called 'big impact' according to the research work [9]. In this context, emerging analytics research opportunities are born. They can be classified into five critical technical areas, which include data analytics, text analytics, web analytics, network analytics, and mobile analytics.

B. Social Media and Big Data

Social media big data is primarily generated from the available users' behavior data that express browsing histories, what they buy, and all the information they are accessing online. The research in [10] provides a comprehensive view of how to manage the knowledge in social media big data. There are some common problems in applying big data techniques in social media data analysis [11]. These problems are data-driven versus theory - driven approaches, measurement validity, multi-level longitudinal analysis, and data integration.

Many studies have focused on big data analysis techniques with social media data [12-14]. Due to this focus and in terms of organization, big data analytics on social media is categorized into different classes to grasp their characteristics [15].

C. Machine Learning and Social Media Big Data

In the context of big data analysis, one must choose the appropriate method. Machine learning methods today are extensively used to analyze big data. The machine learning concept can generally be defined as "the computer software learns from the experience E of a particular task T with a P performance metric, if its performance for tasks T is improved by using measure P through experiment E [16]. T represents the problem to be addressed and resolved, such as classification problem. P is the performance metric that is related to the type of Task T to solve. While experience E represents the data used to conduct the learning process.

Machine learning is a subset of artificial intelligence that can modify itself by learning without any human intervention [17].

Different machine learning methods to analyze big data are discussed in detail in [18]. It compared many machine learning methods and found that to analyze big data, neural networks have advantages over other methods. A discussion about the challenges of learning with big data and the corresponding

possible solutions in recent researches was given in [19]. It states that, for big data, there is no single machine learning method that can be generalized. This is because we need different machine learning methods for different data. In addition, an analysis of the machine learning methods for big data is given in [20]. It emphasizes on the new characteristics of machine learning used to analyze big data by proposing a machine learning framework and to analyze big data based on parallel computing and distributed storage.

D. Deep Learning

In this section, the researchers explain the main concept of deep networks. In addition to that, different deep learning network architecture will be examined. The focus will be on the architecture that suitable for the research. Neurons and layers in deep learning network are significantly greater than in traditional neural networks. This increase in neurons and layers supports solving problems that are more complex. Moreover, there are developments in how layers communicate with each other and support for automatic feature extraction. Based on this principle, deep learning is now frequently used to analyze big data.

Various studies have applied deep learning to analyze social media big data. There are many applications using deep learning such as anticipating the behavior of users via social networks, sentiment analysis of text and business analysis. The research [21] concentrated on twitter data; it investigated the behavior of shared learning via Twitter data by examining the behavior of subscribers. Traditional learning methods used to analyze social networking data in crisis were criticized in [22] and deep learning has been tried to overcome these difficulties, since it proposes a new approach to improve the social media analysis incises situations.

Analyzing sentiment by using social media data is an important process. A framework in [23] for sentiment analysis was developed to build a self-developed military sentiment dictionary using deep learning. A real-time syndromic surveillance system is introduced in [24]. This system relies on Twitter social-media data analysis in order to be aware of a syndromic surveillance. Deep learning is used in this system to analyze data and determine the prevalence of a disease. Lamos and Cristianini are used in [25] social network users as sensors to predict real-time events such as rainfall prediction.

Furthermore, the prediction of earthquakes by the same idea is developed in [26] by using deep learning. Social media post analysis is introduced in [27] to extract various meaning information to predict any events. In addition, deep learning is used to make classifications and recommendation accurately. A massive amount of posts was analyzed in [28]. It studies on hashtag, retweet, characters in each tweet, and so on to classify users by their ages. It uses Deep Convolutional Neural Network after displaying the results of some traditional machine learning techniques like SVM. The impact of the recommendation and the adaptation process has become very large today, especially data on social media are increasingly growing. For instance, Multi-view Deep Neural Networking [29] was proposed to make recommendations on computer games that one can play based on data from Sony and Microsoft.

All these previous studies have appropriately contributed to the analysis of social media data, but one can see that there is still a gap and the research presented in this paper works to bridge this gap. The business community wants to make the most of social media data, and it requests that this benefit be translated into recommendations and plans through an integrated system. Therefore, the present research concludes that the design of an integrated system which is used by the business community to improve sales and advertising campaigns and inventory management is one of the important and required actions today. This research also discusses three deep neural networks that can be used extensively because they can handle huge data. They involve Deep stacking network, long short-term memory and Deep Boltzmann machines.

E. Deep Stacking Network (DSN)

Layers in this type of deep learning architecture are arranged in a hierarchical form as illustrated in Fig. 3 [30]. This deep learning type supports parallel learning, where the training of each module is held in isolation from the other blocks. In Fig. 3, there are only four modules, but hundreds of these modules are supposed to exist according to the type of the problem to be solved. DSN is mainly used to implement huge amount of data like social network data. The DSN has been developed and extended to the Tensor Deep Stacking Network (T-DSN) [31]. They are similar to a DSN but are different by adding some new layers such as a hidden layer.

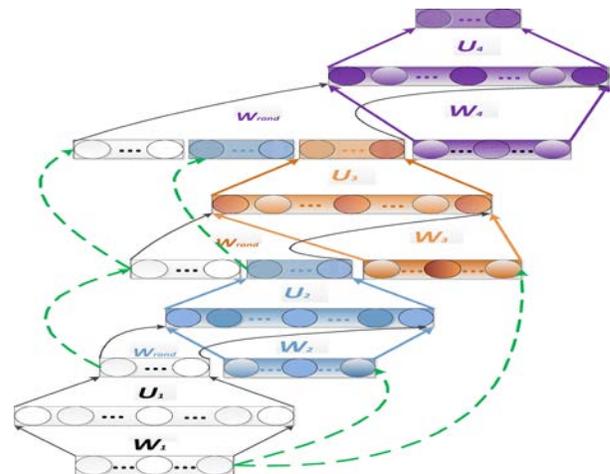


Fig. 3. A DSN Architecture using Input-output Stacking. Four Modules are Illustrated, each with a Distinct Color [30].

One advantage of DSN is that the lower module output combines with higher module inputs to be the final input for this higher unit. Applications designed based on DSN for data retrieval are described in the research [32].

F. Deep Boltzmann Machines (DBM)

Deep Boltzmann Machine (DBM) has several hidden layers [33]. Geoff Hinton defines the general Boltzmann machines as follows: "A network of symmetrically connected, neuron-like units that make stochastic decisions about whether to be on or off." DBM is an undirected model, it is used for data regression and time series. The hidden variables in DBM are used to learn the probability distribution of the input data. In another way, a network prevents any communication between nodes at the

same layer in Restricted Boltzmann Machines (RBM) [34] as illustrated in Fig. 4. In research [35], regression model was built using DBM to complete shapes. This network is successful in analysing heterogeneous data.

G. Long Short-Term Memory (LSTM)

Recurrent Neural Network is used for many purposes like speech and voice recognition, time series prediction, and natural language processing. This network is used to process sequential data such as that used to implement the autocomplete feature of words that predicts the rest of the words. Long short-term memory (LSTM) is an extension of Recurrent Neural Network [36], where it can consider much longer input sequences and overcome some obstacles. The structure of the LSTM is illustrated in Fig. 5.

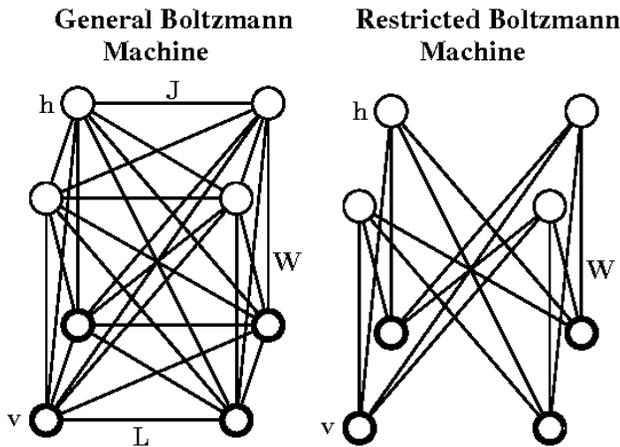


Fig. 4. Deep Boltzmann Machines [34].

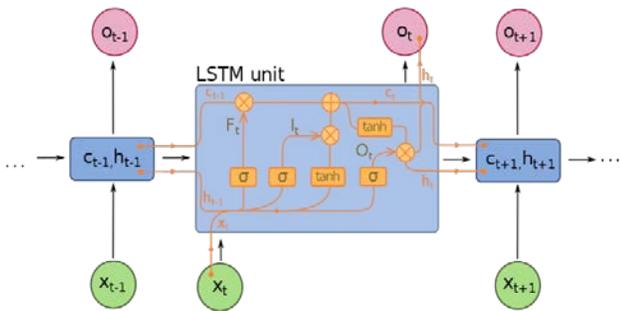


Fig. 5. Long Short-term memory LSTM[36].

The researchers in [37] proposed a system built on the reviews provided by Amazon Company. The system used neural network with Support Vector Machine (SVM) as a classifier and the accuracy achieved in this work is 81 %. The researchers of the current research call this system NN-SVM system for short. It is worth mentioning that this system will be compared later with the system proposed in this paper.

III. PROPOSED INTELLIGENT SYSTEM FOR SOCIAL MEDIA ANALYSIS

The proposed framework in Fig. 6 consists of a data base (data set), a number of users, and the proposed Intelligent Social Media Analysis System (ISMAS).

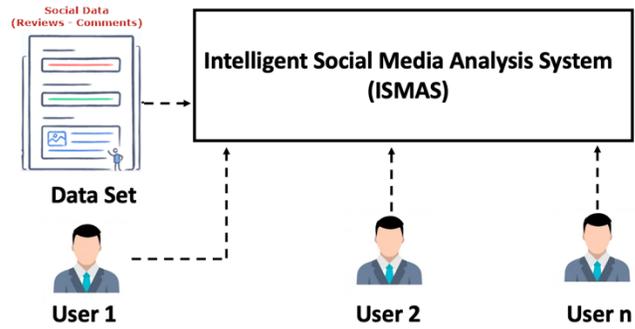


Fig. 6. Framework of the Proposed System ISMAS.

A. Data Set

The data base is used to train and test the ISMAS. The data used in this study is a set of approximately 3.5 million product reviews collected from Amazon.com by Fang et al. [38]. Thousands of Amazon product reviews were obtained for use in training and testing the deep learning model. A large group of reviews (28000) was collected in order to be classified according their ratings; each review consists of 40 words long. We adopted this length because most of these reviews fall within the scope of this length. These reviews have been labeled with ratings from 1 to 5. The resulted data are divided into two parts, the first part is used to train the model and its subset is 80% of the dataset, and the remaining part is used to test the accuracy of the model after its training. The ratings 1 and 2 are considered negative review, while the ratings 3, 4, and 5 are considered a positive review. Fig. 7 displays a review of the dataset.

Training and test data texts are extracted with their attached ratings, which are considered to their labels as illustrated in Fig. 8. The text of the reviews is divided into words and some preliminary processes are performed to prepare texts for the analysis process.

B. Proposed Intelligent Social Media Analysis (ISMAS)

The ISMAS consists of three main components. Table I summarizes the components, the task of each component, and the place where they are installed.

```
rating: 5.0 out of 5 stars
product_ID: B00DS842HS
helpfulness: 4/4
ID: A28R8UNBXGLFOR
review_by: Melliemel
title: It's working!
review_time: 20140308
review: So far so good. I bought this because I wanted to start oil pulling. It's
been working great. Great taste (while swishing it around and NOT
swallowing it). Put some on my arm that was very dry. It helped. Haven't
cooked with it yet, but I'm sure it will be great!
```

Fig. 7. Example Amazon.com Review [38].

Reviews.Rating	Reviews.Text
3	order 3 of them and one of the items is bad quality. Is missing backup spring so I have to put a pcs of aluminum to

Fig. 8. Reviews Text, and their Rating.

TABLE I. COMPONENTS

Component Name	Task
Text Analyzer	Text Analyzing.
Deep Learning Model	Determining degree of acceptance.
Decision Maker	Making recommendations.

Fig. 9 illustrates the architecture of the proposed ISMAS.

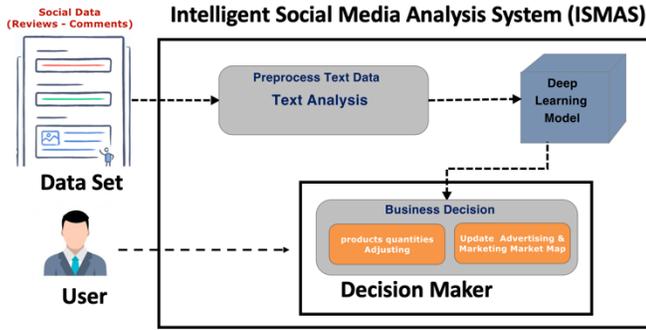


Fig. 9. The Architecture of the Proposed ISMAS.

1) *Text analyzer*: In general, texts are divided into tokens, for example, the sentence “training and test data” are divided into ‘training’ – ‘and’ – ‘test’ – ‘data’ tokens. Then, for these tokens, the researchers add Part-Of-Speech, for example adjective, noun, verb, etc. The lemmatization process is applied to reduce each token to their root form, for example “building” word can be reduced to “build”, also “words” transformed to “word”. The punctuation can be erased to adjust the accuracy. Again, words like “and”, “to” may add some noise to the texts, so the researchers need to delete those words before analyzing the product reviews. Besides, they have deleted any word that consists of two letters or less.

Upon this, the text analyzer performs a kind of text mining, where the last step of this process is extracting the knowledge (which is achieved by the deep learning model). Fig. 10 illustrates the three main steps of text mining [39].

In details, step 1 establishes the corpus, whereby all relevant unstructured data are collected. Then, the researchers digitize and standardize the collection. Finally, they place the collection in a common place (a directory consists of separate files).

In step 2, the Term-by-Document Matrix (TDM) is created. All terms are included in the TDM, such as stop words, synonyms, homonyms, and stemming.

In step 3, patterns/knowledge is extracted relying on deep learning model.

2) *Deep learning model*: Here, the intelligent model (ANN) is trained on the data that are analysed and stored in the TDM. This component performs the ILbM to determine the acceptance degree related to a certain product. The ILbM relies on the bagging technique to enable incremental learning, as shown in Fig. 11.

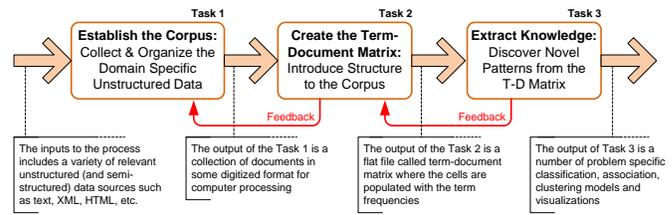


Fig. 10. Three-step Text mining Process [39].

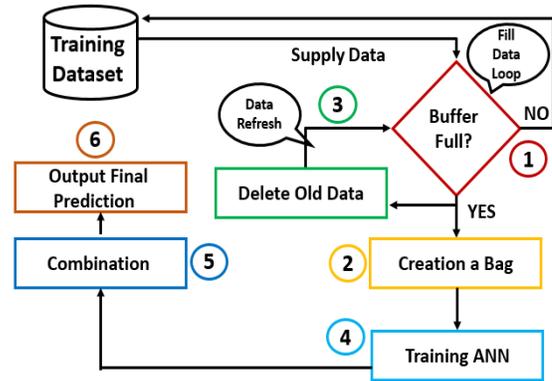


Fig. 11. Flow Chart of the Proposed Deep Learning Model.

As shown in Fig. 11, there are six steps in the proposed model, starting with filling a data loop and finishing with the final prediction step. Below is a description of each step.

a) *Filling data loop*: Here, a buffer is used as a temporal data storage and it is linked with the training data set. At the beginning the buffer is empty and gradually is filled with data. The data is supplied to the buffer from the dataset. The loop continues until the buffer is filled. The size of the buffer ($Zise_{buf}$) is calculated according to the following formula:

$$Zise_{buf} = \frac{size_{OTDB}}{10} \quad (1)$$

Where $size_{OTDB}$ denotes the size of the original training data set. When this condition is satisfied, the next step is executed.

b) *Creation a bag*: The bag represents a small training dataset that contains a part of the original dataset. The created current bag is used to train the first classifier (in the fourth step). When the buffer is filled again in the next iteration, a new bag is created.

c) *Old data deletion*: This step refers to empty the buffer. In other words, the old data used to create the current bag is deleted and a new data is supplied. This process is called data refresh. This process gives an advantage of enabling the bags in parallel as it is described later when it comes to enhancing the performance of the proposed system in terms of response time.

d) *Training artificial neural network*: In this step, the first classifier is trained on the data contained in the first bag (current one). The training process depends on two main stages, which are (1) extracting the features of the analysed text generated by the text analyser component; and (2) pooling the extracted features to draw a deep description of a product. The

extracted features are mainly inspired by strong and weak words included in the review of the products. Fig. 12 illustrates how to extract features.

In Fig. 12, the features are represented by a vector that contains highlighted words taken from the review of a product. The words included in the vector are taken from the review provided by Fig. 8. The words that have a positive trend are considered strong features that support the product, while the words that have negative trend are considered weak features. In Fig. 12, the words ‘Good and Great’ are strong features while the words ‘Very and Dry’ are weak features. Relying on the vector of features, the ANN model (classifier) is trained.

After finishing the extraction stage, the pooling stage starts. In the pooling stage, both the strong and weak features are formed to construct a sentence. The sentence is used to train the ANN to give an output (i.e., the class of the sentence to be positive or negative). The training process lasts until no reviews are found in the current bag. Consequently, the output of training ANN step is a trained classifier that can predict the class of a given review. Fig. 13 illustrates the structure of the ANN.

To generate the outputs, an activation function is used. In this work, the Softmax function is employed since it has a valuable feature called multiclass. In other words, the Softmax function has the property of handling more than two classes [40]. Therefore, the two classes (positive and negative) can be generated as a numeric data. Visually, the Softmax activation function is illustrated in Fig. 14.

The Softmax function is attached before the output layer, where positive class = 1 and negative class = -1, as shown in Fig. 15.

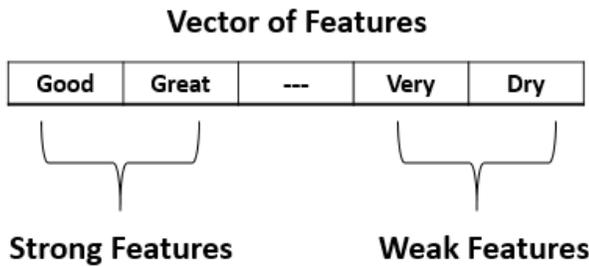


Fig. 12. Features Extraction.

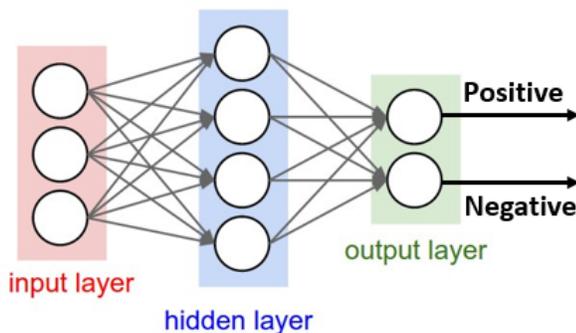


Fig. 13. Structure of ANN.

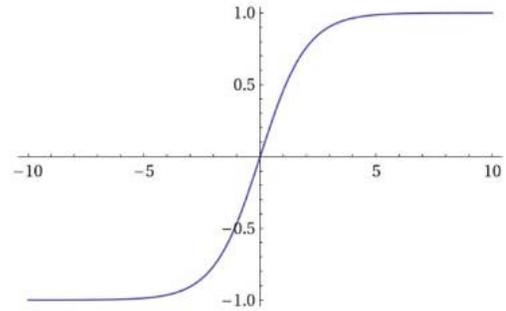


Fig. 14. Softmax Function [39].

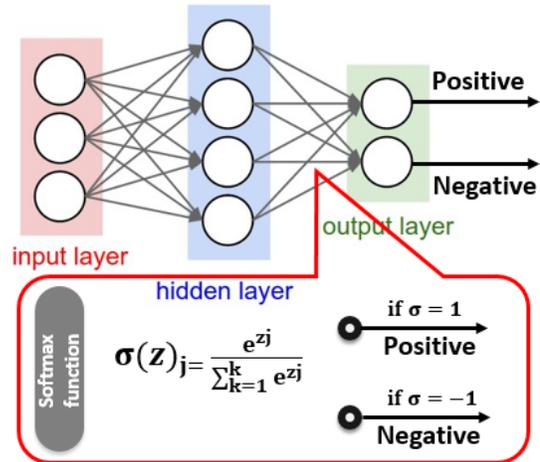


Fig. 15. Using the Softmax Function to Generate Outputs.

e) *Combination*: Based on formula 1, there will be 10 bags. Each bag is used to train an ANN classifier and consequently there will be 10 classifiers. Since each classifier generates a result, a combination process is performed to generate the result (prediction). Fig. 16 illustrates the combination of the result of each classifier.

It is worth mentioning that each classifier takes the knowledge learned by the previous one to be added as new knowledge. For example, classifier 1 learns by adding to the knowledge of classifier 2. Thus, the final classifier has the incremental knowledge learned by all the nine previous classifiers.

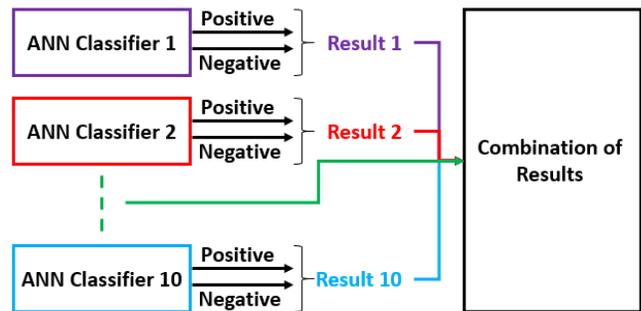


Fig. 16. Combination of results.

f) *Output final prediction:* In this step, the final prediction generated by the trained ANN model is shown as an input for the decision-maker component. This step includes marking the review by positive or negative for further manipulation by the evaluator.

3) *Decision maker:* This component is responsible for providing some recommendations. These recommendations fall in the business domain and are related to make some adaptation according to the situation that is currently dominating the market. In the context of the business domain, adaptation refers to a process that targets the most amount of gain. This can be achieved by providing some recommendations related to stop generating or supplying the market with more items of a specific product (the product that is predicted to have a positive or negative trend). Since there are many products under review each time, the decision maker groups the positive products in a cluster and does the same for the negative products. This facilitates to provide one recommendation to all positive products and all negative ones. Fig. 17 shows the clustering of products.

C. *Employing Hadoop for Better Performance*

Hadoop is an open-source distributed processing framework that manages data processing and storage for big data applications in scalable clusters of computer servers [41]. It's at the center of an ecosystem of big data technologies that are primarily used to support advanced analytics initiatives, including predictive analytics, data mining and machine learning. Hadoop systems can handle various forms of structured and unstructured data, giving users more flexibility for collecting, processing, analyzing and managing data than relational databases and data warehouses provide [42]. It is worthwhile to mention that the data supplied to the Hadoop platform should follow the security and privacy agreements. Security and privacy issues are highly highlighted according to many researches [43-52]. In this work, privacy and security of data are considered out of scope, where they will be considered in future work.

Hadoop allows parallel execution of jobs. This means that the job that is intended to be executed on Hadoop must be paralyzed. In this work, the job that is executed in parallel on Hadoop is the training phase of the model. That is because each bag is used to train one classifier, the number of bags was determined in proportion to the data in the training and consequently each classifier can be trained separately in a single thread. Fig. 18 illustrates using Hadoop to train the classifiers using threads.

Hadoop uses MapReduce technique to perform parallel execution. The major advantage of MapReduce is that it is easy to scale data processing over multiple computing nodes [53]. The MapReduce technique contains two important tasks, namely Map and Reduce. The main mission of the Map is to divide the job into parts and then distribute the parts over computational nodes (servers) for execution. The main mission of the Reduce is to collect the results of the executed parts and return them to the master node. Fig. 19 illustrates the mechanism of MapReduce technique used for training the classifier.

As shown in Fig. 19, there are 10 computation nodes. Each bag is loaded on a node (i.e., the training data set used to train a classifier). After processing (i.e., training the ANN on each node), the results are gathered by the Reduce to return them back to the master node. Here, the results are represented by the rules that the classifiers are learned. At the master node side, all the rules are combined to form the final trained classifier.

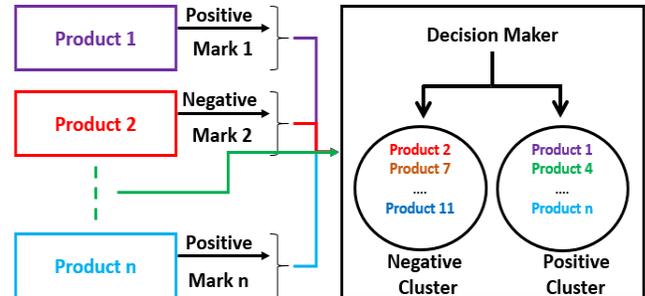


Fig. 17. Clustering of Products.

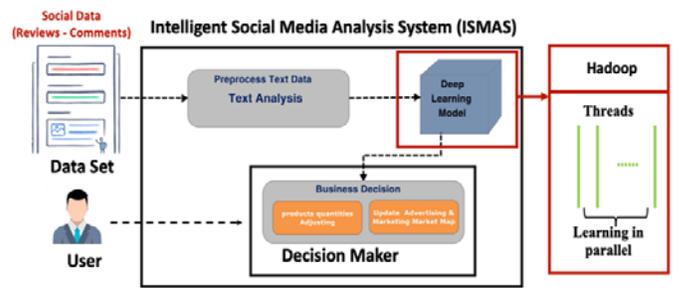


Fig. 18. Employing Hadoop for Parallel Training.

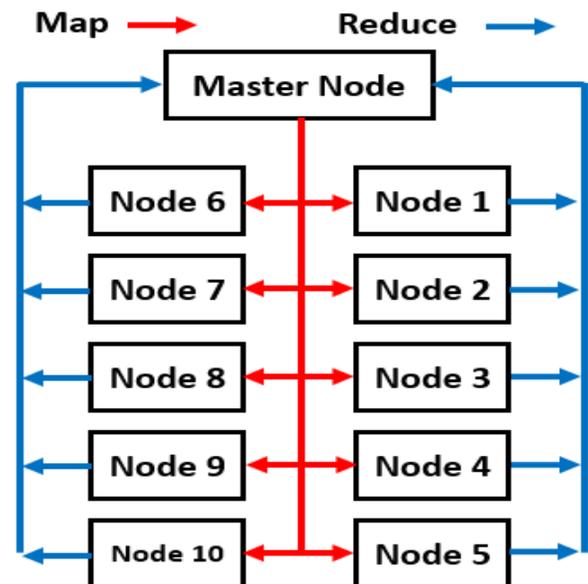


Fig. 19. Map Reduce Mechanism.

The steps of the ILbMare are shown in the following algorithm.

Algorithm 1: Incremental Learning-based Method (ILbM)

```

1: begin
2:   Read Input ( text = reviews);
3:   Pre-processing TextAnalyzing (reviews);
4:   ResultSet = ∅;
5:   for (i = 1 to 10) do
6:     begin
7:       creation (bagi) ;
8:       Training (Ci);
9:       Generate (resulti = rulesi);
10:      ResultSet = ResultSet ∪ resulti;
11:     end;
12:   Generate ResultSet;
13:   Class = Classify (new review);
14:   recommendation (class);
15: end;

```

IV. USED METRICS

Two types of metrics are presented for use in the evaluation process. They are AI-based metrics, and performance-based metrics.

A. AI-based Metrics

Basically, the confusion matrix (CoMa) is an effective benchmark for analyzing how well a classifier can recognize the images of different classes. The CoMa is formed considering the following terms [41]:

- 1) True positives (TP): positive images that are correctly labelled by the classifier.
- 2) True negatives (TN): negative images that are correctly labelled by the classifier.
- 3) False positives (FP): negative images that are incorrectly labelled as positive.
- 4) False negatives (FN): positive images that are mislabeled as negative.

Table II shows the CoMa in terms of the TP, FN, FP, and TN.

TABLE II. CONFUSION MATRIX

Actual class (Predicted class)	Confusion matrix		
	CI	¬ CI	total
C1	True Positives (TP)	False Negatives (FN)	TP + FN = P
¬C1	False Positives (FP)	True Negatives (TN)	FP + TN = N

Relying on the CoMa, the accuracy (Acc), precision (Pre), and recall (Rec) metrics are driven. For a given classifier, the accuracy can be calculated by considering the recognition rate, which is the percentage of the test set images that are correctly classified. The accuracy is defined as [41]:

$$Acc = \frac{(TP+TN)}{\text{number of all records}} \quad (2)$$

Accuracy-based evaluation. In this context, a higher accuracy corresponds to a better classifier output. The maximum value of the accuracy metric is 1 (or 100%), which is achieved when the classifier classifies the images correctly without any error during the classification process.

For precision, it refers to the exactness (what % of tuples that the classifier labelled as positive that are actually positive). It is given by [41]:

$$Pre = \frac{TP}{TP+FP} \quad (3)$$

Precision-based evaluation. In this context, a higher precision corresponds to a better classifier output. The maximum value of the precision metric is 1 (or 100%), which is achieved when FP=0.

Recall refers to the completeness (what % of positive tuples did the classifier label as positive?). It is given by [41]:

$$Rec = \frac{TP}{TP+FN} \quad (4)$$

Recall-based evaluation. In this context, a higher recall corresponds to a better classifier output. The maximum value of the recall metric is 1 (or 100%), which is achieved when FN=0.

B. Performance-based Metrics

Time of response (ToR) is used to evaluate approaches in terms of performance. The ToR is calculated based on the total time of the four main steps that are illustrated in Fig. 12 (i.e. fill data loop, bag creation, ANN training, combination, and output steps). ToR is given as:

$$ToR = T_{fdl} + T_{br} + T_{tr} + T_{cob} + T_{opt} \quad (5)$$

Where T_{fdl} , T_{br} , T_{tr} , T_{cob} , T_{opt} denote the time consumed by the fill data loop, bag creation, ANN training, combination, and output steps, respectively.

It is important to notice that the shorter the ToR is, the better the performance of this approach is.

V. EXPERIMENTAL RESULTS AND EVALUATIONS

This section is arranged so that it first presents the setup and the approach that is intended to be compared. Then, the actual results and the corresponding discussions are provided.

A. Setup

The proposed AI-based system is implemented on a laptop that has the specifications shown in Table III.

The researchers selected one system (NN-SVM system), mentioned in the related work section, to compare with the proposed ISMAS system. It is illustrated in Table IV.

TABLE III. SPECIFICATIONS OF LAPTOP

Item	Details (value)
Operating system	MacOS Cataling version 10.15.5
Processor	Dual-Core Intel Core i5
RAM	8 GB
Speed	2.9 GHz

TABLE IV. SELECTED APPROACHES

Ref	Selected System		
	Used technique	Journal	Year
[37]	NN-SVM	International Journal on Soft Computing, Artificial Intelligence and Applications	2019

B. Results and Discussion

The results are provided according to both the AI-based metrics as well as the performance metric.

1) *Results depending on the AI-based metrics:* Since there are 10 ANN classifiers due to using 10 bags, the results of the AI-based metrics are arranged in Table V.

Table V shows that the accuracy of the 10 classifiers vary within the range [90 %, 98%] and the corresponding average of all accuracies is 96.7 %. The precession achieves lower values, where they vary within the range [90 % - 97%] and the corresponding average of all precisions is 95.1%. The recall of the 10 classifiers varies within the range [88%, 94%] and the corresponding average of all recall is 91.5%.

TABLE V. 10 CLASSIFIERS EVALUATION

Iteration (Bag NO)	AI-based Metrics		
	Acc	Pre	Rec
1	98 %	97 %	93 %
2	96 %	90%	90 %
3	97 %	95 %	94 %
4	98 %	96 %	93%
5	98%	95 %	90 %
6	98 %	96.5 %	94 %
7	90 %	93 %	90 %
8	98 %	95%	88 %
9	97%	97 %	91 %
10	97 %	96 %	92 %
Average	96.7 %	95.7 %	91.5 %

TABLE VI. AI-BASED METRIC VALUES

System	AI-based Metrics		
	Acc	Pre	Rec
ISMAS	96.7 %	95.1%	91.5 %
NN-SVM	81.29 %	58.61 %	40.85 %

Table VI summarizes the comparison between the proposed ISMAS system and the NN-SVM system according to the AI-based metrics.

AI-based metrics discussion. The proposed ISMAS system achieves better scores in terms of accuracy, precision, and recall. In details, there is (96.7-81.29=15.41) extra degree for the accuracy, (95.1-58.61=36.49) extra degree for the precision, and (91.5-40.85=50.65) extra degree for the recall. The reason for these enhancement of the AI-based metrics is related to the incremental learning of the classifier. In other words, the cumulative knowledge gained by incremental learning is more efficient when compared to one learning iteration in the NN-SVM system.

2) *Results depending on the performance-based metric:* In the context of this evaluation, the ToR is calculated for each classifier, and then the average time is obtained. Fig. 20 shows the results.

Fig. 20 shows that the ToR metric has a value that varies within the range [31 - 41] seconds, where the value 31 refers to the performance of the 6th classifier, while 41 refers to the performance of the 3rd classifier. The third classifier has the worst ToR value because of some operations related to the running of the program. The average ToR values refers to the final performance of the final classifier, which is calculated as:

$$AVG = \frac{40+36+41+35+32+31+34+39+37+35}{10} = 36 \text{ second}$$

In terms of evaluating the proposed system under the comparison with the NN-SVM system, Fig. 21 shows the results.



Fig. 20. Performance of the 10 Classifiers.

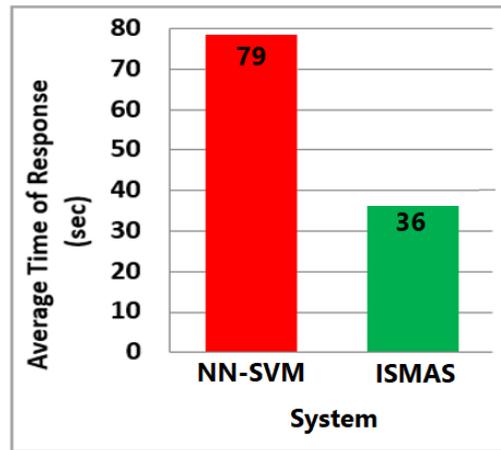


Fig. 21. Performances of Systems.

REFERENCES

Performance-based metric discussion. It is obvious from Fig. 21 that the proposed ISMAS system outperformed the NN-SVM system. The speed of the performance of the proposed ISMAS can be calculated as:

$$\frac{79}{36} = 2.19$$

This means that the proposed ISMAS system has more than double speed when compared to the NN-SVM system. The reason behind the enhancement of the performance of the proposed ISMAS system is employing Hadoop to perform the training in parallel. Theoretically, the ToR of the proposed system should be 10 times lesser than the ToR of the NN-SVM system. However, the procedure of creating Mappers to distribute the job over the nodes as well as the time of creating reducers and combining the sub-results to form the final results is the reason why the ToR of the proposed system is about 2 times lesser than the NN-SVM system.

VI. CONCLUSION

Recently, analysing reviews presented by clients to products that are provided by e-commerce companies, such as Amazon, to produce efficient recommendations has received a wide attention and is considered as a vital research in the business field. In this paper, the researchers propose an artificial intelligence recommendation system for analysing reviews in the business area. The proposed system consists of three main components, which are text analyser, deep learning model, and decision-maker. The text analyser component uses the Term-by-Document Matrix (TDM) in the process of analysing the reviews. The reviews are obtained from a data base provided and available by Amazon e-commerce Company. The deep learning model (classifier) is trained based on the bagging technique. Ten bags are created with a data inspired from the data base and ten classifiers are trained. The knowledge (represented by the rules of classifying the reviews) that is gained by the first classifier is added to the next one until the 10th classifier. In regard to enhancing the response time (the performance), the ten classifiers are trained in parallel using the Hadoop platform depending on the MapReduce technique. The decision-making component clusters the final results of the classifiers to generate an efficient recommendation for better gain purpose. The proposed system is evaluated based on two kinds of metrics, which are the artificial intelligence metrics (accuracy, precision, and recall) and the performance metrics (time of response). Compared to the NN-SVM similar system, the proposed system showed better results as follows: accuracy (96.7), precision (95.1), recall (91.5), and time of response (36 seconds).

In terms of limitations, the proposed system did not consider the sentiment analysis that may be found in the reviews provided by the customers.

In future work, the researchers intend to take into account the sentiment analysis to obtain a higher accuracy percentage. In addition, they intend to apply the proposed system on a real dataset in real time taking into consideration privacy and security issues.

- [1] Zikopoulos Paul, Chris Eaton, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*, McGraw-Hill Osborne Media, 2011.
- [2] Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, 2011. http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation (2011).
- [3] Iafate, Fernando. *Artificial intelligence and big data: the birth of a new intelligence*. John Wiley & Sons, 2018.
- [4] Sharda, Ramesh, et al. "Business intelligence and analytics." System for Decision Support (2014).
- [5] Benioudakis, Myron, Apostolos Burnetas, and George Ioannou. "Lead-time quotations in unobservable make-to-order systems with strategic customers: Risk aversion, load control and profit maximization." *European Journal of Operational Research* 289, no. 1 (2021): 165-176.
- [6] Karim, Ahmad, et al. "Big data management in participatory sensing: Issues, trends and future directions." *Future Generation Computer Systems* 107 (2020): 942-955.
- [7] Baars, H., & Kemper, H.-G. (2008). Management support with structured and unstructured data – An integrated business intelligence framework. *Information Systems Management*, 25(2), 132–148. <http://dx.doi.org/10.1080/10580530801941058>.
- [8] A. Gandomi and M. Haider, Beyond the hype: Big data concepts, methods, and analytics, *International Journal of Information Management*, 35(2) (2015), pp.137-144.
- [9] Chen, Hsinchun, Roger HL Chiang, and Veda C. Storey. "Business intelligence and analytics: From big data to big impact." *MIS quarterly* (2012): 1165-1188.
- [10] Bello-Orgaz, G., Jung, J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. *Information Fusion*, 28, 45--59.
- [11] QIU, Lin, CHAN, Sarah Hian May, & CHAN, David.(2017). Big data in social and psychological science: Theoretical and methodological issues. *Journal of Computational Social Science*, 1(1), 59-66.
- [12] Stahl, Frederic, Mohamed Medhat Gaber, and Mariam Adedoyin-Olowe. "A survey of data mining techniques for social media analysis." *Journal of Data Mining & Digital Humanities* 2014 (2014).
- [13] Kim, J., & Hastak, M. (2018). Social network analysis. *International Journal of Information Management: The Journal for Information Professionals*, 38(1), 86-96.
- [14] Stieglitz, Stefan, et al. "Social media analytics—Challenges in topic discovery, data collection, and data preparation." *International journal of information management* 39 (2018): 156-168.
- [15] Ghani, Norjihjan Abdul, et al. "Social media big data analytics: A survey." *Computers in Human Behavior* 101 (2019): 417-428.
- [16] Mitchell, T. om, *Machine Learning*. McGraw-Hill Science/Engineering/Math, 1997.
- [17] Ethem Alpaydin. 2010. *Introduction to Machine Learning* (2nd ed.). The MIT Press.
- [18] Lidong Wang, "Data Mining, Machine Learning and Big Data Analytics" , *International Transaction of Electrical and Computer Engineers System*, 2017, Vol. 4, No.2, 55-61.
- [19] Yuhua Xu and ShuoFeng ,JunfeiQiu, Qihui Wu, Guoru Ding, "A survey of machine learning for big data Processing", *EURASIP Journal on Advances in Signal Processing* (2016) 2016:67.
- [20] Sreenivas R. Sukumar, "A Research on Machine Learning Methods for Big Data Processing", Conference Paper August 2014.
- [21] H.Aramo-Immonen,J.Jussila,andJ.Huhtamäki, "Exploringco-learning behavior of conference participants with visual network analysis of Twitter data," *Comput. Hum. Behav.*, vol. 51, pp. 1154–1162, Oct. 2015.

- [22] Lazreg, Mehdi Ben, Morten Goodwin, and Ole-Christoffer Granmo. "Deep learning for social media analysis in crises situations." In *The 29th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS) 2-3 June 2016, Malmö, Sweden*, p. 31. 2016.
- [23] Chen, Liang-Chu, Chia-Meng Lee, and Mu-Yen Chen. "Exploration of Social Media for Sentiment Analysis Using Deep Learning." *Soft Computing* (October 8, 2019). doi:10.1007/s00500-019-04402-8.
- [24] Şerban O, Thapen N, Maginnis B, Hankin C, Foot V. Real-time processing of social media with SENTINEL: A syndromic surveillance system incorporating deep learning for health classification. *Inf Process Manage* 2019;56 (3):1166–84. 10.1016/j.ipm.2018.04.011.
- [25] Lampos, Vasileios, and Nello Cristianini. "Nowcasting events from the social web with statistical learning." *ACM Transactions on Intelligent Systems and Technology (TIST)* 3, no. 4 (2012): 1-22.
- [26] Sakaki, Takeshi, Makoto Okazaki, and Yutaka Matsuo. "Earthquake shakes twitter users: real-time event detection by social sensors." In *Proceedings of the 19th international conference on World wide web*, pp. 851-860. 2010.
- [27] M. B. Lazreg, M. Goodwin, and O.-C. Granmo, "Deep learning for social media analysis in crisis situations," in Proc. 29th Annual Workshop Swedish Artif. Intell. Soc. (SAIS), Malmö, Sweden, Jun. 2016, pp. 1–46.
- [28] R. G. Guimarães, R. L. Rosa, D. De Gaetano, D. Z. Rodríguez, and G. Bressan, "Age groups classification in social network using deep learning," *IEEE Access*, vol. 5, pp. 10805–10816, 2017.
- [29] A.M.Elkahky, Y.Song, and X.He, "A multi-view deep learning approach for cross domain user modeling in recommendation systems," in Proc. 24th Int. Conf. World Wide Web, 2015, pp. 278–288.
- [30] Li Deng; Dong Yu, "Deep Learning: Methods and Applications," in *Deep Learning: Methods and Applications*, now, 2014, pp.
- [31] B. Hutchinson, L. Deng, and D. Yu, "Tensor deep stacking networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1944–1957, Aug. 2013.
- [32] L. Deng, X. He and J. Gao, "Deep stacking networks for information retrieval," 2013 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, 2013, pp. 3153-3157. doi: 10.1109/ICASSP.2013.6638239.
- [33] R. Salakhutdinov and G. Hinton, "Deep Boltzmann machines," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*, vol. 24, pp. 448–455, 2009.
- [34] Hinton. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 14(8):1711–1800, 2002.
- [35] Wang, Zheng, and Qingbiao Wu. "Shape Completion Using Deep Boltzmann Machine." *Computational Intelligence and Neuroscience* 2017 (2017). Gale Academic OneFile (accessed March 12, 2020).
- [36] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.
- [37] Shrestha, Nishit, and Fatma Nasoz. "Deep Learning Sentiment Analysis of Amazon. Com Reviews and Ratings." *International Journal on Soft Computing, Artificial Intelligence and Applications*, 8(1), pp.01-15. (2019).
- [38] Fang, Xing, and Justin Zhan. "Sentiment analysis using product review data." *Journal of Big Data* 2.1 (2015): 5.
- [39] Turban, Efraim, et al. *Jue Ce Zhi Chi Yu Shang Wu Zhi Neng Xi Tong = Decision Support and Business Intelligence Systems*. Zhong Guo Ren Min Da Xue Chu Ban She, 2015.
- [40] Khanvilkar, G., & Vora, D. (2018). Activation functions and training algorithms for deep neural network. UGC approved journal, *International Journal of Computer Engineering In Research trends*, 5(4), 98-104.
- [41] Vohra, Deepak. *Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-Related Frameworks and Tools*. Apress, 2016.
- [42] Rao, Boyina Subrahmanyeswara. "Dynamic Histogram Equalization for contrast enhancement for digital images." *Applied Soft Computing* 89 (2020): 106114.
- [43] Alrahhal, Mohamad Shady, Maher Khemakhem, and Kamal Jambi. "Agent-Based System for Efficient kNN Query Processing with Comprehensive Privacy Protection." *International Journal Of Advanced Computer Science And Applications* 9.1 (2018): 52-66.
- [44] Alrahhal, Mohamad Shady, et al. "AES-route server model for location based services in road networks." *International Journal Of Advanced Computer Science And Applications* 8.8 (2017): 361-368.
- [45] Alrahhal, Mohamad Shady, Maher Khemakhem, and Kamal Jambi. "A Survey on Privacy of Location-Based Services: Classification, Inference Attacks, and Challenges." *Journal of Theoretical & Applied Information Technology* 95.24 (2017).
- [46] Alrahhal, Mohamad Shady, Maher Khemakhem, and Kamal Jambi. "Achieving load balancing between privacy protection level and power consumption in location based services." (2018).
- [47] Alrahhal, H.; Alrahhal, M.S.; Jamous, R.; Jambi, K. A Symbiotic Relationship Based Leader Approach for Privacy Protection in Location Based Services. *ISPRS Int. J. Geo-Inf.* 2020, 9, 408.
- [48] Al-Rahal, M. Shady, Adnan Abi Sen, and Abdullah Ahmad Basuhil. "High level security based steganography in image and audio files." *Journal of theoretical and applied information technology* 87.1 (2016): 29.
- [49] Alluhaybi, Bandar, et al. "A Survey: Agent-based Software Technology Under the Eyes of Cyber Security, Security Controls, Attacks and Challenges." *International Journal of Advanced Computer Science and Applications (IJACSA)* 10.8 (2019).
- [50] Fouz, Fadi, et al. "Optimizing Communication And Cooling Costs In Hpc Data Center." *Journal of Theoretical and Applied Information Technology* 85.2 (2016): 112.
- [51] Mona Alfifi, Mohamad Shady Alrahhal, Samir Bataineh and Mohammad Mezher, "Enhanced Artificial Intelligence System for Diagnosing and Predicting Breast Cancer using Deep Learning" *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(7), 2020. <http://dx.doi.org/10.14569/IJACSA.2020.0110763>.
- [52] Alrahhal, Mohamad Shady, and Adnan Abi Sen. "Data mining, big data, and artificial intelligence: An overview, challenges, and research questions." (2018).
- [53] Mohamed, Ehab, and Zheng Hong. "Hadoop-MapReduce job scheduling algorithms survey." *2016 7th International Conference on Cloud Computing and Big Data (CCBD)*. IEEE, 2016.

Comprehensive Analysis for Sensor-based Hydraulic System Condition Monitoring

Ahmed Alenany¹, Ahmed M. Helmi², Basheer M. Nasef³

Computer and Systems Department, Faculty of Engineering, Zagazig University, Zagazig 44519, Egypt^{1,2,3}

Faculty of Computer Science, Nahda University in Beni Suef, New Benisuef City, Egypt¹

Engineering and Information Technology Collage, Buraydah Private Colleges, Buraydah, KSA.²

College of Science and Humanities Studies in Al-Quwaieyah, Shaqra University, Al-Quwaieyah, KSA³

Abstract—Condition monitoring of equipment can be very effective in predicting faults and taking early corrective actions. As hydraulic systems constitute the core of most industrial plants, predictive maintenance of such systems is of vital importance. Due to the availability of huge data collected from industrial plants, machine learning can be used for this purpose. In this work, a hydraulic system condition monitoring (HSCM) is addressed via a public dataset with 17 sensors distributed throughout the system. Using a set of 6 features extracted from sensory data, the random forest classifier was proven, in the literature, to achieve classification rate exceeding 99% for four independent target classes, namely, Cooler, Valve, Pump and Accumulator. In this paper, sensor dependency is examined and experimental results show that a reduced set of important sensors may be sufficient for the addressed classification task. In addition, feature importance as well as implementation issues, i.e. training time and model size on disk, are analyzed. It is found that the training time can be reduced by 25.7% to 36.4% while the size on disk is reduced by 70.3% to 85.5%, using the optimized models, with only important sensors employed, in comparison with the basic model, with full set of sensors, while maintaining classification precision.

Keywords—Condition monitoring; sensory data analysis; machine learning; classification

I. INTRODUCTION

This Faults or failure of equipment in an industrial plant may have serious consequences ranging from threatening the safety of operators, causing the plant to shut down for long periods of time, and lowering production rate and revenue [1]. For these reasons, maintenance plays a crucial role in process industries.

The simplest strategy for maintenance is to wait until the fault occurs and then start reacting. In this strategy, there is cost for replacing the damaged equipment and additional cost for the loss of production during equipment downtime. More advanced strategy is scheduled maintenance which is performed periodically. This approach, despite being effective, may take corrective actions which are unnecessary and costly. The most advanced approach is predictive maintenance in which the condition of equipment is continuously monitored and faults are predicted and necessary corrective actions are taken [2].

Currently, equipment condition monitoring (CM) is possible thanks to the advances in sensor technology as well as

machine learning techniques which can process huge bulks of data from sensors distributed throughout the plant. These techniques, which extract key features from data and correlate them to possible faults [3] are successfully applied in condition monitoring of e.g. gearbox [4-6], rotating machinery [7], motor bearings [8], centrifugal pumps [9,10], hydraulic systems [11], cutting tools [12], grinding mill liners [13], and semiconductor failures [14].

Hydraulic systems are core components in most fields of industries such as water treatment plants, vehicle, aerospace [15], and other industries. The failure of hydraulic system can cause a whole plant to shut down or threaten operators' safety [16-18]. In addition, hydraulic systems are not operator-friendly environments for condition monitoring [19]. Due to these reasons, condition monitoring of hydraulic systems gain a lot interest in the past two decades. For example, Liu [16] developed a tree structure model for fault diagnosis of out-of-sync oil cylinder. El-Betar et al. [17] proposed a neural network scheme for fault diagnosis of actuator leakage and valve spool blockage. Tian et al. [18] applied support vector machines (SVM) for predicting pump faults. On the other hand, Jegadeeshwaran and Sugumaran [20] employed both SVM and decision trees for fault detection of hydraulic brake systems. Helwig et al. [21] developed a hydraulic test rig with several induced fault types. Key features are extracted from sensors' data and the most highly correlated with a given fault are determined. Linear discriminant analysis (LDA) was employed to reduce feature space. The same test rig was further examined by Chawathe [22] who applied naïve Bayes, decision trees, and random forests (RF) [23]. RF classifier achieved classification accuracy of about 99% for all classes. Furthermore, it was noted that accuracy can be retained using only small set of features. Quatrini et al. [15] have studied the same dataset and used Pearson's correlation coefficient to rank the features correlated with a given fault. Algorithms such as SVM, ANN, RF, and logistic regression are tested and again RF outperforms the other techniques for most fault types. On the other hand, König and Helmi [24] applied convolutional neural network (CNN) successfully for the same dataset [21]. In the contrary to previous studies, CNN has the ability to automatically extract key features. In addition, an analysis of misclassifications is also conducted.

The main objectives of this work are to determine the sensors and features which are more effective in detecting a

given type of fault in HSCM system, and to analyze the implementation issues of the optimized classifier models.

The paper is organized as follows: the benchmark hydraulic system and dataset are described in Section 2. Section 3 illustrates the classifier model. The experimental results are then presented in Section 4. Section 5 presents a detailed discussion and the most important findings are highlighted. Finally, conclusions and future work are given in Section 6.

II. HYDRAULIC SYSTEM DATASET

For the purpose of setting up an environment to test and diagnose common faults in hydraulic systems, Helwig, Pignanelli, and Schütze [21] developed the hydraulic test rig shown in Fig. 1. In this system, several reversible faults, with different degrees of severity, can be induced and the data of the sensors distributed throughout the system are recorded. By correlating features extracted from sensor data and known faults, an automated mechanism for fault detection can be developed.

The system consists of two hydraulic circuits: the primary working (Fig. 1, top) and the cooling-filtration circuit (Fig. 1, bottom). The two circuits are connected through an oil tank. The primary circuit contains the main pump (MP1) and a relief valve (V11) which can be used to generate different load levels in the circuit, as well as four-compartments accumulator (A1,

A2, A3, and A4) for pressure storage. The secondary circuit contains the cooler unit (C1).

The test rig is equipped with 14 sensors to measure pressure (PS1 – PS6), flow (FS1, FS2), temperature (TS1 – TS4), electrical power (EPS1), and vibration (VS1). The measurements are recorded using the standard industrial 20 mA current loop interfaces connected to a data acquisition system within Beckhoff CX5020 PLC. Additionally, three virtual sensors are designed to provide estimates for system efficiency (SE), cooling efficiency (CE), and cooling power (CP). The data is collected with sampling rates of 100 Hz for pressure and motor power, 10 Hz for flow rate, and 1 Hz for other variables.

In addition to recording 17 sensors' data, the state or condition of the following targets: Cooler, Valve, Pump and Accumulator, are also recorded. A total of 2205 cycles or training examples are collected, each is 60 second long. The training examples contains cases for each state of the four targets ranging from being fully operating to close to complete failure. A list of targets, degrees of faults, their abbreviations, and the corresponding number of training examples are given in Table I. As can be seen, the problem at hand can be considered as four separate classification problems, one for each target.

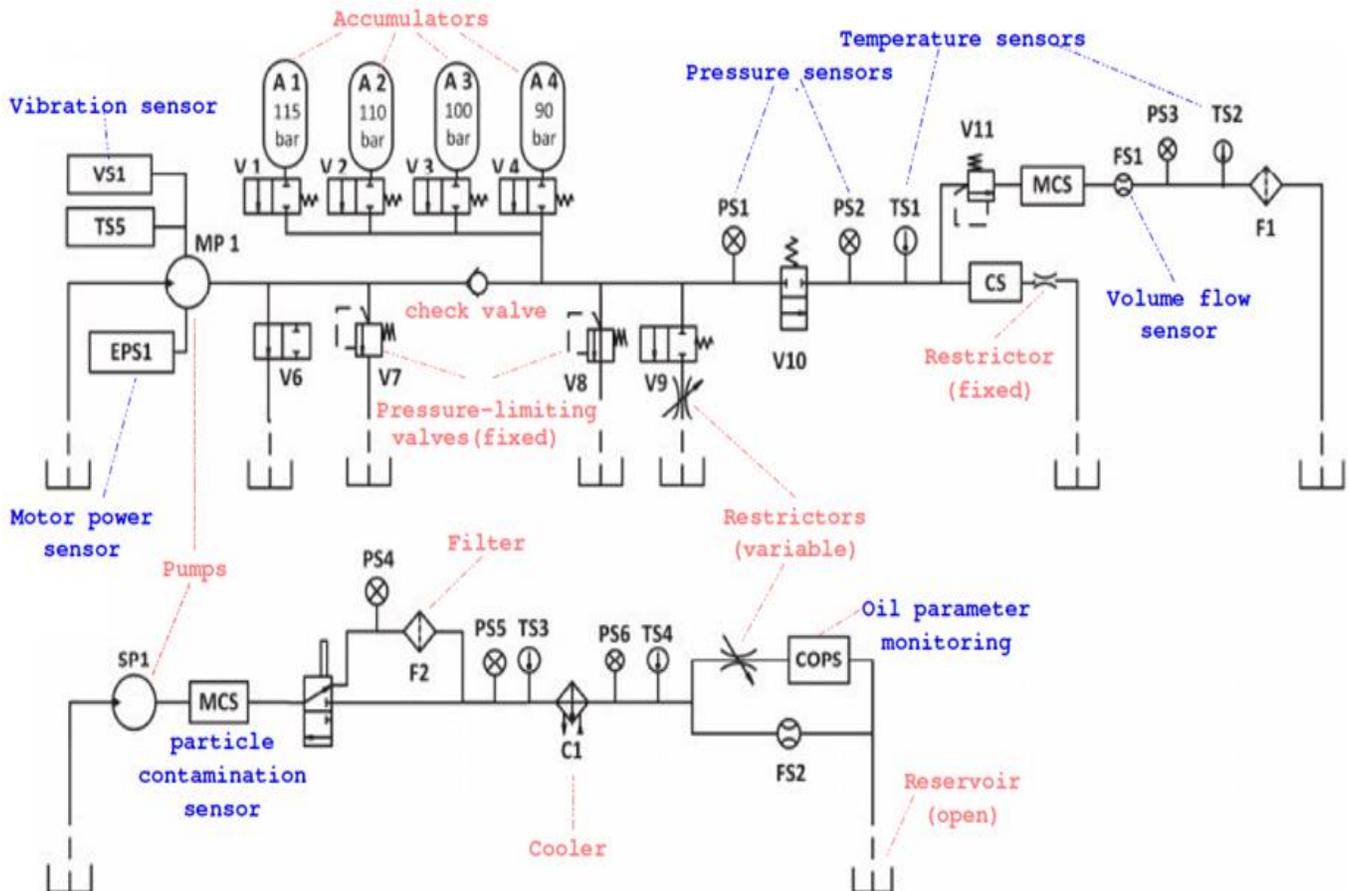


Fig. 1. Hydraulic System Under Study. Basic Monitored Units are in Light Red While Sensors are in Blue [22].

TABLE I. LIST OF TARGETS, FAULT STATES, AND THE NUMBER OF TRAINING EXAMPLES FOR EACH, IN THE BENCHMARK DATASET [21]

Condition	States	Abbr.	Examples
Cooler	Close to Complete Failure	CF	732
	Reduced Efficiency	RE	732
	Full Efficiency	FE	741
Valve	Close to Complete Failure Severe Lag	CF	360
	Small Lag	SvL	360
	Optimal Switching	SmL	360
		OS	1125
Pump	Severe leakage	SL	492
	Weak leakage	WL	492
	No leakage	NL	1221
Accum.	Close to Total Failure	CF	808
	Severely Reduced Pressure	SvP	399
	Slightly Reduced Pressure	SIP	399
	Optimal Pressure	OS	599

III. THE CLASSIFIER MODEL

The two main components of a classification task are the extraction of features and the use of a suitable type of classifier.

Features are key representative attributes of raw sensor data. They can be time domain or frequency domain. The set of features introduced by Quatrini et al. [15] are reused here to implement the experimental models. Each operation cycle is represented by 6 features, namely mean (m), standard deviation (sd), skewness (sk), kurtosis (k), slope of linear fit (slf) and position of maximum (p). The first four features characterize the distribution density of sensory signals, while slf and p features can capture the shape of the signal. These features proved very useful for fault recognition in such applications. Thus, for the 17 sensors, there are 102 features in total.

On the other hand, according to previous studies, random forest (RF) and artificial neural networks show outstanding results in the addressed dataset with a slight preference of RF [15, 24]. As the objective of the current work is not to compare different classifiers but to determine which sensors and features are important in detecting a given type of fault, random forest is employed in this paper.

Random forest classifier is an ensemble of decision trees. Each tree is fed with a set of features and provides a decision which represents a corresponding class. Within the forest, the most voted class is selected as the final classifier output [23]. To use a random forest, two parameters need to be set: the number of trees and the maximum number of splits allowed in each tree which controls the tree depth.

IV. EXPERIMENTAL RESULTS

In this section, several experiments are conducted to test the performance of RF classifier using the full set of sensors and features, and then a reduced set of them.

In all experiments, RF model is implemented with 100 decision trees and maximum number of splits equals 10. The latter parameter is selected out of values between 2 and 10. For each target, samples are randomly split into 75% for training and validation, and 25% for testing. A number of 100 independent computer runs are carried out per experiment in order to well characterize the average performance of classification models.

To evaluate the classification performance, the following set of metrics are used: *accuracy* (*Acc*), *precision* (*Pre*), *recall* (*Rec*) and *F-measure* (*F*). They are defined as follows:

$$Acc = \frac{TP+TN}{NS}, \quad (1)$$

$$Pre = \frac{TP}{TP+FP} \quad (2)$$

$$Rec = \frac{TP}{TP+FN} \quad (3)$$

$$F = \frac{2 \times Pre \times Rec}{Pre + Rec} \quad (4)$$

Where *NS* is the total number of training examples, and *TP*, *FP*, *TN*, and *FN* denote the number of true positive, false positive, true negative, and false negative examples of a given class, respectively. The environment of Matlab 2018 is used for the implementation and testing on a machine with core i5, 2.6 GHz CPU and 10 GB RAM. The implementation of RF algorithm follows [23].

A. RF for HSCM Task

In this preliminary experiment, the RF classifier is used with the full set of sensors and features. The average classification rates and implementation issues (model size, training and inference times) for the RF model are presented in Table II. The average rates of all metrics are above 0.99. High accuracy and F-measure ensure the effectiveness and robustness of employed attributes or features together with the classifier.

TABLE II. PERFORMANCE OF RF CLASSIFIER FOR FAULTS OF DIFFERENT TARGETS: COOLER (C), VALVE (V), PUMP (P), AND ACCUMULATOR (A)

	Acc	Pre	Rec	F	Size (KB)	Tr. Time (sec)	Inf. Time (sec)
C	0.998	0.998	0.998	0.998	4028	1.4	0.09
V	0.996	0.995	0.995	0.995	4510	1.76	0.091
P	0.994	0.992	0.992	0.992	4540	1.71	0.09
A	0.992	0.992	0.991	0.992	5399	2.35	0.098

Although the addressed dataset is unbalanced, the proposed model performs equally well for the recognition of different conditions of each target class. This behaviour can be confirmed by looking at the confusion matrix for the average performance of RF classifier for each target as shown in Tables III to VI. For better readability, fractions are truncated.

TABLE III. CONFUSION MATRIX FOR COOLER TARGET

	CF	RE	FE
CF	183	0	0
RE	0	183	0
FE	0	0	185

TABLE IV. CONFUSION MATRIX FOR VALVE TARGET

	CF	SvL	SmL	OS
CF	90	0	0	0
SvL	0	89	0	0
SmL	0	1	89	0
OS	0	0	0	281

As can be seen from Table V, only one observation of “Weak leakage” is misclassified as “Severely leakage” for the Pump target and vice versa. Also, from Table IV, two observations are misclassified for “Slightly reduced pressure” and one for “Severely reduced pressure” of the Accumulator target. From Table II, the size of classifier model is 4028 KB for the Cooler, with an increment of 12%, 12.7% and 32.5% for the Valve, Pump and Accumulator targets, respectively. Such variance may be explained by the different degrees of difficulty for classifying each target conditions. In addition, the training time for cooler model is the fastest with time 1.4 sec, while models of Valve and Pump require about 1.7 sec. The Accumulator model takes more training time with 2.35 seconds. Clearly, the inference time (i.e., testing the model for one observation) is almost the same for any of the four models and equals approximately 0.09 seconds.

TABLE V. CONFUSION MATRIX FOR PUMP TARGET

	NL	WL	SL
NL	305	0	0
WL	0	122	1
SL	0	1	122

TABLE VI. CONFUSION MATRIX FOR ACCUMULATOR TARGET

	CF	SvP	SIP	OP
CF	201	0	1	0
SvP	0	98	0	1
SIP	2	0	97	0
OP	0	0	0	149

Summing up, the proposed RF model alone achieves outstanding recognition rates for all targets. Unlike the work of Quatrini et al. [15], notable performance was achieved by two classifiers, ANN for the Pump with rate 0.998 for the pump target while RF was better for the Cooler, Valve and Accumulator with classification rates of 0.998, 1 and 0.991, respectively.

B. Sensor and Target Correlation

The sentiment analysis of the impact of individual sensors on recognition of severe operating conditions is extensively studied in this section. The features of each sensor are

introduced to the RF model. It might be more important to examine an employed sensor capability of recognizing every probable failure condition ignoring false alarms. Regarding this concern, precision metric is applied in this experiment. Each sensor test is repeated for 100 times in order to build up rigorous conclusions. Then, the highest-precision sensors on average are determined.

For the Cooler target, the classification task seems straightforward in agreement with previous studies [15, 24]. Precision achieved using only one of the following sensors: TS2, CE, TS1, CP, PS6, PS5, TS3 and TS4, exceeds 0.995. The use of one sensor minimizes the model size from 4028 to 651 KB and the training time from 1.4 to 1.04 sec as shown in Fig. 2(a) for the cooler. Similar behaviour is reported for the Valve target. The sensors PS3 and PS2 give Precision rates of 0.997 and 0.991, respectively. The next important sensors are PS1 and FS1 with Precision 0.957 and 0.953, respectively. Fig. 2(b) shows the model performance using the following groups of sensors: G_{V1} (PS3, PS2), G_{V2} (PS3, PS2, PS1) and G_{V3} (PS3, PS2, PS1, FS1). Model size is reduced to 892, 1137, 1432 KB for G_{V1} , G_{V2} and G_{V3} , respectively, instead of 4510 KB for the basic model with full set of sensors.

Pump and Accumulator targets are more challenging. The most effective sensors, when used individually, are FS1, SE, PS1, EPS1 and PS3 with precision 0.981, 0.979, 0.943, 0.941, and 0.929, respectively, for the Pump. The best combination of sensors are FS1, SE, PS1, giving an optimal precision 0.993. In this case, model size occupies only 657 KB and takes 1.15 seconds for training compared to 4540 KB and 1.71 seconds, respectively, when the full set of sensors are used. Fig. 2(c) summarizes the performance of different groups of sensors for Pump target: G_{P1} (FS1, SE), G_{P2} (FS1, SE, PS1), G_{P3} (FS1, SE, PS1, EPS1) and G_{P4} (FS1, SE, PS1, EPS1, PS3).

Similarly, the top 6 effective sensors when used individually, for the classification of Accumulator, are PS3, PS1, SE, FS1, TS1 and PS2 with precision 0.923, 0.88, 0.837, 0.828, 0.802 and 0.8, respectively. Fig. 2(d) illustrates the performance of combining attributes of sensors in the groups: G_{A1} (PS3, PS1), G_{A2} (PS3, PS1, FS1), G_{A3} (PS3, PS1, FS1, SE), G_{A4} (PS3, PS1, FS1, SE, PS2) and G_{A5} (PS3, PS1, FS1, SE, PS2, and TS1). The best precision of 0.986 is obtained using group G_{A5} with model size 1602 KB (compared with 5399 KB for the basic model) and training time of 1.73 sec (compared with 2.35 seconds, for basic model). This experiment reveals that the Accumulator target conditions are the hardest for classification.

In summary, this experiment justifies, to a great extent, the preliminary discrimination of target classes in this dataset into easy (Cooler and Valve) and hard (Pump and Accumulator) classifiable classes [21]. For the Cooler and Valve targets, it is sufficient to apply only one sensor for monitoring the different conditions of each one. For the pump target, however using FS1 alone can achieve precision of 0.981 but the group G_{P2} (FS1, SE, PS1) improves it to 0.993. Focusing on the most effective sensors can optimize the RF model in terms of model size and training time. Finally, several sensors are needed to give an acceptable recognition precision of Accumulator conditions. The sensors G_{A5} (PS3, PS1, FS1, SE, PS2, TS1)

can achieve a high precision of 0.986. However, this is still below the precision 0.993 of the basic model which employs all sensors.

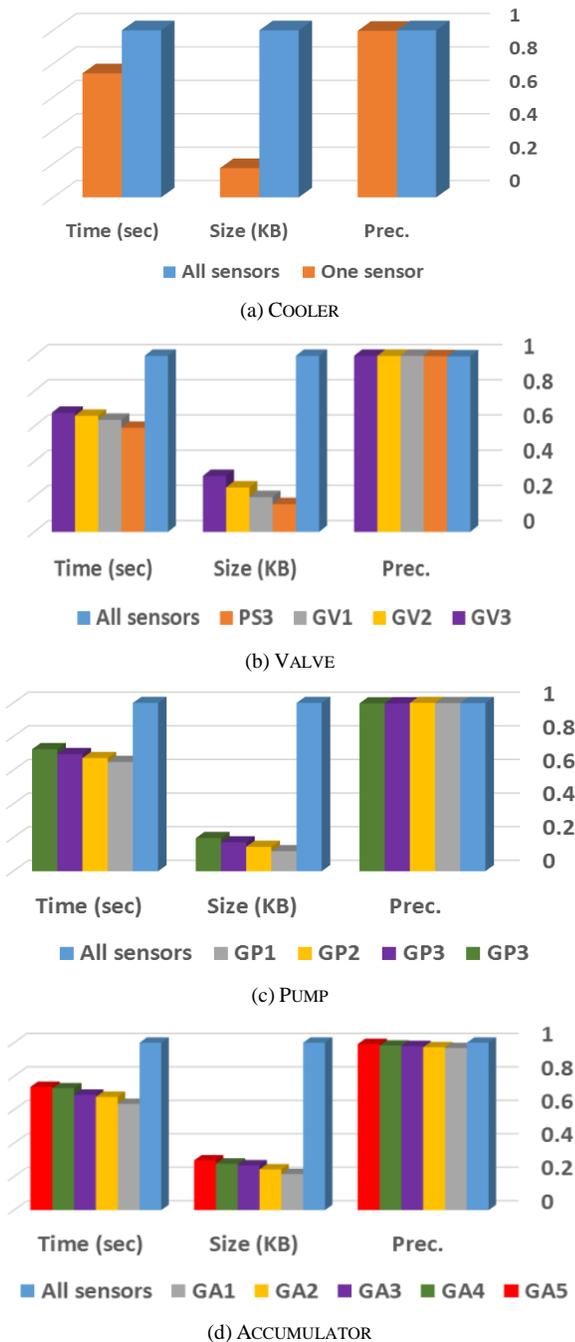


Fig. 2. Normalized Numerical Results of Sensor Correlation Tests for Different Targets.

C. Feature Effectiveness

Based on the findings of the previous experiment, it is interesting to investigate the most effective attributes per sensor. Therefore, in section, only the most effective sensors for each target are considered. The wrapper-based approach is followed where each individual attribute is provided for the classifier. Thus, a direct judgement of the discriminating power of each attribute is obtained. For this purpose, RF model using 5 decision trees with a maximum number of split equals 2 is sufficient.

Fig. 3(a) shows that only the mean of the CE sensor (CE_m) can classify the conditions of the Cooler target with average precision of 0.998. Also, the mean of CP sensor (CB_m) can achieve precision 0.992. Moreover, the mean of each of PS5, PS6, TS1 and TS2 results in precision exceeding 0.96. Slop of linear fit of the TS1 sensor is also useful achieving a precision of 0.942.

Other effective features are determined for the Valve condition. The kurtosis and skewness of PS2 (PS2_k and PS2_sk) give precision of 0.994 and 0.972, respectively. Position of the maximum of PS3 (PS3_p) achieves 0.99 precision. Fig. 3(b) shows the effectiveness of attributes of PS2 and PS3 denoted by G_{V1} and give the highest rate for the Valve target.

Attribute effectiveness for the most useful sensors for the Pump target, namely G_{P2} (FS1, SE, PS1), is presented in Fig. 3(c). The mean of SE (SE_m) gives 0.952 and the mean of FS1 gives 0.924 precision. The rest of attributes achieve lower rates, in particular the position of the maximum of each of FS1 and SE (FS1_p and SE_p) are definitely useless for Pump target class.

For the Accumulator target, no individual attribute of the sensors group G_{A5} (PS3, PS1, FS1, SE, PS2, TS1) can exceed a precision level of 0.7 as shown in Fig. 3(d). The mean of each of TS1 and PS1 are the highest two attributes with precision 0.683 and 0.59, respectively. Such observation shows again that recognition the conditions of the Accumulator target is harder than others in this application. Also, the position of maximum of each of FS1 and SE (FS1_p and SE_p) seems useless for this target class.

Summarizing these findings, it is interesting to discover that only one attribute of one sensor can be efficient for the addressed classification task in this work for some targets. The mean of CE sensor (CE_m) and the kurtosis of PS2 (PS2_k) achieve precision of 0.998 and 0.992 for Cooler and Valve targets, respectively. Conversely, some features are useless for classification such as (FS1_p and SE_p) for both Pump and Accumulator targets.

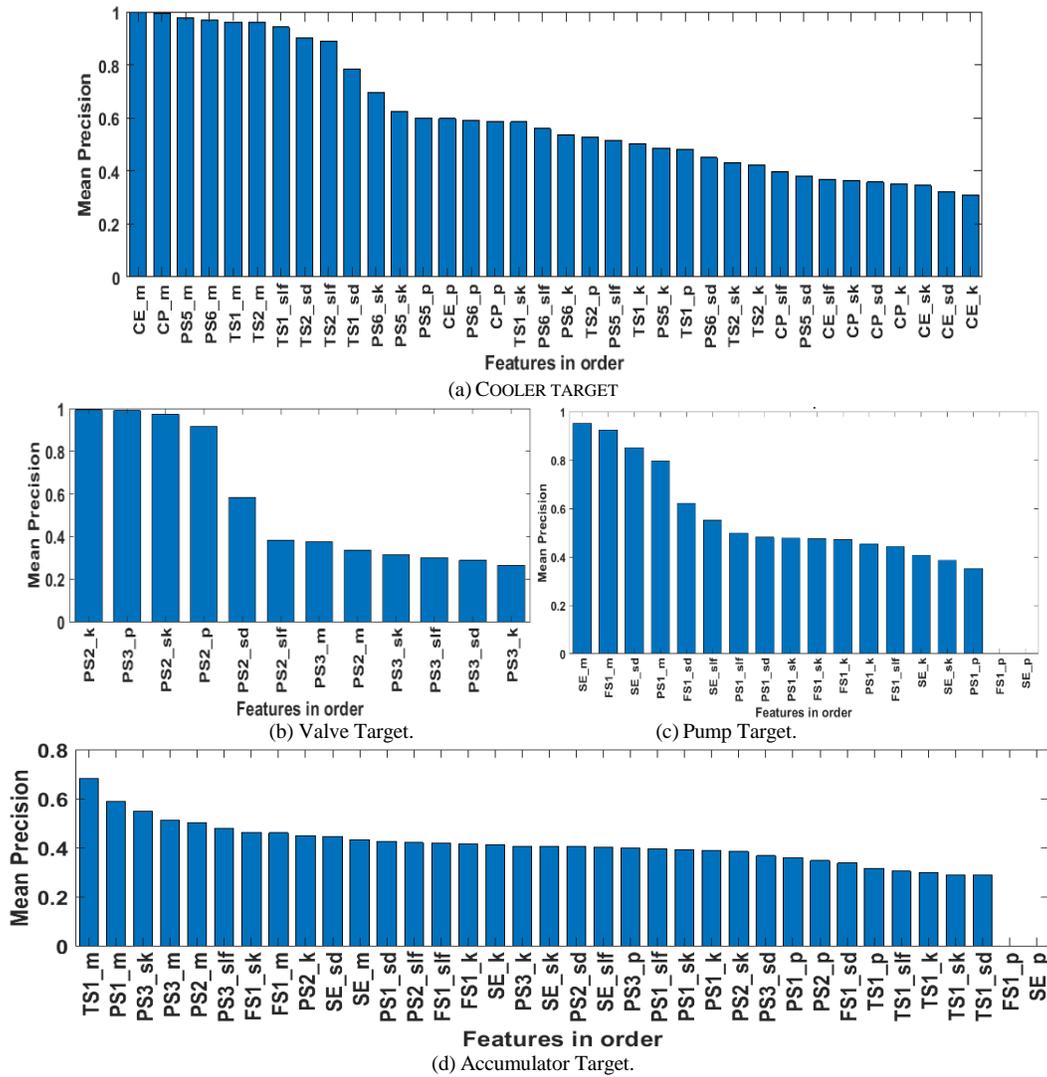


Fig. 3. Effective Attributes for each Target Class.

V. DISCUSSION AND LIMITATIONS

Condition monitoring of hydraulic systems via sensors fixed inside the system is suspicious to some hazard situations if one or more sensors become out of service. Thus, the importance of current study stems from investigating the role of each sensor in the assigned classification task. Besides, emphasizing the most effective attributes leads to optimizing the classification model size and training time.

Table VII summarizes the outcomes of this study. The optimized RF model for the Cooler target uses only one sensor of (TS2, CE, TS1, CP, PS6, PS5, TS or TS4), keeps performance (denoted by + in Table VII) of basic model that uses all sensors. It follows that model size and training time are reduced by 83.8% and 25.7%, respectively.

TABLE VII. SUMMARY OF EFFECTIVE SENSORS AND FEATURES FOR EACH TARGET, THE IMPACT ON MODEL PRECISION (PRE. ↑), PERCENTAGE REDUCTION IN MODEL SIZE (RS) AND TRAINING TIME (RT)

	Sensors	Features	Pre.	Pre. ↑↓	Size (KB)	RS(%)	Tr. Time (sec)	RT(%)
Cooler	TS2, CE, TS1, CP, PS6, PS5, TS, TS4	Mean	0.995	+	651	83.8	1.04	25.7
Valve	PS3, PS2	Kurtosis, Skewness	0.997	+	892	80.2	1.12	36.4
Pump	FS1, SE, PS1	Mean	0.993	+	657	85.5	1.15	32.7
Accumulator	PS3, PS1, SE, FS1, TS1, PS2	Mean	0.986	-	1602	70.3	1.73	26.4

For the Valve target, the RF model, using PS3 and PS2, is as efficient as basic model with reduction in model size by 80.2% and training time by 36.4%. Using only FS1, SE and PS1 attributes can achieve slightly better performance than the basic model for the Pump target with reduction ratios of 85.5% and 32.7% for model size and training time, respectively. A different finding is reported for the Accumulator target. Using a reduced set of sensors such as PS3, PS1, SE, FS1, TS1 and PS2 might result in performance degradation (marked with “-” in Table VII). However, model size reduction ratio reaches 70.3% and training time becomes 26.4% less than using all sensors. For the attributes, it is figured out that simple time-domain features such as mean, kurtosis and skewness are very useful and efficient for such classification problem.

It is important to emphasize that, in this work, sensor role and feature importance is studied from a pure machine learning point of view. It is interesting to interpret the validity of the results obtained with the aid of an expert of such hydraulic plant. Moreover, the proposed model, and its optimized versions, should be tested in real environment where ad-hoc devices like PLC units are in charge for monitoring the system conditions. It also lacks to consider the effect of noise on recorded sensor signals, in particular low-frequency sensors.

VI. CONCLUSIONS AND FUTURE WORK

The employment of machine learning techniques for hydraulic system condition monitoring proves effective for automatic recognition of faults and severe conditions. It is common to fix various sensors in the system in order to collect enough readings for different operating conditions. The considered system in this study is provided by 17 (14 physical + 3 virtual sensors) for measuring quantities such as motor power, volume flow, pressure, temperature and vibration. These sensor signals are represented by a set of six simple time-domain attributes to classify four targets, namely, Cooler, Valve, Pump and Accumulator. The random forest classifier is very suitable for such classification task and its performance exceeds 99% for all targets. Moreover, the conducted experimental work reveals the impact of each sensor in classification of each target conditions. Using of all sensors is not essentially effective and efficient. Only one temperature sensor is sufficient for the Cooler conditions classification. The same observation holds for the Valve target where only two pressure sensors are sufficient. Interestingly, the volume flow, pressure and efficiency factor sensors can achieve better recognition rate than using all sensors for the Pump target.

On the contrary, for the Accumulator target, the use of the attributes of all sensors looks mandatory in order to achieve high performance. However, using a reduced set of pressure, volume flow, temperature, and efficiency factor sensors still gives acceptable classification rate for this target.

Using few sensors optimizes the classification model size and training time and, furthermore, minimizes the cost of purchasing and maintenance of many sensors while some of them can be sufficient. It is worth to investigate the applied methodology here for other similar applications to determine the most important sensors for a given fault.

The effect of noise commonly present in sensor measurements on the classification model is challenging and can be investigated in a future work. Moreover, noting that the classification performed in the current work employs windows of data of size 60 seconds, it is a reasonable extension to study the possibility of using shorter windows e.g. 5, 10, 20 or 30 seconds. This can have significant effect on the quick detection of severe fault conditions.

REFERENCES

- [1] Z. Li, Y. Wang, and K. Wang, “Intelligent Predictive Maintenance for Fault Diagnosis and Prognosis in Machine Centers: Industry 4.0 Scenario,” *Advanced Manufacturing*, Vol. 5, pp. 377–387, 2017.
- [2] T. P. Carvalho, F. Soares, R. Vita, R. Francisco, J. Basto, and S. Alcalá, “A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance,” *Computers & Industrial Engineering*, 137, 106024, 2019.
- [3] W. Zhang, D. Yang, and H. Wang, “Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey,” *IEEE Systems Journal*, Vol. 13, No. 3, pp. 2213–2227, 2019.
- [4] Y. Li, J. X. Gu, D. Zhen, M. Xu, and A. Ball, “An Evaluation of Gearbox Condition Monitoring Using Infrared Thermal Images Applied with Convolutional Neural Networks,” *Sensors*, Vol. 19, No. 9, 2205, 2019.
- [5] L. Jing, M. Zhao, P. Li, and X. Xu, “A Convolutional Neural Network Based Feature Learning and Fault Diagnosis Method for the Condition Monitoring of Gearbox,” *Measurement*, Vol. 111, pp. 1–10, 2017.
- [6] P. Wang Ananya, R. Yan, and R. X. Gao, “Virtualization and Deep Recognition for System Fault Classification,” *Journal of Manufacturing Systems*, Vol. 44, pp. 310–316, 2017.
- [7] M. Xia, T. Li, L. Xu, L. Liu, and C. W. de Silva, “Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks,” *IEEE/ASME Transactions on Mechatronics*, Vol. 23, No. 1, pp. 101 – 110, 2018.
- [8] L. Wen, X. Li, L. Gao, and Y. Zhang, “A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method,” *IEEE Transactions on Industrial Electronics*, Vol. 65, No. 7, pp. 5990–5998, 2018.
- [9] A. Osman, M. H. Gobran, and F. F. Mahmoud, “Vibration Signature of Gear Pump of Missing One and Two Teeth,” *Egyptian Journal for Engineering Sciences and Technology*, Vol. 30 (Mechanical Engineering), pp.39–50, 2020.
- [10] A. Osman, A. Salman, and K. Fawzy, “Vibration Signature of Misaligned Rotors of Centrifugal Pump,” *Egyptian Journal for Engineering Sciences and Technology*, Vol. 27, pp.30–42, 2019.
- [11] N. Helwig, S. Klein, and A. Schützea, “Identification and Quantification of Hydraulic System Faults Based on Multivariate Statistics Using Spectral Vibration Features,” *Procedia Engineering*, Vol. 120, pp. 1225–1228, 2015.
- [12] A. Kothurua, S. P. Nooka, and R. Liu, “Application of Deep Visualization in CNN-based Tool Condition Monitoring for End Milling,” *Procedia Manufacturing*, Vol. 34, pp. 995–1004, 2019.
- [13] F. Ahmadzadeh and J. Lundberg, “Application of multi regressive linear model and neural network for wear prediction of grinding mill liners,” *International Journal of Advanced Computer Science and Applications*, Vol. 4, No.5, pp. 53–58, 2013.
- [14] Y. El Mourabit, Y. El Habouz, H. Zougagh, Y. Wadiai, “Predictive System of Semiconductor Failures based on Machine Learning Approach,” *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 11, No. 12, pp. 199–203, 2020.
- [15] E. Quatrini, F. Costantino, C. Pucci, and M. Tronci, “Predictive Model for the Degradation State of a Hydraulic System with Dimensionality Reduction,” *Procedia Manufacturing*, Vol. 42, pp. 516–523, 2020.
- [16] X. Liu, “Study on Knowledge -based Intelligent Fault Diagnosis of Hydraulic System,” *Telkommnika*, Vol.10, No.8, pp. 2041–2046, 2012.

- [17] A. El-Betar, M. Abdelhamed, A. El-Assal, and R. Abdelsatar t al., "Fault Diagnosis of a Hydraulic Power System Using an Artificial Neural Network," *JKAU: Eng. Sci.*, Vol. 17, No. 1, pp. 117–137, 2006.
- [18] H. L. Tian, H. R. Li, and B. H. Xu, "Fault prediction based on data fusion," in *Advanced Material Research*, 2013, Vol. 712–715, pp. 2084–2088.
- [19] A. Steinboeck, W. Kemmetmüller, C. Lassl, and A. Kugi, "Model-based condition monitoring of an electro-hydraulic valve," *J. Dyn. Syst. Meas. Control. Trans. ASME*, Vol. 135, No. 6, 2013.
- [20] R. Jegadeeshwaran and V. Sugumaran, "Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines," *Mech. Syst. Sig. Process.*, Vol. 52, pp. 436–446, 2015.
- [21] N. Helwig, E. Pignanelli, and A. Schütze, "Condition Monitoring of a Complex Hydraulic System Using Multivariate Statistics," *IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Pisa, Italy, pp. 210–215, 2015.
- [22] S. S. Chawathe, "Condition Monitoring of Hydraulic Systems by Classifying Sensor Data Streams," *IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, pp. 0898–0904, 2019.
- [23] Breiman, L. "Random Forests." *Machine Learning*, Vol. 45, pp. 5–32, 2001.
- [24] C. König and A. Helmi, "Sensitivity Analysis of Sensors in a Hydraulic Condition Monitoring System Using CNN Models," *Sensors*, Vol. 20, No. 11, 3307, 2020.

A New Communication Protocol for Drones Cooperative Network: 5G Site Survey Case Study

Youssef Shawky Othman¹, Dr. Mohamed Helmy
Megahed², Prof. Dr. Mohamed Abo Rezk³
Computer Science
Arab Academy for Science and Technology and Maritime
Transport, Cairo, Egypt

Prof. Dr. Fathy Ahmed Elsayed Amer⁴
Prof. Dr. Comp. Sci. Dep
College of Info. Sys., Comp. Sci., Oct. 6, Univ
Cairo, Egypt

Abstract—Directing the antennas of the 5th generation mobile network optimally became a hard and tedious process due to the abundance of the antennas that the 5th generation networks relay on. Also, due to the traditional way of measuring the signal strength of the 5th generation networks. The process that could take weeks of working until it is done. So, the solution is to make an automated process to measure signal strength and to direct antennas using "Drones" rather than human power. This way will ease the process of directing antennas in a quite shorter amount of time which will be some hours instead of weeks. Plus the low cost and the high accuracy achieved. So, a cooperative network between "Drones" and a new communication protocol to support that network will be designed. "Drones" will communicate with each others and with antennas through exchanging messages by "MQTT cloud" using new designed communication protocol. "Raspberry pi" platform will be used as a server to control the direction of antennas. "Drones" will carry a "4G mobile" and a "Raspberry pi" with a "Building Identification System (BIS)" setup on it. The "(BIS)" gives every building a number and will recognize the entrance of buildings to measure signal strength at every floor. Then, performance analysis metrics (Throughput) will be measured. OMNeT++ "will be used for simulation", and "Raspberry Pi" platform will be used to implement the system and measure the performance of the new Communication Protocol.

Keywords—Drones; communication protocol; Message Queuing Telemetry Transport (MQTT); cooperative network

I. INTRODUCTION

The fifth generation of mobile communication systems aims to provide a ubiquitous mobile service with better quality. Automobiles, public transportation, medical care, energy, public safety, agriculture, entertainment, manufacturing, and other vertical industrial applications are expected to benefit from this technology. User density, traffic volume, and data rate are all expected to increase dramatically. In the coming decade, innovative solutions to the needs of both smartphone users and vertical industries are needed.

There are many challenges to realize 5G networks, i.e., high system capacity (1000× capacity per km²), high data rates (targeting 1 Gbps per user everywhere and 100× user throughput increase) and massive connectivity (100× connected users) [1]. The advent of fifth generation (5G) technology will accelerate the growth of V2V communications. A recent report from Juniper Research predicted that as 5G

technology provides lower latency and high range, automotive Original Equipment Manufacturer (OEMs) will use it as the main technology for V2V communications over other technologies [2].

The traditional way that is used now in The fifth generation (5G) of mobile communication system to measure the coverage cell phone network signal strength in cellular base stations is through the human power using smartphones, where specialists go to the antennas above the base stations to direct it. The different teams of specialists are distributed over the whole places and buildings that is under the coverage of the base station. Then they take the measure of signal strength via a mobile application like "Network cell info lite" then they communicate with the specialists on the base station and deliver the measures taken to them so that they can direct the antennas according to what they received. This process that repeated over and over again in an attempt to get the optimal direction of the antennas. But finally it doesn't get achieved for several reasons.. First is the high cost. Second is the high consumption of time "that could take days and sometimes a few weeks" and third is the lack of accuracy from the human power. Besides some changes that may happen to the coverage places in terms of the buildings and its height or the continuous variation in the number of users of the network. Which leads us to repeat the whole process many times in small period of times. Which is a big obstacle to telecommunication companies.

With advances in technology and commercially available vehicles, the interest is shifting toward collaborative drones. cooperation necessitates effective communication as well as a sound strategy for avoiding obstacles and drone collisions. Considering that drones must swap positions in order to maintain the swarm formation, cover a large area, and avoid colliding [3].

Several projects searched into the challenges of designing UAV systems for various applications. The general design principles of a multi-UAV system in civil applications are an open issue and are still being researched [4].

According to all the mentioned before, it becomes necessary to think & research to find an alternative way (using "Drones" to direct antennas) instead of the traditional way that has lots of disadvantages like the high cost, the high consumption of time and the lack of accuracy.

These disadvantages that led to the weak signal or no signal at all in mobile networks due to the poor distribution of antennas and not directing them optimally that finally make the mobile users suffer when they use the network specifically in higher floors in skyscrapers. So to avoid all these problems, the proposal is:

Replacing the human power with "Drones" by creating a cooperative network between each others. As "Drones" will be sent to the buildings and places that are under the coverage to take the measures of the signal strength. And by the connection between the "Drones" and the "Server" placed on the base station, the measures will be sent to it through MQTT Cloud (or off load when facing no connection situation). Then, by the "Servo motors" the antennas will be directed to the optimal angle.

As a result the solution presented will lead us to avoid all the disadvantages mentioned before there for there is no chance to face any high cost caused by repeating the process in the traditional way.

Contributions :

- Building Cooperative Network Between Drones.
- Building New Communication Protocol To support Cooperative Network Between Drones
- Design and implementation for mobile application on MQTT cloud
- Make a Hardware Implementation using raspberry pi for receiving the data from drones to do tilting of antenna.

Outlines of paper:

Section II presents "Related work" (Cooperative network of drones, 5G, ordinary site survey, image processing, mobile application, MQTT, hardware implementation). Section III presents "Overall Proposed model". Section IV presents "Drones Cooperative network". Section V presents "Proposed new communication protocol". Section VI presents "Mobile application". Section VII presents "Building Identification System (BIS) and Directing Antenna System (DAS)". Section VIII presents "Simulation results and Performance analysis". Section IX presents "Hardware implementation". Section X presents "Comparison with other works". Section XI presents "Conclusion and future work".

II. RELATED WORK

The rapid growth of Unmanned Aerial Vehicles (UAVs) for various applications such as package delivery, inspection, defence, and disaster relief has opened up a plethora of commercial business opportunities. For successful use of UAVs for various applications, it is critical that the vehicles be configured to operate autonomously and cooperatively.

In this section, related work will be reviewed. Also the Cooperative network of drones, 5G, ordinary site survey, image processing, mobile application, MQTT, and hardware implementation of Cooperative Network of Drones will be reviewed.

In [3] Bekhti, et al. presented The application of a swarm of drones to the intrusion detection and tracking problem. They devised an algorithm to keep the drones in a team formation in order to address the issue of coverage and collisions. Omnet simulator was used to test the algorithm in various scenarios.

In [5] Gu, et al. they proposed a cooperative network platform and system architecture of multi-UAV surveillance, they discussed the establishment of suitable algorithms based on machine learning.

In [4] presented by Yanmaz, et al. they proposed drones are being utilized in monitoring, transport, safety and disaster management, other domains. they described a high-level architecture for the design of a collaborative aerial system consisting of drones with on-board sensors and embedded processing, coordination, and networking capabilities.

In [6] Dong, et al. they studied optimizing the deployment density of Drone Small Cells (DSCs) to achieve the maximum coverage performance, they propose an algorithm to get the optimal deployment density with low complexity. they conducted both field experiments and Matlab simulations to verify the correctness of theoretical analysis.

In [1] Nirwan and Liang they investigated the problem of Placement and communications in the Drone-mounted base-stations (DBS) to provide ubiquitous connections and high spectrum efficiency.

In [7] Chakrabarty, et al. they simulated a complete urban operations in a high fidelity simulation environment. they designed a V2V communication protocol and all the vehicles "Drones" participating communicate over this system.

In [8] Manasa, et al. they discussed a new approach to the wireless communication between Ground Control Station (GCS) to Micro Air Vehicle (MAV) and vice versa.

In [9] Iranmanesh, et al. they discussed congestion problem in cellular networks through the assistance of parcel delivery drones. To this end, they proposed a novel algorithm, called CARLO.

In [10] Ma, et al. they utilized drones as air routers to establish a LAN, to collect information from pipeline networks and transmit it to pipeline inspectors. to achieve optimal drone deployment. A two-phase evolution optimal 3-D drone layout algorithm was proposed to deploy drones.

In [11] Buksz, et al. they presented an approach for autonomous radio base station inspection using cooperative drones based on Intent-driven Strategic Tactical Planning (ISTP).

In [12] Lee. He proposed a cooperative drones positioning measuring in internet-of-drones in order to obtain high accuracy positions for both egodrones and obstacles.

In [13] Bartolini, et al. They work addressed the problem of assigning location based tasks to a fleet of drones in an emergency critical scenario, to ensure early inspection of target locations.

In [14] Pu, and Logan Carpenter, they investigated the problems of service scheduling for drones and Zone Service

Providers (ZSPs) in the Internet of Drones. For performance evaluation, they conducted extensive simulation experiments using OMNeT++.

In [15] Guan, et al. they studies distributed algorithms for controlling self-organizing flying Unmanned Aerial Vehicles with massive MIMO networking capabilities, dubbed mDroneNet. The distributed solution algorithm converges in tens of iterations and can achieve around 90% of the global optimum, according to the Results.

In [16] Verri, et al. they discussed the drones' overall performance. They proposed a different approach to the problem of airspace route planning. They used a simplified simulation model.

In [17] Haas, and Zhong, they discussed the collaborative communication of multiple drones, as to create a reliable and powerful communication links, which as part of the network (they called "Network in the Sky").

In [18] Valianti, et al. they proposed Using a swarm of pursuer drones to track down and apprehend one or more rogue drones. Extensive performance assessment results show that the problem can be solved efficiently and effectively using the evolved distributed algorithm.

In [19] Khosravi, et al. they proposed designing drones trajectories that efficiently perform some transportation operation (for example, package delivery) while also providing uniform coverage over a neighbourhood area, which is required for applications such as network coverage, data collection from Internet of Things devices, wireless power transfer, and surveillance. They proposed a trajectory process that would allow for uniform coverage while maintaining transport (delivery) efficiency.

In [20] Dao, et al. they studied the Aerial radio access networks (ARANs) as a potential strategy to supplement existing terrestrial communication systems. Unmanned aerial vehicles, drones, and satellites are among the airborne components involved. The development of seamless mobile communication systems is expected to be aided by ARANs as part of a comprehensive sixth-generation (6G) global access infrastructure.

In [21] Real, et al. they presented a group of multiple Drones that would work together on autonomous construction missions. They tested and simulated the system's performance.

In [22] Huang, et al. they illustrated how to navigate a group of unmanned aerial vehicles (UAVs) to monitor traffic on a road. They tested their method using computer simulations.

A. 5g Mobile Communication System

In [23] Habibi, et al. they discussed the fifth generation of mobile communication systems aims to provide a ubiquitous mobile service with better quality (QoS). Automobiles, public transportation, medical care, energy, public safety, agriculture, entertainment, manufacturing, and other vertical industrial applications are expected to benefit from this technology. User density, traffic volume, and data rate are all expected to increase dramatically. In the coming decade, innovative

solutions to the needs of both smartphone users and vertical industries are needed, as shown in Fig. 1.

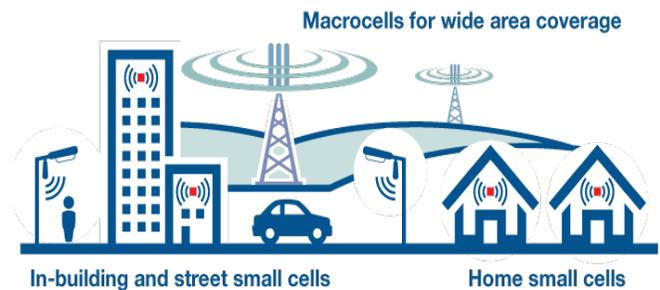


Fig. 1. Abundance of the Antennas that the 5th Generation Networks Relay on.

B. Ordinary Site Survey

The human power is the main factor in the traditional way of measuring the signal strength of the 5th generation mobile networks. Lots of specialists work using their smartphones (using mobile application like "Network Cell info Lite") [23]. They are all responsible for making that process get done. Like going to all the places and coverage areas, measuring the signal strength and even directing the antennas on the base stations by themselves. They communicate with each other to send and receive the measures and the required data to control and direct the antennas.

So according to all that lots of problems occurred through this process like the high cost, high consumption of time and the lack of accuracy. Plus the process needs to get repeated so many times in small period of time to reach the direction required. Which is not achieved at the end.

C. Image Processing

In [24] Chen, et al. they framed this drone-view building identification as building retrieval problem: given a building (multimodal query) with its images, geolocation and drone's current location, they aim to retrieve the most likely proposal (building candidate) on a drone-view image. Despite few annotated drone-view images to date, there are many images of other views from the Web, like ground-level, street-view and aerial images. Thus, they proposed a cross-view triplet neural network to learn visual similarity between drone-view and other views.

D. Mobile Application (To Measure Signal Strength)

Will be measure the signal Strength by "Network Cell Info lite app" [23]. It will measure the signal strength and display the location of your signal on a map.

E. MQTT Protocol / MQTT Cloud

In [25] Kang, et al. they discussed MQTT protocol operation. The basic concepts of it is publish/subscribe and client/broker and its basic functionality is connect, publish, and subscribe.

In [26] Mishra and Attila, they discussed MQTT protocol, For message delivery, MQTT provides three levels for quality of service (QoS 0 - QoS 1 - QoS 2). The design principles of this protocol focus on minimizing network bandwidth and device resource requirements ensuring reliable delivery.

In [27] presented by Krishna, et al. they discussed MQTT Dashboard, a mobile app that runs on a phone and can be used to track data sent from sensor nodes using the MQTT protocol. MQTT Dash is another choice. M2M, Arduino, Raspberry Pi, Microcontrollers, and other devices are supported by the application.

F. Hardware Implementation

In [28] Mathur, et al. proposed a novel approach to digital image processing by using the histogram of oriented gradients (HOG) features descriptor with the OpenCV library, which was programmed in Python and booted with the Raspberry Pi, which was equipped with a RaspyiCam to capture moving images of objects passing under the cam.

III. OVERALL PROPOSED MODEL

As shown in Fig. 2. "Proposed Model" the human power has been replaced with "Drones" by creating a cooperative network "that based on new communication protocol between drones. A Cooperative Network between Drones will be built and a "4G Mobile" will be set up and a "Raspberry pi" will be set up on every drone. digital map will be used with drones so drones are ready to be sent to the locations to take the measures of the signal strength. All buildings will be identified by the " (BIS) ". The signal strength will be measured by a mobile application "Network Cell Info Lite" set up on the "4G mobile". Data will be sent from "Drones" to "Server" and vice versa using Cooperative network based on a new communication protocol and also the messages will be sent between "Drones" and "Server" and vice versa by "MQTT Cloud". So, Finally the "Server" will control the direction of the antennas using "Servo motors".

The Model Structure:

1) *Building cooperative network between drones:* "Drones" will communicate with each others and Data will be sent from "Drones" to "Server" and vice versa using Cooperative network based on our new designed communication protocol, as shown in Section 5.

2) *Building new communication protocol:* The Drones communicate with the Server or other Drones is realized in defined frame structure of new communication protocol supporting cooperative network, which consists of (payload data, header) for identification. The frame structure consists of the (header, payload) accompanied by the checksum. The protocol is a designed communication protocol to support the cooperative networks between drones to be used in 5G site survey.

3) *Mobile app SW (network cell info lite):* It will measure the signal strength and display the location of your signal on a map.

4) *Building Identification System (BIS) on drones:* As shown in [24]. The work of the "BIS based on CNN" is to give every building a specific number to recognize the entrance of the building to measure signal strength at every floor.

5) *Directing antenna system (DAS):* The (DAS) will be set up on "Raspberry pi" and programmed using "Python language". It will receive messages "with measures "from "Drones" and according to these measures, the parameters of the angles of the antennas will be given to the three groups of "Servo motors" so that the antennas will be directed to the optimal direction.

6) *MQTT (broker/client - publisher/subscriber):* "Drones" will communicate with each others and with the antennas through exchanging messages by "MQTT cloud".

7) *Drones:* The "4G Mobile" and "Raspberry pi" will be set up on Drones, then the signal strength measured will be sent to the base station by the drones.

8) *Phones:* Setup the "4G Mobile" on Drones and Setup a mobile App (Network Cell Info Lite – Mobile & Wi-Fi Signal) on smart phone "4G Mobile" to measure signal Strength.

9) *Raspberry Pi:* This versatile single board machine can be used for a variety of tasks. two "Raspberry pis" will be used : First with an (DAS) and second with an (BIS).

10) *Servo motors:* The designed system requires maximum precision, in 5G, servo motors will be used to control antenna direction. The Server will give orders to the three groups of "Servo Motors" to control the direction of the antennas (every group consists of 2 servo motors to move one antenna "Pan1, Tilt1").

As shown in Fig. 3. "Proposed model flow chart" A Cooperative Network will be built between Drones and a "4G Mobile" and a "Raspberry pi" will be set up on all drones. The digital map will be used on Drones so they are ready to be sent to their locations for taking the measures of the signal strength. The "(BIS)" will identify all buildings. Then the signal strength will be measured by a mobile application like "Network Cell Info Lite" which set up on the "4G mobiles". Data will be sent from "Drones" to "Server" and vice versa by Cooperative network based on a new communication protocol and also the messages will be sent between "Drones" and "Server" and vice versa by "MQTT Cloud". So at the end the "Server" will control the direction of all antennas using "Servo motors". And the process will be repeated again and again till achieve our target.

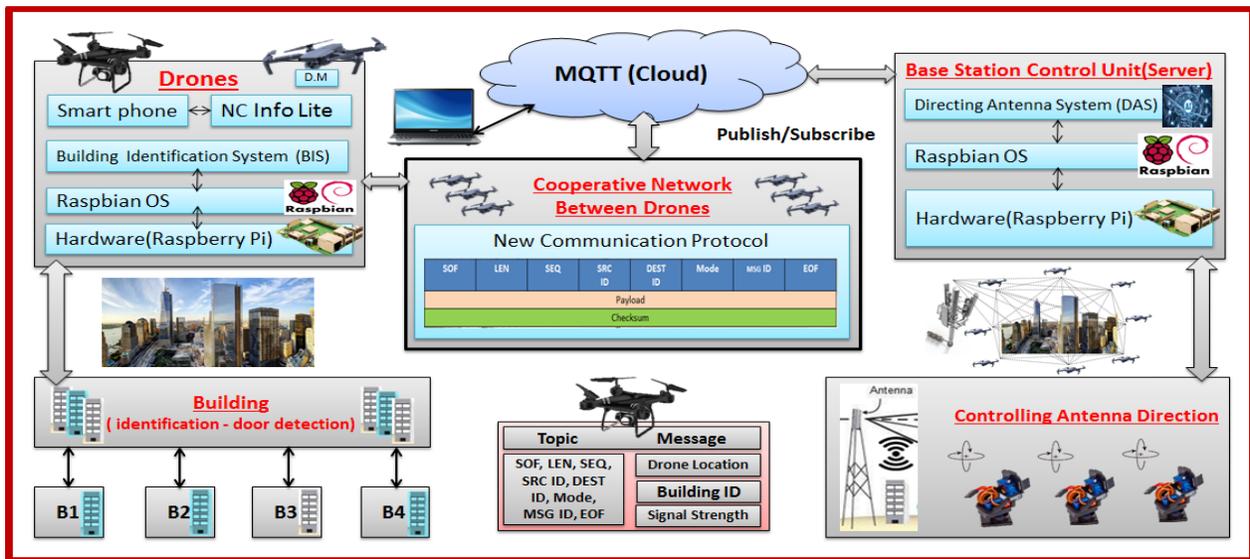


Fig. 2. Proposed Model.

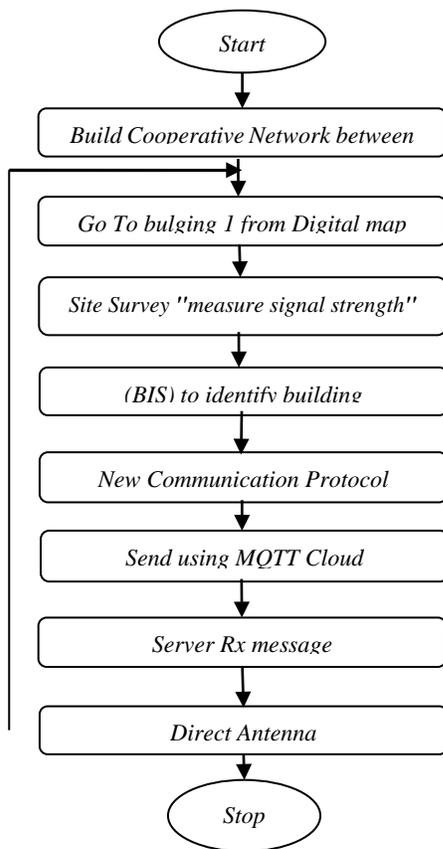


Fig. 3. Proposed Model Flow Chart.

IV. DRONES COOPERATIVE NETWORK

As shown in Fig. 4, the Cooperative Network: (Peer to Peer) in a P2P network, the "peers" are Drone which are connected to each other via the Cooperative network. Drones can be connected directly with other drones in the network without the need of a Master through MQTT Cloud through an internet connection.

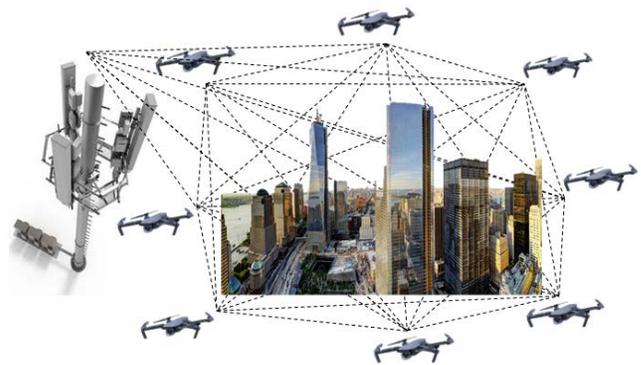


Fig. 4. Peer-to-Peer Cooperative Network between Drones.

V. PROPOSED NEW COMMUNICATION PROTOCOL

- In this paper, a comprehensive solution was proposed by design a new communication protocol to support the cooperative networks between drones to be used in different fields and applications."Drones" will communicate with each others and with the antennas through exchanging messages by "MQTT cloud".
- MQTT has three components (Publisher "Drones/Server", Subscriber "Drones/Server", Broker MQTT Cloud).
- Mosquitto MQTT is used as a broker.
- The Drones communicate with the Server or other Drones is realized in defined frame structure of new communication protocol supporting cooperative network, which consists of (payload data, header) for identification. The frame structure consists of the (header, payload) accompanied by the checksum.
- The frame structure of the new communication protocol is shown in Fig. 5.

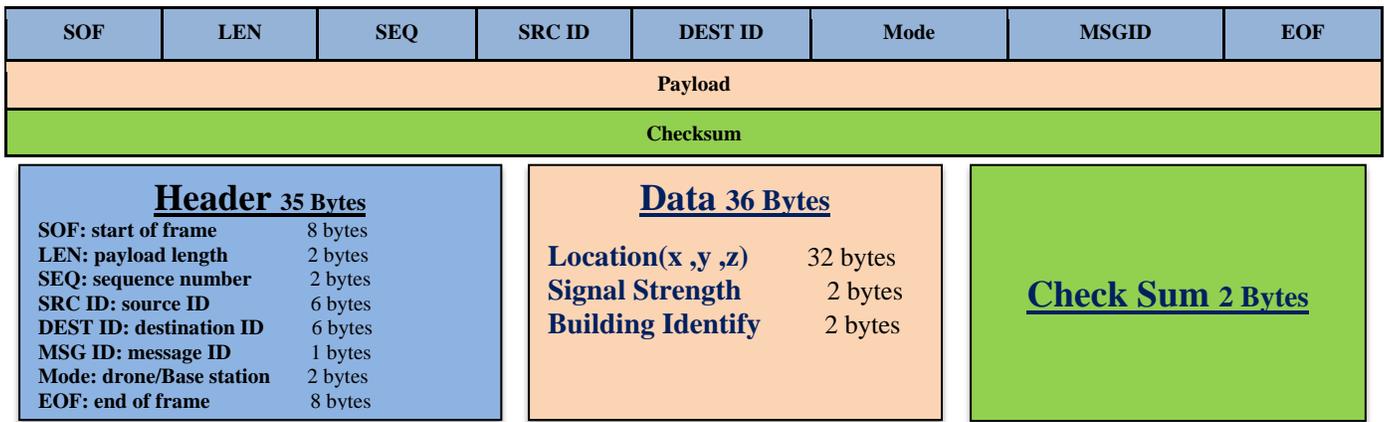


Fig. 5. Communication Frame Structure.

The frame structure fields will be explained in the following below:

a) *Header*: The protocol is including a header that contains all of the information about the payload data that must be exchanged between the Drones and the Server (Base Station antenna's) or Drones and Drones. The header is of 35 bytes, it has a set length and consists of the fields mentioned below.

- SOF: In continuous transmission, this Start of Frame implies the beginning of a frame. The field length is 8 Bytes. which indicates the beginning of new packet.
- EOF: In continuous transmission, this End of Frame indicates the end of a frame. The field is 8 bytes long. This indicates that the previous frame has come to an end.
- LEN: The length of the payload is indicated by this value. (Location(x,y,z)/Signal Strength/Building Identify) will be the payload, payload length max 7200 bytes which uses 2 bytes, (signal strength is 2 bytes and Building Identity is 2 bytes and GPS data is 32 bytes, taking 200 readings for each building with total of $36 * 200 = 7200$ bytes), The original data or message to be sent is referred to as the payload.
- SEQ: The 2 byte Sequence The long message is divided into a number of packets and numbered, allowing it to be assembled at the receiving end.
- SRC ID: Source identification is used to determine the system's uniqueness among the various Drones that are connected to one another. source ID has a length of 6 bytes.
- DEST ID: The receiving destination's address is stored in the Destination ID field, destination ID represents Drone/Server number. The length of the destination ID is 6 bytes. By including the data's destination address, it can be ensured that the data intended for a specific drone arrives there and others drone may reject them.
- MSG ID: 1 byte Message Identification of the parameter is used to distinguish parameters from the many that are available (control msg/cooperative

network/Drones and the Server). Each of these parameters has a unique ID prefix, which aids in faster and more reliable communication.

- MODE: send to drone or base station 2 bytes, indicating the number of Drones in swarm flight .

b) *Payload*: This is the original flight data, Payload of one data information is : 36 bytes.

(Navigation information of drone and camera information is not included in new communication protocol).

c) *Checksum 2byte*: This is used to verify the data's integrity. At the transmitter, the frame's checksum is calculated and appended to the frame at a specific location in the frame structure. The checksum for the received frame will be calculated and compared to the value that comes with the frame. A value mismatch between the Transmitter and the Receiver indicates that the data is incorrect and frame should be discarded, with a request for re-transmission. ITU X.25 is used to calculate the frame's checksum [8].

- There is an example of a one "Drone" sending a message from a drone to the server, message consists of three frames. Each frame measures the signal strength from a different position. (Server ID: 4, Drone 1 ID: 1, No. of frames: 3) as shown in Table I.

Header is 35 bytes

(SOF: 8 bytes, EOF: 8 bytes, LEN: 2 bytes, SEQ: 2 bytes, SRC ID: 6 bytes, DEST ID: 6 bytes, MSG ID: 1 byte, Mode: 2 bytes)

Payload of one data information is 36 bytes

(Building Identify 2 bytes, Signal Strength 2 bytes, GPS location of Drone (x, y, z) is 32 bytes)

Checksum is 2 bytes

The size of Message = total size of frames = (frame size * number of frames)

The size of one frame = Header size + Payload size + Checksum size = $35 + 36 + 2 = 73$ Bytes

The size of Message = $73 * 3 = 219$ Bytes.

TABLE I. EXAMPLE FOR TRANSLATE THE DATA

No. of frame	Header								Payload				Check Sum	
	SOF	LEN	SEQ	SRC ID	DEST ID	Mode	MSG ID	EOF	Location(x ,y ,z)			Signal Strength		Building Identify
1	50	3	1	1	4	3	90	500	398	195	104	1757	B1	2304
2	50	3	2	1	4	3	90	500	52	451	91	3843	B1	2349
3	50	3	3	1	4	3	90	500	423	97	250	7331	B1	2973

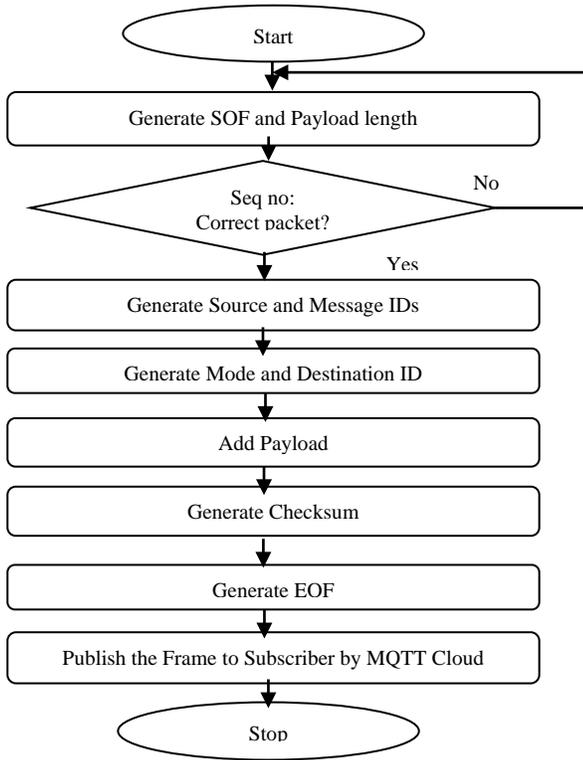


Fig. 6. Flowchart of the Publisher.

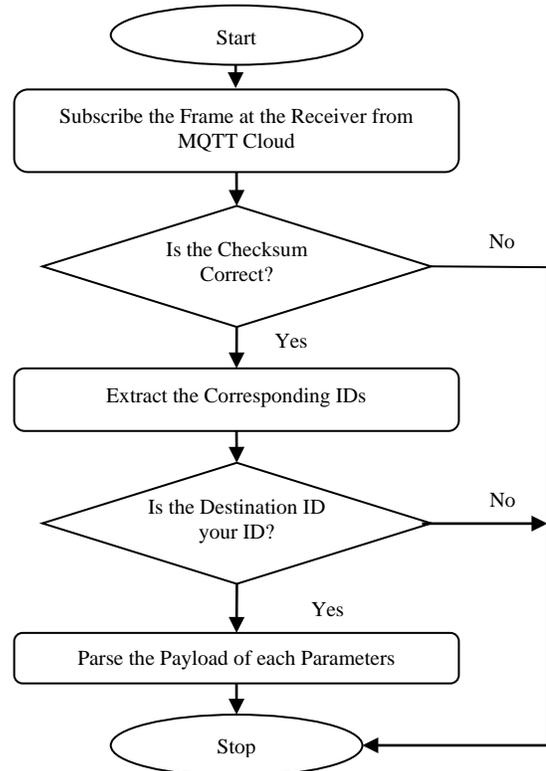


Fig. 7. Flowchart of the Subscriber.

As shown in Fig. 6. “Flowchart of the Publisher” at the transmitter "publisher" side, a fixed SOF indicates the start of new packet. Once the payload data is arrived for first time, the SEQ field is initializes and length of the payload is calculated and fed into the LEN field in the frame. The SRC ID and DEST ID is generated and appended to the frame. MODE is decided and appropriate value is appended in the frame, to this header, the payload data is added and checksum is calculated for the entire frame and a fixed EOF indicates the End of packet and appended at the last position of the frame. Now the frame is ready for transmission "publish" to the MQTT Cloud.

As shown in Fig. 7. “Flowchart of the Subscriber” At the receiver "Subscriber" side the decoding procedure is followed as given in the Fig. 9. Subscribe the Frame at the Receiver from "MQTT Cloud" Once the packet is received, it is cross checked with the transmitted checksum to check the integrity of the transmitted data. If the checksum value matches the transmitted ones, the corresponding payload parameter is parsed by extracting the corresponding IDs. The DEST ID is checked with the host ID to check the transmitted packet is intended for host or not. A mismatch in checksum indicates the packet wrong data.

VI. MOBILE APPLICATION

Network Cell Info Lite - Mobile & WiFi Signal (To measure signal Strength), with testing cell signal strength, it's easier to see how subtle differences in signal strength can affect the performance of the mobile communication system. Lots of people actually consider "Network Cell Info lite app" the best of all apps to measure signal Strength. And for good reason. It covers every cellular network.

VII. BUILDING IDENTIFICATION SYSTEM (BIS) AND (DAS)

A. Proposed Building Identification System

Fig. 8 shows the steps of Building Identification System of detecting the building entrance as shown in [24], which setuped on drone for building identification. The Raspberry Pi model “zero w” will be run by Rasbian OS, Install Building Identification System based on CNN using Python and Install OpenCV for Building Identification. First, send digital map to drones to be able to match the building image with it, then generate "SOF, LEN, SEQ", then check the SEQ no, then generate "SRC ID, MSG ID, Mode, DEST ID, Checksum, and EOF". then creating Cooperative Network between Drones

then detect the drone location then proceed to make the Building Identification by giving the OpenCV sample of Buildings photo as a training sample and matching the building image with digital map, then read the measures of the signal strength by "NC Info Lite app" which setuped on Smart phone, and finally will Publish the frames on "MQTT Cloud".

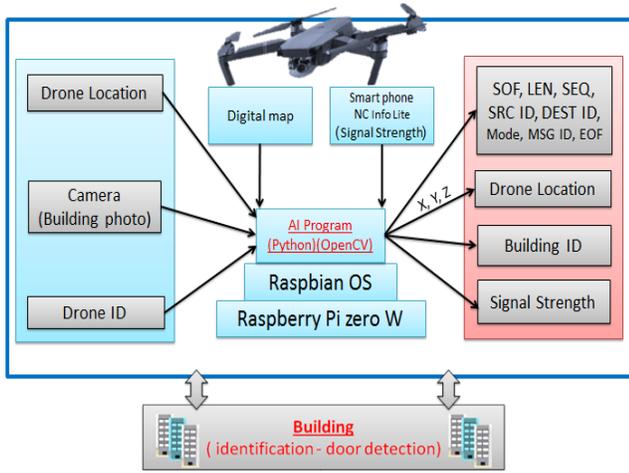


Fig. 8. Implemented Building Identification System.

Algorithm 1: The algorithm for Building Identification System to detect building entrance

Input:
 digital map sent to drone
 Declaration the parameters of header, SOF, LEN, SEQ, SRC ID, MSG ID, Mode, DEST ID, and EOF
 Declaration the parameters of payload (location x y z, building ID, Signal Strength).
 Declaration the parameters of checksum.

Output:
 SOF, LEN, SEQ, SRC ID, MSG ID, Mode, DEST ID, EOF, (location x y z, building ID, Signal Strength) and checksum
 Publish the data on MQTT Cloud

- 1: generate SOF, LEN, SEQ.
- 2: generate SRC ID
- 3: generate, MSG ID
- 4: generate Mode
- 5: generate DEST ID
- 6: generate Checksum
- 7: generate EOF
- 8: read drone location
- 9: building identification
- 10: read Signal Strength
- 11: return SOF, LEN, SEQ, SRC ID, MSG ID, Mode, DEST ID, EOF, (location x y z, building ID, Signal Strength) and checksum

B. Proposed Directing Antenna System (DAS)

Fig. 9 shows Directing Antenna System. Which setuped on "Raspberry Pi" on the base station to control the antennas, the Raspberry Pi will be run by Raspbian OS, Install (DAS) using Python to control antenna.

First, (DAS) will subscribe the Frame information, payload, and checksum value from "MQTT Cloud" that has sent from (BIS) setuped on Drones, then generate checksum value and match the checksum value received from drones with the checksum value generated. If they matched, the corresponding IDs will be extracted. And if not, the process will be repeated again (subscribe the Frame information, payload, and

checksum value from MQTT Cloud). Then generate servo motors degrees (Pan1, Tilt1, Pan2, Tilt2, Pan3, Tilt3) based on the measures of the signal strength to be sent to servo motors.

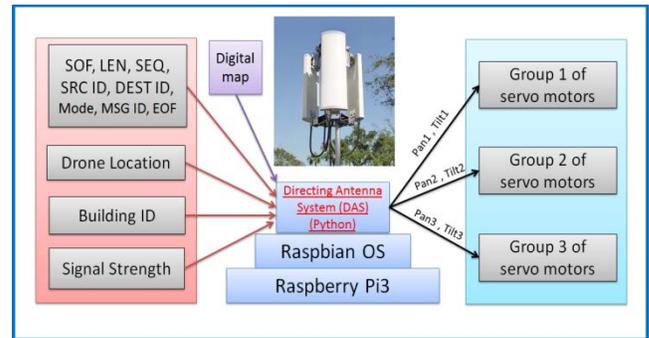


Fig. 9. (DAS) to Control Antenna.

Algorithm 2: The algorithm for (DAS) to controlling antenna

Input:
 The parameter of header, SOF, LEN, SEQ, SRC ID, MSG ID, Mode, DEST ID, and EOF.
 The parameter of payload, (location x y z, building ID, Signal Strength).
 The parameter of checksum.

Output:
 Write servo motors degree Pan1, Tilt1, Pan2, Tilt2, Pan3, Tilt.
 Publish the data on MQTT Cloud.

- 1: subscribe the data from MQTT Cloud
- 2: generate checksum
- 3: **if** Destination checksum is correct **then**
- 4: Extract the corresponding IDs
- 5: **else**
- 6: **Goto** step 1
- 7: **end if**
- 8: **if** the Destination ID == ID **then**
- 9: **Goto** Step 13
- 10: **else**
- 11: **Goto** step 1
- 12: **end if**
- 13: generate servo motors degree Pan1, Tilt1, Pan2, Tilt2, Pan3, Tilt3
- 14: **return** Pan1, Tilt1, Pan2, Tilt2, Pan3, Tilt3.

VIII. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

• Simulation tools

The simulation experiment is to verify the feasibility and performance of the communication protocol and the system. In order to validate our work, OMNeT++ (Objective Modular Network Testbed in C++) framework will be used. OMNET++ is an open source discrete event simulator with modular, component-based C++ simulation library and framework. [29]

The implementation leverages a series of shared libraries over INET, an OMNeT++ model suite for the simulation of wired, wireless and mobile networking protocols. [30].

• Simulation scenario and results

Three Drones are used and there is a cooperation network between them. And through the 2 antennas, the signals will be sent between Drones and the server. Drones measure the signal strength within all the buildings areas and send it to the server

to direct the antennas to the optimal direction. Then measuring the network throughput.

The video in this link shows our simulation scenario

https://drive.google.com/file/d/1KLneJvXlqTj2ZZzqFvQWzEzY_o u06y4Bh/view?usp=sharing

Bandwidth is measured for nine different bandwidth throttles within the speed range of [1 - 16] Mbps downlink and [1- 5] Mbps uplink, and data size 0.11138916 Mbits.

The Data Size is calculated as the following equation where The Data Size (in Mbits) equals the size of the frame (in byte) multiplied by 8 bits to convert it to bits, then multiply size of the frame (in bits) by the number of reading, And then divide by 1024/1024 to convert to Mbits, after that the result will be multiplied by the number of Buildings.

Data Size (Mbits) = (((Size of frame (Byte) * 8 bits * reading) / KB / MB) * for one Building)

$((73 * 8 * 200) / 1024 / 1024) * 1 = 0.11138916$ Mbits

Time_Downlink Throughputs = Data size (Megabits) / Downlink (Mbps) = (s)

Time_uplink Throughputs = Data size (Megabits) / uplink (Mbps) = (s)

Network throughput is one of the key network performance metrics for QoS among end-users. However, demonstrating reliability and validity of network throughput measurement deployed in vendor end-point is a critical task [31].

As shown in Fig. 10, the results of uplink and Downlink speed measurements. Each stripped Brick color bar represents the measured uplink speed that is starting to rise from 1 to 5 Mbps. Similarly, each blue bar represents downlink throughput measurements. The speed is constantly increasing as the Bandwidth increases. that is starting to rise from 1 to 16 Mbps.

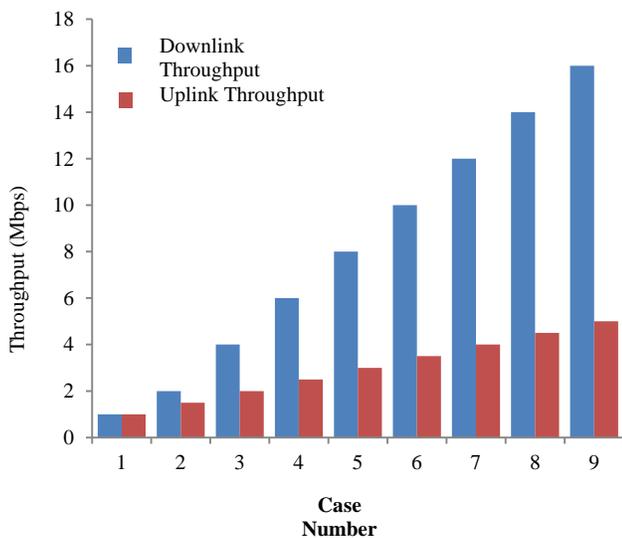


Fig. 10. Measured Downlink /uplink Throughputs for 9 Cases.

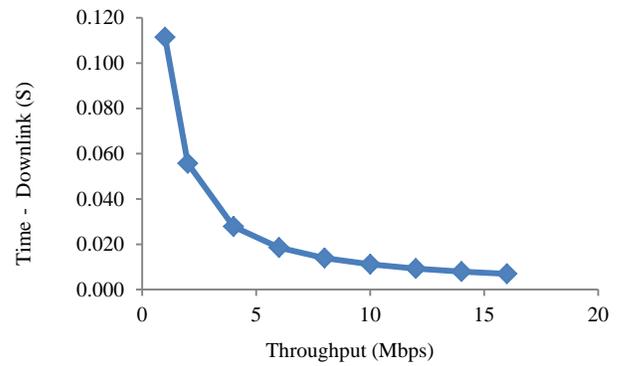


Fig. 11. Time_Downlink Throughputs.

Throughputs_Downlink

As shown in Fig. 11 the Downlink Time within the time range of [0.000 - 0.120] (S) for receiving the amount of data "0.11138916 Mbits" of measuring the signal strength for one building in nine different bandwidth throttles within the speed range of [1 - 16] Mbps. In order to better demonstrate accuracy of the results, will depicting overlay line diagrams for Downlink Time of 9 Cases results.

The figure also illustrates the results of Downlink time and different values for Downlink speed measurements that start to rise from 1 to 16 Mbps. Each blue dot represents the measured downlink speed for the nine cases. The speed is constantly increasing as the Bandwidth increases, On the contrary, the Downlink time decreases as the Bandwidth increases. As shown in the previous example when the data size equal 0.11138916 Mb and speed is 1 Mbps, time becomes 0.111389160 S while the speed increased to 16 Mbps, time decreased to 0.006961823 S.

Throughputs_Uplink

As shown in Fig. 12, the Uplink Time within the time range of [0.000 - 0.120] (S) for sending the amount of data "0.11138916 Mbits" of measuring the signal strength for one building in nine different bandwidth throttles within the speed range of [1.0 – 5.0] Mbps. In order to better demonstrate accuracy of the results, will depicting overlay line diagrams for Uplink Time of 9 Cases results.

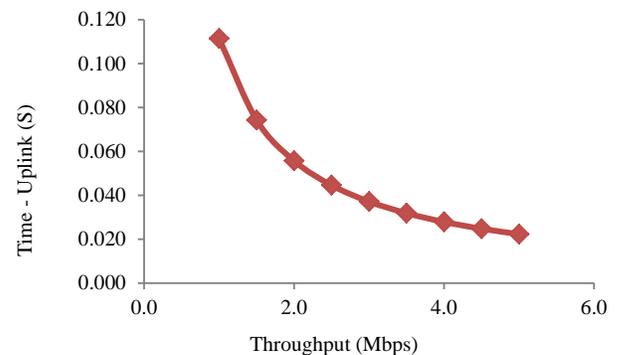


Fig. 12. Time_uplink Throughputs.

The figure also illustrates the results of Uplink time and different values for Uplink speed measurements that start to rise from 1 to 16 Mbps. Each red dot represents the measured uplink speed for the nine cases. The speed is constantly increases as the Bandwidth increases, On the contrary, the Downlink time decreases as the Bandwidth increases. As shown in the previous example when the data size equal 0.11138916 Mb and the speed is 1 Mbps, time becomes 0.111389160 S but when the speed increased to 5 Mbps, time decreased to 0.022277832 S.

As shown in these figures, the results are significantly correlated with each other and can support each other which advocate validity of the performance evaluations.

IX. HARDWARE IMPLEMENTATION

In order to position the antennas, "I²C Servo Driver" will be used (using I²C protocol) connected to "Raspberry pi" (which used as a server) to control all groups of servo motors which are responsible for fully control and direct the antennas. There is three groups of "Servo motors". Every group consists of two "Servo motors" to control the antennas with PAN/TILT mechanism. "Raspberry bi" will give orders to the groups of "Servo Motors" to direct the antennas till it reaches the optimal direction.

Raspberry Pi as a Server. The I2C will be used as a communication protocol between the Server and the servo motors, and Three Groups of servo motors to controlling in three antennas, each group consist of two servo motor (Pan , Tilt).

- The prototype to control antenna

There are three groups of "Servo motors". Every group consists of 2 "Servo motors" to control the antennas with PAN/TILT mechanism. "Raspberry bi" will give orders to the groups of "Servo Motors" to direct the antennas till it reaches the optimal direction.

The video in this link shows our prototype for part of Hardware Implementation

<https://drive.google.com/file/d/1xqOLXLoxJDSu7dtEJyFQNdcIrfaSI2/view?usp=sharing>

X. COMPARISON WITH OTHER WORKS

“To the best of our knowledge, our proposal is the first work to control the antennas direction of mobile communication using drones”, the Table II shows the limitations of the previous communication protocol and cooperative network between drones against the new proposed work.

TABLE II. THE LIMITATIONS OF THE PREVIOUS WORKS AGAINST THE NEW PROPOSED WORK

Reference No.	subject (problem)	Communication Protocol	Throughput measuring	Packet Reception Ratio	Cooperative Network	Simulation	SW system	HW Implementation
The proposed work	5G site survey	√	√	x	√	(Omnet++) √	DAS / BIS	√
[1]	Mobile Communications	x	x	√	x	x	x	x
[3]	Intrusions detection	x	x	x	(Swarm) √	(Omnet++) √	x	x
[4]	disaster, rescue, aerial monitoring	x	x	x	√	x	√	√
[5]	Moving Targets Surveillance	x	x	x	√	√	x	√
[6]	Coverage of Drone Small Cells	x	x	x	√	(Matlab) √	x	√
[8]	communication between Ground Station and drones	√	x	x	(master/slaves) x	x	x	√
[9]	Parcel delivery	x	x	x	(Multi-Drones) x	√	x	x
[10]	Interference in Pipeline Networks	x	x	x	√	(Matlab) √	x	x
[11]	radio base station inspection	x	x	x	√	(UNITY) √	√	√
[12]	positioning measuring in IOD	x	x	x	√	√	x	x
[13]	Assigning location	x	x	x	(Multi-Drones) x	√	x	x
[18]	Disabling a Rogue Drone	x	x	x	√	√	x	x
[19]	Transport Applications	x	x	x	(Multi-Drones) x	√	x	x
[21]	Cooperative Construction	x	x	x	√	√	√	√
[22]	monitor the traffic	x	x	x	(Network) √	√	x	x

TABLE III. PROPOSED MODEL AGAINST TRADITIONAL WORK

		The proposed work	Traditional way
1	Time	Hours	Weeks
2	Accuracy	Accurate	Not accurate
3	Human Power	No persons	Lot of persons
4	Cost	Low cost	High cost

As a result all the disadvantages that mentioned before, As shown in Table III, will be avoided and we will no longer face any high in the cost caused by repeating the process in the traditional way.

XI. CONCLUSION AND FUTURE WORK

In this paper, a comprehensive solution was proposed by designing a standard communication protocol to support the cooperative networks between drones to be used in different fields & applications. The system is for controlling the antenna direction of mobile communication using drones to resolve the problem as the 5G mobile users on higher floors of high buildings suffer from the weak signals. The proposed model included new communication protocol, cooperative network between drones, Building Identification System (BIS) to detect building entrance based on CNN, Directing Antenna System (DAS) to control antenna and hardware controlling antenna. the "Python Programming language" is used for Programming the (BIS and DAS), and " Raspberry Pi" platform to implement the system, and "Raspbian" operating system for " Raspberry Pi" and OpenCV, CNN algorithm for the Building Identification and give the OpenCV sample of Buildings photos as a training sample. MQTT Cloud used to transfer data between drones and server, "OMNeT++" for making the simulation, and the performance of the new Communication Protocol and Network throughput will be measured. At the end we will no longer face the disadvantages like the high cost, the high consumption of time and the lack of accuracy.

Build more cooperative Network to use Applications Between Drones, and Secure the New Proposed Communication Protocol of Drones Cooperative Network.

REFERENCES

[1] Ansari, Nirwan, and Liang Zhang. "Flexible backhaul-aware DBS-aided HetNet with IBFD communications." *ICT Express* 6.1 (2020): 48-56.
[2] Zeadally, S., J. Guerrero, and J. Contreras. "A tutorial survey on vehicle-to-vehicle communications." *Telecommunication Systems* 73.3 (2020): 469-489.
[3] Bekhti, Mustapha, Nadjib Achir, and Khaled Boussetta. "Swarm of Networked Drones for Video Detection of Intrusions." *International Wireless Internet Conference*. Springer, Cham, 2017.
[4] Yanmaz, Evşen, et al. "Drone networks: Communications, coordination, and sensing." *Ad Hoc Networks* 68 (2018): 1-15.
[5] Gu, Jingjing, et al. "Multiple moving targets surveillance based on a cooperative network for multi-UAV." *IEEE Communications Magazine* 56.4 (2018): 82-89.
[6] Dong, Chao, et al. "Optimal deployment density for maximum coverage of drone small cells." *China Communications* 15.5 (2018): 25-40.

[7] Chakrabarty, Anjan, et al. "Vehicle to Vehicle (V2V) communication for Collision avoidance for Multi-copters flying in UTM-TCL4." *AIAA Scitech 2019 Forum*. 2019.
[8] Manasa, K. M., et al. "Design and development of communication protocol for micro air vehicle." *2015 Twelfth International Conference on Wireless and Optical Communications Networks (WOCN)*. IEEE, 2015.
[9] Iranmanesh, Saeid, et al. "Improving Throughput of 5G cellular Networks via 3D placement Optimization of Logistics Drones." *IEEE Transactions on Vehicular Technology* (2021).
[10] Ma, Dazhong, et al. "An Optimal Three-Dimensional Drone Layout Method for Maximum Signal Coverage and Minimum Interference in Complex Pipeline Networks." *IEEE Transactions on Cybernetics* (2021).
[11] Buksz, Dorian, et al. "Intent-driven strategic tactical planning for autonomous site inspection using cooperative drones." *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2020.
[12] Lee, Chao-Yang. "Cooperative Drone Positioning Measuring in Internet-of-Drones." *2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2020.
[13] Bartolini, Novella, Andrea Coletta, and Gaia Maselli. "A Multi-Trip Task Assignment for Early Target Inspection in Squads of Aerial Drones." *IEEE Transactions on Mobile Computing* (2020).
[14] Pu, Cong, and Logan Carpenter. "\$ Psched \$: A Priority-Based Service Scheduling Scheme for the Internet of Drones." *IEEE Systems Journal* (2020).
[15] Guan, Zhangyu, et al. "Distributed Joint Power, Association and Flight Control for Massive-MIMO Self-Organizing Flying Drones." *IEEE/ACM Transactions on Networking* 28.4 (2020): 1491-1505.
[16] Verri, Filipe Alves Neto, et al. "An Analysis on Tradable Permit Models for Last-Mile Delivery Drones." *IEEE Access* 8 (2020): 186279-186290.
[17] Haas, Zygmunt J., and Zhong Zheng. "Engineering a Network in the Sky." *2021 International Conference on Information Networking (ICOIN)*. IEEE, 2021.
[18] Valianti, Panayiota, et al. "Multi-Agent Coordinated Close-in Jamming for Disabling a Rogue Drone." *IEEE Transactions on Mobile Computing* (2021).
[19] Khosravi, Mohammadjavad, et al. "Multi-Purpose Drones for Coverage and Transport Applications." *IEEE Transactions on Wireless Communications* (2021).
[20] Dao, Nhu-Ngoc, et al. "Survey on Aerial Radio Access Networks: Toward a Comprehensive 6G Access Infrastructure." *IEEE Communications Surveys & Tutorials* (2021).
[21] Real, Fran, et al. "Experimental Evaluation of a Team of Multiple Unmanned Aerial Vehicles for Cooperative Construction." *IEEE Access* 9: 6817-6835 (2021).
[22] Huang, Hailong, Andrey V. Savkin, and Chao Huang. "Decentralised Autonomous Navigation of a UAV Network for Road Traffic Monitoring." *IEEE Transactions on Aerospace and Electronic Systems* (2021).
[23] Habibi, Mohammad Asif, et al. "A comprehensive survey of RAN architectures toward 5G mobile communication system." *IEEE Access* 7 (2019): 70371-70421.
[24] Chen, Chun-Wei, et al. "Drone-View Building Identification by Cross-View Visual Learning and Relative Spatial Estimation." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018.
[25] Kang, Do-Hun, et al. "Room temperature control and fire alarm/suppression IoT service using MQTT on AWS." *2017 International Conference on Platform Technology and Service (PlatCon)*. IEEE, 2017.
[26] Mishra, Biswajeetan, and Attila Kertesz. "The Use of MQTT in M2M and IoT Systems: A Survey." *IEEE Access* (2020).
[27] Krishna, P. Gopi, et al. "Implementation of MQTT protocol on low resourced embedded network." *Int. J. Pure Appl. Math. IJPAM* 116 (2017): 161-166.

- [28] Mathur, Shubham, et al. "Human detector and counter using raspberry Pi microcontroller." 2017 Innovations in Power and Advanced Computing Technologies (i-PACT). IEEE, 2017.
- [29] Jin, Chenghao, Bing Chen, and Feng Hu. "A Drone Formation Transformation Approach." International Conference on Machine Learning and Intelligent Communications. Springer, Cham, 2019.
- [30] Barbeau, Michel, Joaquin Garcia-Alfaro, and Evangelos Kranakis. "Geocaching-inspired resilient path planning for drone swarms." IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2019.
- [31] Ablfazli, Saeid, et al. "Throughput measurement in 4G wireless data networks: Performance evaluation and validation." Computer Applications & Industrial Electronics (ISCAIE), 2015 IEEE Symposium on. 2015.

Real Time Face Expression Recognition along with Balanced FER2013 Dataset using CycleGAN

Fatma Mazen Ali Mazen^{1*}, Ahmed Aly Nashat², Rania Ahmed Abdel Azeem Abul Seoud³

Electronics and Communication Engineering Department
Faculty of Engineering, Fayoum University
Fayoum 63514, Egypt

Abstract—Human face expression recognition is an active research area that has massive applications in medical field, crime investigation, marketing, online learning, automobile safety and video games. The first part of this research defines a deep neural network model-based framework for recognizing the seven main types of facial expression, which are found in all cultures. The proposed methodology involves four stages: (a) pre-processing the FER2013 dataset through relabeling to avoid misleading results and getting rid of non-face and non-frontal faces; (b) design of an efficient stable Cycle Generative Adversarial Network (CycleGAN), which provides unsupervised expression-to-expression translation. The CycleGAN has been designed and trained with a new cycle consistency loss. (c) Generating new images to overcome the class imbalance and finally (d) building the DNN architecture for recognizing the face sign expression, using the pretrained VGG-Face model with vggface weights. The second part encompasses the design of a GPU-accelerated face expression recognition system for real time video sequences using NVIDIA's Compute Unified Device Architecture (CUDA). OpenCV library has been compiled from scratch with CUDA and NVIDIA CUDA Deep Neural Network library, cuDNN. For face detection stage Haar Cascaded and deep learning were used and tested using both CPU and GPU as backend. Results show that the designed model run time to recognize a face sign is 0.44 seconds. Besides, the average test accuracy has been increased from 64% for the original FER2013 dataset to 91.76% for the modified balanced version using the same transfer learning model.

Keywords—Facial expressions detection and recognition; multi-task cascaded convolutional networks; transfer learning; residual neural network; CycleGAN; FER2013; GPU and CUDA; HAAR

I. INTRODUCTION

Face sign expressions are one of the most active research areas in computer vision since they are a form of nonverbal communication for people in the deaf community. It is, also, used for understanding human behavior, mental disorder detection, cognition of human emotions, and lie detection. They convey non-verbal cues, which play an important role in interpersonal relations. Automatic recognition of facial expressions can be a vital component of natural human-machine interfaces; it may also be used in different areas such as artificial intelligence, computer vision, psychology, physiology, behavioral science and in clinical practice. Some robots can also benefit from the ability to recognize expressions [1]. Automated analysis of facial expressions for behavioral science or medicine is another possible application

domain. It can be used to detect the state of the learner in E-learning, help doctors to understand and analyze the behavior of children with Autism Spectrum Disorder (ASD) who are known to have difficulty in producing and perceiving emotional facial expression [2]. The system of facial expression recognition is divided into three steps: optimal preprocessing, feature extraction or selection, and classification, particularly under conditions of input data variability to attain successful recognition performance.

Dataset preprocessing is an elementary step in machine learning systems especially when the dataset contains wrongly labeled images due to source crowding and suffers from class imbalance, which leads to biased learning toward majority classes. In this paper, an innovative strategy has been adopted to overcome this shortage. CycleGAN is a promising data augmentation scheme recently adopted to solve the class imbalance problem by generating new samples of the minority class. An efficient stable CycleGAN has been designed and trained to perform style transfer from a reference domain, which is the neutral class, to a target domain, which could be any of the other face expression classes. After creating a clean balanced dataset, the transfer learning approach has been adopted for model design. Designing a model from scratch is a time-consuming process and will never outperform efficiently, since the pretrained models have been trained on datasets containing millions of images which is a huge number compared to our dataset.

Many machine learning algorithms have been proposed for successful recognition performance, but they have not reached the optimal performance due to lack of the optimal set of features required for classification. Feature extraction is the process of using domain knowledge of the data to create features that make machine learning algorithms work. Coming up with features is difficult, time-consuming, requires expert knowledge. Deep learning aims at learning feature hierarchies where features from higher levels of the hierarchy are formed by lower-level features, so it eliminates the need to feature extraction and as a result achieving better performance in less time than traditional machine learning algorithms. Convolutional Neural Networks, CNN, are currently one of the most prominent algorithms for deep learning with image data. Whereas for traditional machine learning relevant features must be extracted manually. In computer vision machines, storing the knowledge learned by solving one problem and applying it to another similar problem is challenging. The cross-domain transfer learning models have received enormous

* Corresponding Author

attention in face expression classification and generally artificial intelligence applications, where the training and test data are drawn from different feature space and different distribution. The knowledge learned from the learning dataset can help improve prediction accuracy in the testing domain, especially when the testing data are target scant. Also, knowledge from a labeled domain can help generate labels for an unlabeled domain, which may avoid a costly human labeling process.

In recent years, massive computer visions algorithms and techniques have been developed and employed for designing real time face expression recognition system due to its significance not only in human-to-human interaction but also in Human Machine Interaction (HMI). HMI is an extruded field in computer science that aims at making the computer intelligent such that it interacts with human the way human and human interact. This task can be decomposed into two stages, face detection for face localization and face expression classification. Recently, two main approaches have been adopted for face detection, the traditional machine learning approach, and the deep learning approach.

In this paper, a new framework for real time face expression recognition in video sequences has been designed and tested. It comprises two phases. For face detection phase, two versions of OpenCV library have been employed. The first version is standard OpenCV library that was used for machine learning approach. With the rapid development of GPU, deep learning became more powerful in classification tasks. The second version is compiled from scratch with CUDA and cuDNN support to achieve optimal hardware resources employment through interaction between our Graphical Processing Unit (GPU) and deep learning module of OpenCV which undoubtedly has a great effect in improving the speed of video by increasing number of frames that can be processed per second (FPS) making our system more convenient for real time applications.

The rest of the paper is organized as follows: Section 2 discusses the work related to the face expression classification problem. Section 3 illustrates the dataset and the proposed methodology for the facial expression recognition. Experimental results and discussion are presented in Section 4. Finally, Section 5 concludes the paper and outlines directions for future work.

II. RELATED WORK

In the last decade, face expression recognition has become a hot research area. Several papers have been published using classical machine learning approaches like Support Vector Machine (SVM) and Random Forest, whereas others use modern schemes like deep learning, convolutional neural networks (CNNs), transfer learning and ensemble models.

Ying Zilu and G. Zhang [3] adopted a framework for facial expression recognition based on the combination of non-negative matrix factorization (NMF) and support vector machine (SVM). They achieved a recognition rate of 66.19% with their algorithm applied to Japanese female facial expression database (JAFEE database). In [4], face expression recognition system was designed by Y. Luo, C. M. Wu and Y.

Zhang based on a hybrid scheme of principal component analysis (PCA) and local binary pattern (LBP) for feature extraction and support vector machine (SVM) for facial expression classification. They achieved a recognition rate of 93.75% using a small dataset of 350 face sign images. C. Shan, S. Gong and P. W. McOwan [5] proposed a novel scheme based on local binary pattern (LBP) for facial expression recognition. The Cohn-Kanade Database was used to demonstrate the efficiency of their proposed approach. Their model attains a recognition rate of 79.1%. In [6], Ahmed A. Nashat proposed a new approach for extracting features for facial expression classification. Firstly, the discrete wavelet packet best tree (DWPBT) decomposition is used to remove the spatial and spectral redundancies for each block of the desired face regions. Then, the radial difference LGP (RD-LGP) in radial directions of a circular grid is used as a descriptor for facial expression recognition. The average recognition rate was 83.4%.

Y. Wang, Y. Li, Y. Song and X. Rong [7] have proposed an effective hybrid approach of random forest and convolutional neural networks. Their approach achieved an accuracy of 98.9% on JAFFE dataset, 99.9% on CK+ dataset, 84.3% on FER2013 dataset and 92.3% on RAF-DB dataset. E. Barsoum, C. Zhang, C. C. Ferrer and Z. Zhang [8] worked on using a deep CNN on noisy labels acquired via crowdsourcing for ground truth images. They used 10 taggers to relabel each image in the dataset, and used various cost functions for their DCNN, achieving decent accuracy. They reached accuracy of 84.986% on FER+ dataset. C. Pramerdorfer and M. Kampel [9] formed an ensemble of recent deep convolutional neural networks. They investigated the approaches adopted by six current state-of-the-art papers and ensembles their networks to reach 75.2% test accuracy on FER2013 outperforming previous works without needing auxiliary training data.

In [10], X. Wang, J. Huang, J. Zhu, M. Yang and F. Yang have utilized transfer learning using Keras VGG-Face library and each of ResNet50, SeNet50 and VGG16 as pre-trained models. They managed to achieve test accuracy of 75.8% which surpassed all existing publications at this time. A. Chen, H. Xing, and F. Wang [11] proposed an efficient and secure facial expression recognition method based on the edge cloud framework combined with the improved CycleGAN to solve class imbalance of FER2013 dataset by generation 4000 new sample for disgust class and 800 new samples for surprised class achieving higher recognition rate not only on the data augmented classes but also on the other classes. They achieved average recognition rate of 78.61% on the enhanced dataset. Z. Zhang, P. Luo, C. C. Loy, and X. Tang [12] also achieved 75.1% test accuracy by adding auxiliary data and additional features: a vector of HoG features was computed from face patches and processed by the first FC layer of the CNN. Facial landmark registration has been also employed.

In this work, a neoteric relabeled balanced FER2013 dataset has been introduced. A balanced CycleGAN has been designed and trained using new cycle consistency loss to preserve luminance conditions of reference class to overcome class imbalance problem through generating new images for minority class (Disgust). State of the art results have been achieved through dataset cleaning and pre-processing. A new

deep neural network architecture for recognizing the face sign expression, using the pretrained VGG-Face model with vggface weights and the modified balanced version of the FER2013 dataset is designed and implemented. To enhance our work, A GPU-accelerated system for face expression recognition in real time video stream is introduced achieving 312.01% faster inference for feature-based approach and inference speed improvement by up to 169.74% for deep learning-based approach.

III. METHODOLOGY

A. The Dataset

The proposed methodology was trained and tested using the open-source FER2013 dataset [13-14], which is created for an ongoing project by Pierre-Luc Carrier and Aaron Courville from university of Montreal, then shared publicly for a Kaggle competition, shortly before ICML 2013. The dataset consists of 35.887 labelled 48x48 grayscale human facial expressions. These are afraid (11.42%), angry (11.13%), disgusted (1.21%), happy (20.1%), neutral (13.84%), sad (13.46%) and surprised (8.84%). To train and test the performance of the proposed classification model, we selected 80% of the set for training and the rest for testing. Fig. 1 shows some sample images from the FER2013 dataset.



Fig. 1. Sample Images from the FER2013 Dataset.

B. Data Set Preprocessing Phase

Crowdsourcing has become a widely used approach to gather ground truth labels to form a dataset. However, these labels can be very noisy which causes model misleading and as a result low recognition rate [8]. FER2013 suffers from severe crowdsourcing as it contains non-face images, text images, sleepy faces, profile images which are indistinguishable by humans, and large number of wrongly labelled images. According to [15], the overall accuracy of the facial expression classification using this dataset does not exceed 65%. In this research, three steps have been taken to resolve this problem. First, we deleted non-face images, text images and profile images. Second, we re-labelled wrongly labelled images based on class distribution using convolutional neural networks (CNN). Finally, the class imbalance problem has been solved by designing a CycleGAN used for generating new face expressions for the minority classes. CycleGAN phase will be discussed in detail in the next section. Fig. 2 displays some bad images, which have been deleted.

As shown in Fig. 3, a few examples of wrongly labelled misleading images due to crowd sourcing are displayed with the wrong expression in the red color and the right expression in the blue color.



Fig. 2. Samples of Bad Images that have been deleted from the Dataset.



Fig. 3. Samples of Mislabeled Images.

Besides crowd resourcing, the dataset suffers from class imbalance problem which results in model biasing. As shown in Fig. 4, class1 contains only 436 images which are too low compared to other classes which contain thousands of images. The class imbalance problem has been solved by designing a CycleGAN used for generating new samples for the minority class. CycleGAN phase will be discussed in detail in the next section.

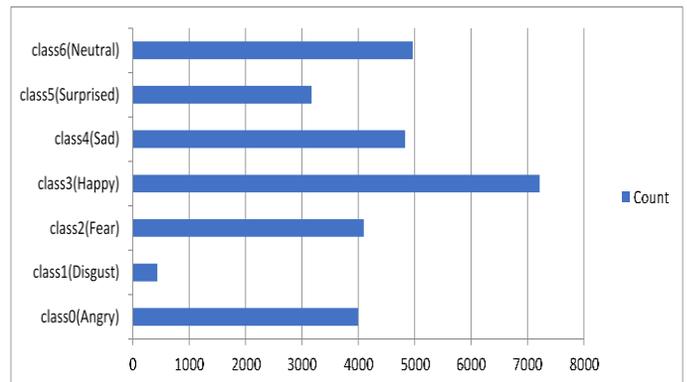


Fig. 4. A Severe Class Imbalance with the Highest Number of Images at Class 3 and the Lowest at Class 1.

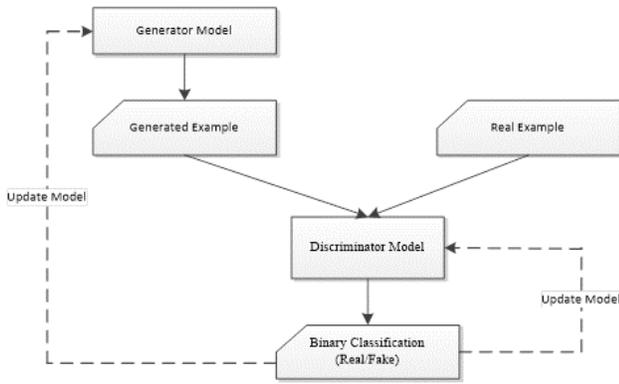


Fig. 5. Flowchart of the Generative Adversarial Network Model Architecture.

C. CycleGAN Phase

It is a challenging task to classify images with multiple class labels using only a small number of labelled examples, especially when the label (class) distribution is imbalanced. Generative Adversarial Networks (GANs) can be used to generate images from an adversarial training. The generator attempts to produce a realistic image to fool the discriminator, which tries to distinguish whether its input image is from the training set or the generated set. The flow chart of the GAN architecture is shown in Fig. 5. J. Y. Zhu, T. Park, P. Isola, and A. Efros [16] proposed the CycleGAN model, which can do image-to-image transition between two unpaired image domains. This model has been employed to build our framework.

The architecture of the discriminator, the generator and the composite model are like those used in [17] except for employing some training tricks like using dropout layers to avoid GAN failure modes, labels flipping (True label was used for generated images while false label was used for Real images) and using Batch Normalization layers to produce sharper generated images. Unlike other GAN models for image translation, the CycleGAN does not require a dataset of paired images.

The discriminator and the composite model architecture of the CycleGAN are shown in Fig. 6 and Fig. 7, respectively.

1) *CycleGAN loss functions*: Each generator model is optimized via the combination of four outputs with four loss functions and are defined by (1), (2), (3), and (4) as follows:

- Adversarial loss (L_2 or mean squared error). It is calculated as the L_2 distance between the model output and the target values of 1.0 for real and 0.0 for fake.

$$\text{Adversarial loss} = L_2 \quad (1)$$

- Identity loss (L_1 or mean absolute error). It is calculated as the L_1 distance between the input and output image for each sequence of translations.

$$\text{Identity loss} = L_1 \quad (2)$$

- Forward cycle loss

$$\text{Forward cycle loss} = \frac{1}{3}(L_1 + L_2 + \text{SSIM loss}) \quad (3)$$

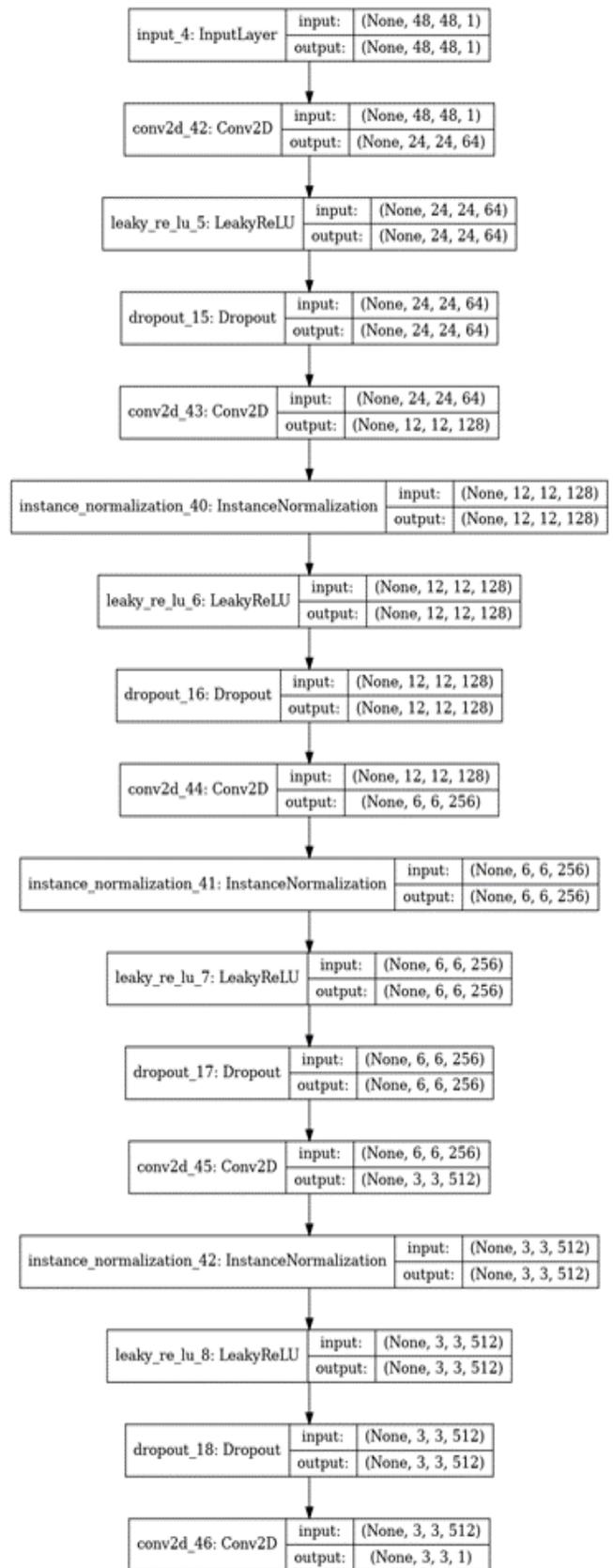


Fig. 6. The Discriminator Architecture.

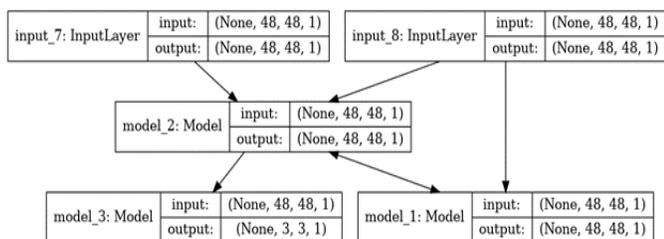


Fig. 7. The Composite Model of the CycleGAN.

- Backward cycle loss

$$\text{Backward cycle loss} = \frac{1}{3}(L_1 + L_2 + \text{SSIM loss}) \quad (4)$$

In this paper, the cycle consistency loss, for a better-quality image, is defined as the average of the L_1 distance and the L_2 distance. The structural similarity index measure (SSIM) loss is defined by (5) as:

$$\text{SSIM loss} = 1 - \text{SSIM index} \quad (5)$$

Where SSIM Index is computed by considering the luminance, contrast, and structural similarity. Since the image pixels are correlated and not independent, L_2 loss is unsuitable choice to improve the quality of generated images. SSIM loss term is added to the cycle consistency loss to compensate for this disadvantage by reducing structural distortions between the input and output images. It is worthy mentioned that SSIM was designed for gray scale images.

2) *Training details:* In our design, after five trials, the forward and the backward cycle loss are given more weight than the adversarial loss whereas the identity loss weight has been set to two. In other words, λ_{cyc} is set to five and λ_{id} equals to two.

For all the experiments, the Adam optimizer has been used with a batch size of one. All networks were trained from scratch with a learning rate of 0.0002. The learning rate was kept constant for the first 100 epochs and linearly decays to zero over the next 100 epochs. Fig. 8 shows two sets of sample images of translation from class 6, (Neutral), to class 1, (Disgust).

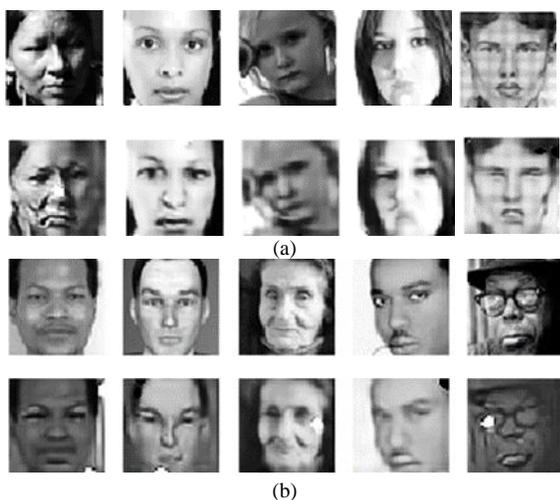


Fig. 8. Two Sets, a and b, of Samples of Neutral to Disgust Translation.

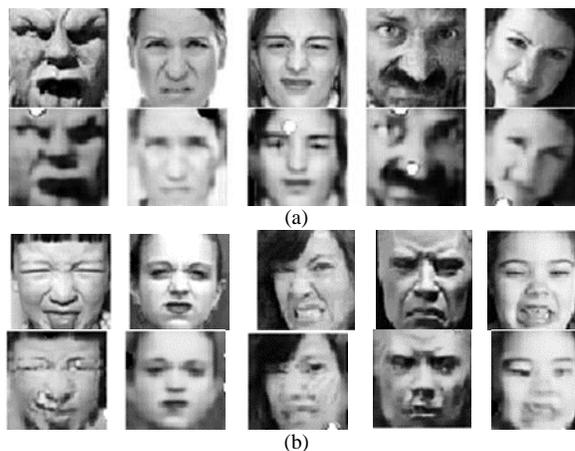


Fig. 9. Two Sets, a and b, of Samples of Disgust to Neutral Translation.

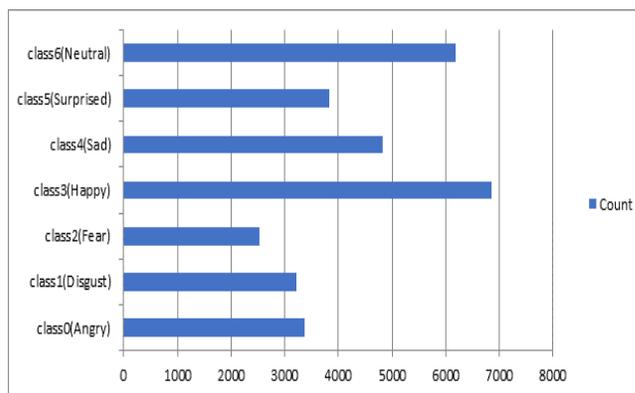


Fig. 10. Data Distribution of the New Balanced FER2013 Dataset.

Fig. 9 displays sample images of class 1 (Disgust) to class 6 (Neutral) translation which are close to real ones.

Fig. 10 shows data distribution of the new balanced FER2013 dataset, after cleaning it from bad images and generating new images using CycleGAN for the minority class.

D. Transfer Learning Phase

In this paper, transfer learning scheme has been adopted as it has a significant role in achieving state of the art accuracy. Keras VGG-Face pre-trained model [18] has been employed with vggface weights. As this network accepts RGB images with size 224x224, the 48x48 gray scale images in FER2013 during training time have been resized and recolored using image data generator.

1) *Training details:* In our design, the top layers of the VGG-Face model, which include last fully connected layers and softmax layer have been removed and a flatten layer was added to extract the bottleneck feature vector. To reduce overfitting due to large network size, the last three fully connected (dense) layers have been replaced by only one fully connected layer with 128 neurons followed by a dropout layer with dropout rate equals to 0.3 and then a softmax layer with seven outputs has been added to match the seven output facial expressions.

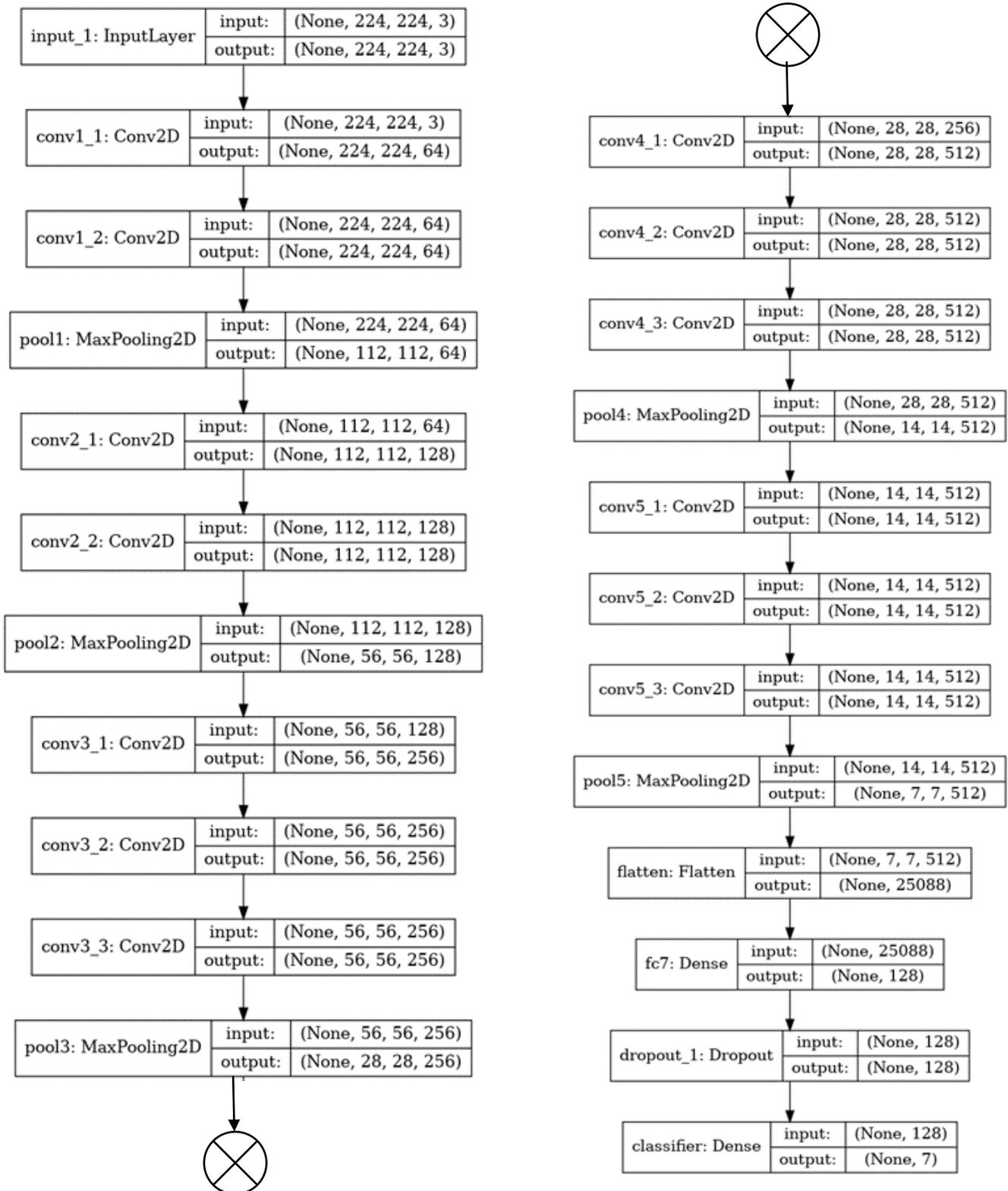


Fig. 11. The Transfer Learning Model.

During training the network, the model has been compiled using Adam optimizer with learning rate equals to 0.001 and Beta_1 equals to 0.5. The input to the network has been divided into minibatches of 128 images and the total number of epochs equals to twenty epochs. All experiments have been conducted using Keras with TensorFlow backend on Alien ware laptop with Nvidia GeForce GTX 1070 GPU. To prevent model overfitting and performance deterioration, early stopping callback function is used to halt training the model at the right time. Model Check Point callback has been used to save the model only when better results are obtained according to predefined performance measures. Reduce on Plateau callback has been used to reduce learning rate when the validation accuracy decays. The detailed description of our transfer learning model is shown in Fig. 11.

The upcoming sections discuss the procedure of designing a real time facial expression recognition system to detect face and recognize the seven standard face expressions in real time video sequences.

E. Compiling OpenCV from Source

In this stage OpenCV with CUDA and cuDNN support has been compiled from source to achieve the best optimum utilization of hardware resources (Alien ware laptop with NVidia GeForce GTX 1070 GPU). OpenCV has been compiled with eight cores which should be adjusted according to the number of processor cores you use. The block diagram of this process is shown in Fig. 12.

The CMake output for NVIDIA CUDA and cuDNN indicates that CUDA version, is 10.1 and cuDNN version is 7.6.5. Another important CMake output that should be verified before proceeding to compilation step are the proper python paths to Interpreter and Libraries.

F. Face Detection Phase

Face detection is a challenging computer vision problem which aims to identify and localize faces in images. It is a primary step for any facial expression recognition (FER) system. It can be accomplished using the classical machine learning (feature based) Haar cascaded classifier or using recent deep learning-based approaches through using the OpenCV library.

1) *Feature-based approach (Haar cascaded classifier):* OpenCV provides the cascade classifier class that can be used for face detection. The constructor can take a filename as an argument that specifies the XML file for a pre-trained model. The model can be used for face detection on an image or video by calling the detectMultiScale function whose output is a list of bounding boxes for all faces detected. ScaleFactor and minNeighbors are two parameters that should be carefully tuned for a given dataset as the ScaleFactor controls how the input image is scaled before detection.

2) *Deep learning-based approach:* After building OpenCV from source with CUDA and cuDNN support, now its Deep Neural Network (DNN) module can be used. It contains a CNN-based face detector which enhances the face detection performance compared to machine learning-based

models like Haar Cascade classifier. It utilizes the Single Shot Detector (SSD) framework with ResNet as the base network. The confidence level is extracted after looping over the detections to suppress weak detections if they do not meet the minimum confidence level. In our experiment, the minimum confidence level was set to 0.5.

G. Face Expression Recognition Phase

In this phase, the system classifies each frame of the video into one of the seven universal expressions - Anger, Disgust, Fear, Happiness, Sadness, Surprise and Neutral as labelled in the FER2013 dataset. The flowchart of the proposed deep learning framework and machine learning framework are shown in Fig. 13 and Fig. 14, respectively.

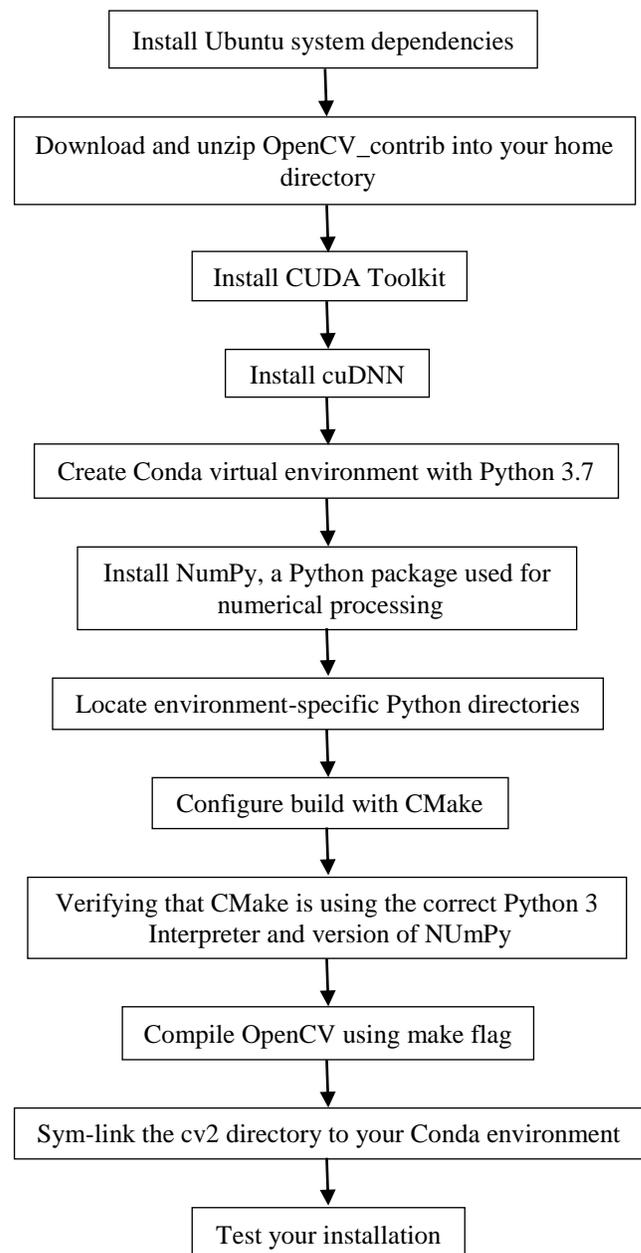


Fig. 12. The Block Diagram of OpenCV Compilation Process.

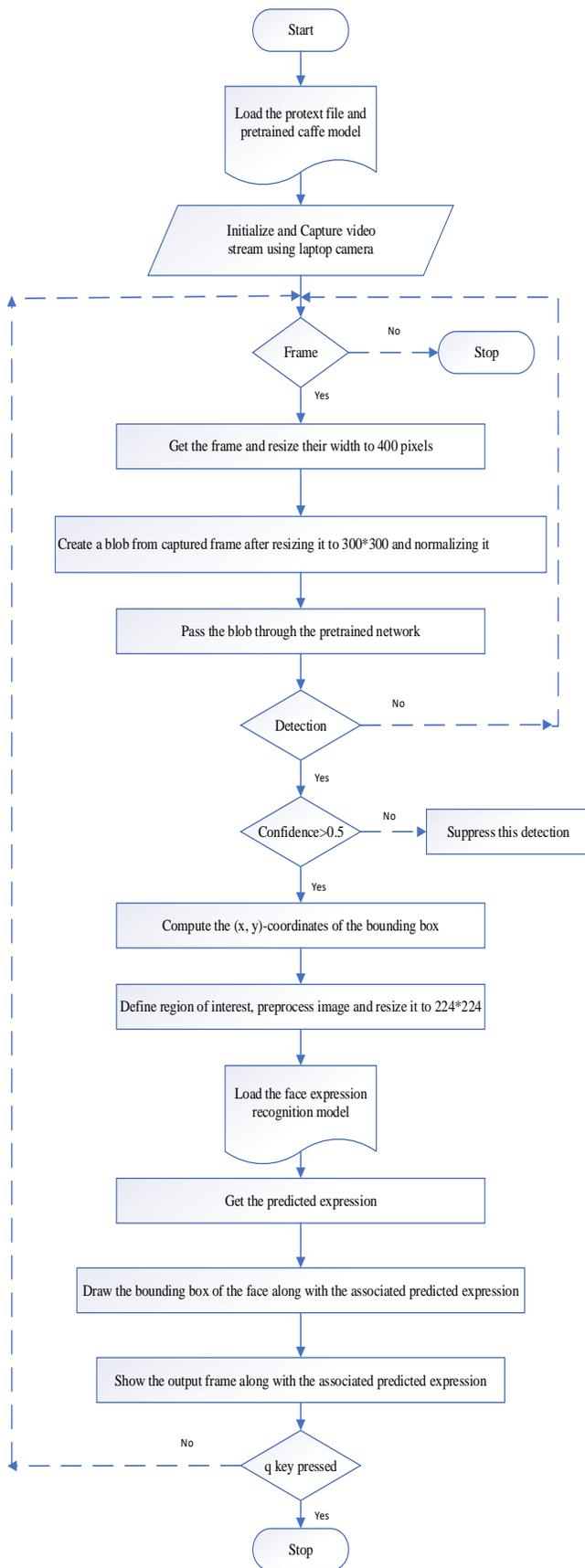


Fig. 13. Flowchart of Deep Learning-based Approach.

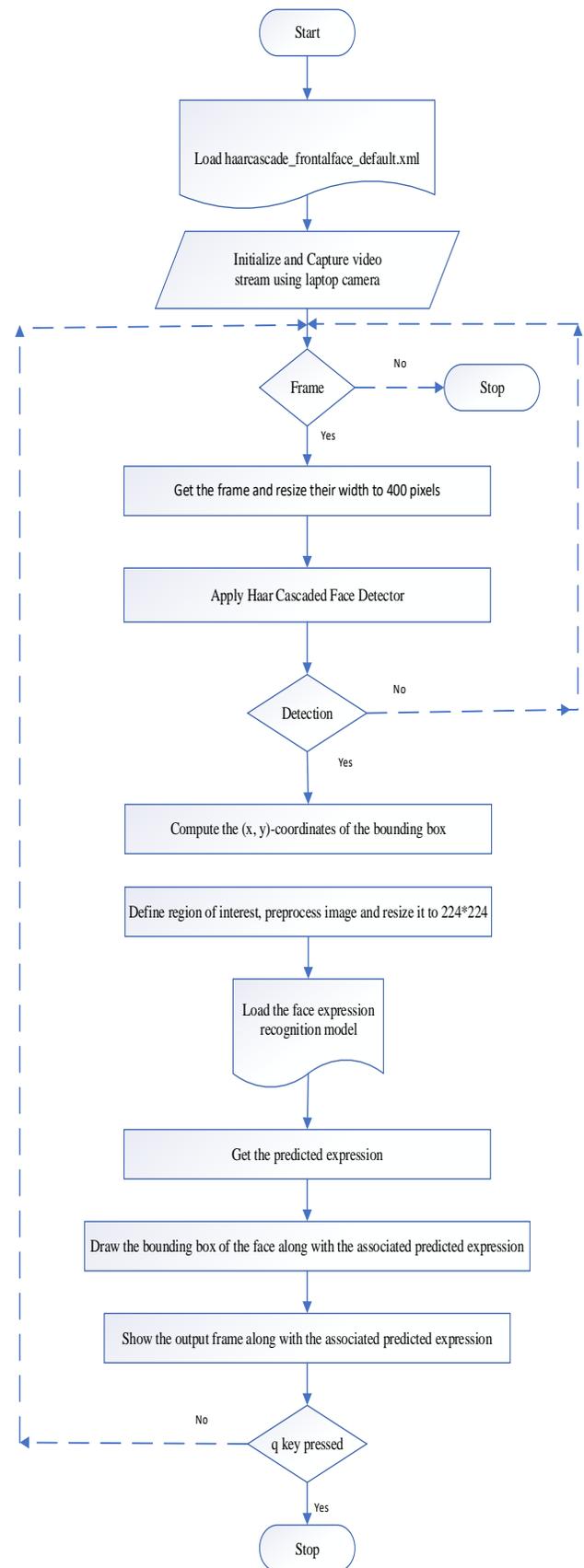


Fig. 14. Flowchart of Machine Learning (Haar Cascaded) based Approach.

IV. RESULTS AND DISCUSSION

The designed FER transfer learning model has been tested on both the original FER2013 dataset and the modified balanced version of FER2013 dataset. Fig. 15 and Fig. 16 show the robustness of the model and the effectiveness of the modified balanced dataset. It is seen that the test loss approaches zero and the overall test accuracy reaches 91.76%, which is a state-of-the-art result, for our model using the modified balanced FER2013 dataset. However, the test loss increases, and the average overall test accuracy drops to 64% for our model using the traditional FER2013 dataset despite using dropout layer and a dense layer with small number of neurons to prevent model overfitting.

Results show that our approach starting from dataset cleaning, re-labelling, and solving class imbalance problem by generating new samples of the minority class using CycleGAN is a promising scheme.

Fig. 17 shows the normalized confusion matrix for our model using the original FER2013 dataset and the modified balanced version of FER2013 dataset. It is seen that the class recognition rate has improved remarkably using the modified balanced FER2013 dataset. The recognition rate is higher by 35% for the angry class, 71% for the disgust class, 50% for the fear class, 14% for the happy class, 34% for the sad class, 21% for the surprised class, and 29% for the neutral class.

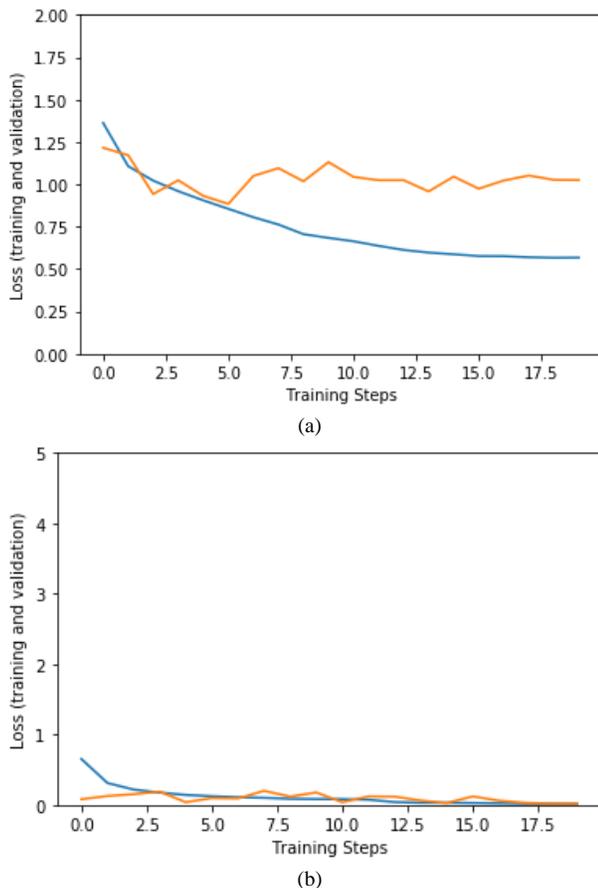


Fig. 15. Loss Curve. (a) Original FER2013 Dataset and (b) Modified Balanced FER2013 Dataset.

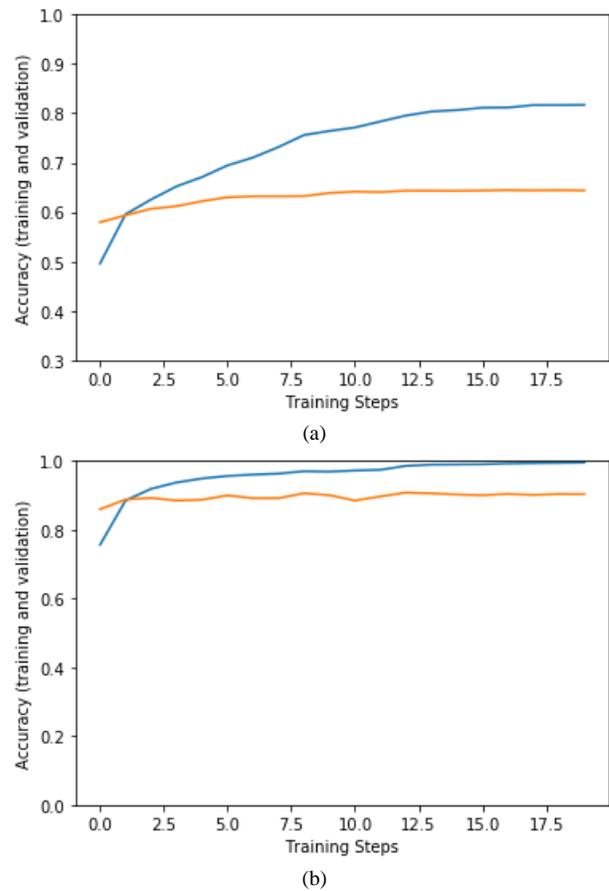


Fig. 16. Learning Curve. (a) Original FER2013 Dataset and (b) Modified Balanced FER2013 Dataset.

TABLE I COMPARISON OF THE STATE OF THE ART ACCURICES ON FER-2013 DATASET

Methodology	Test Accuracy
Zhang et al [12]	75.1%
Christopher Pramerdorfer et al [9]	75.2%
Amil Khanzada et al [10]	75.8%
An Chen et al [11]	78.61%
Barsoum et al [8]	84.986%
Our approach	91.76%

Table I summarizes the top test accuracies of recent approaches applied to FER-2013 dataset, where our approach achieves state of the art results.

To test our model in real time, Multi-task Cascaded Convolutional Networks (MTCNN) [19] has been used to detect the face prior to recognizing the type of the expression. Fig. 18 shows results.

In this stage, the compiled OpenCV DNN module is ready to be used. Four experiments have been conducted using Keras with TensorFlow backend on Alien ware laptop with NVidia GeForce GTX 1070 GPU. The first and third experiments used CPU as backend. On the other hand, the second and fourth experiments utilized GPU as backend.

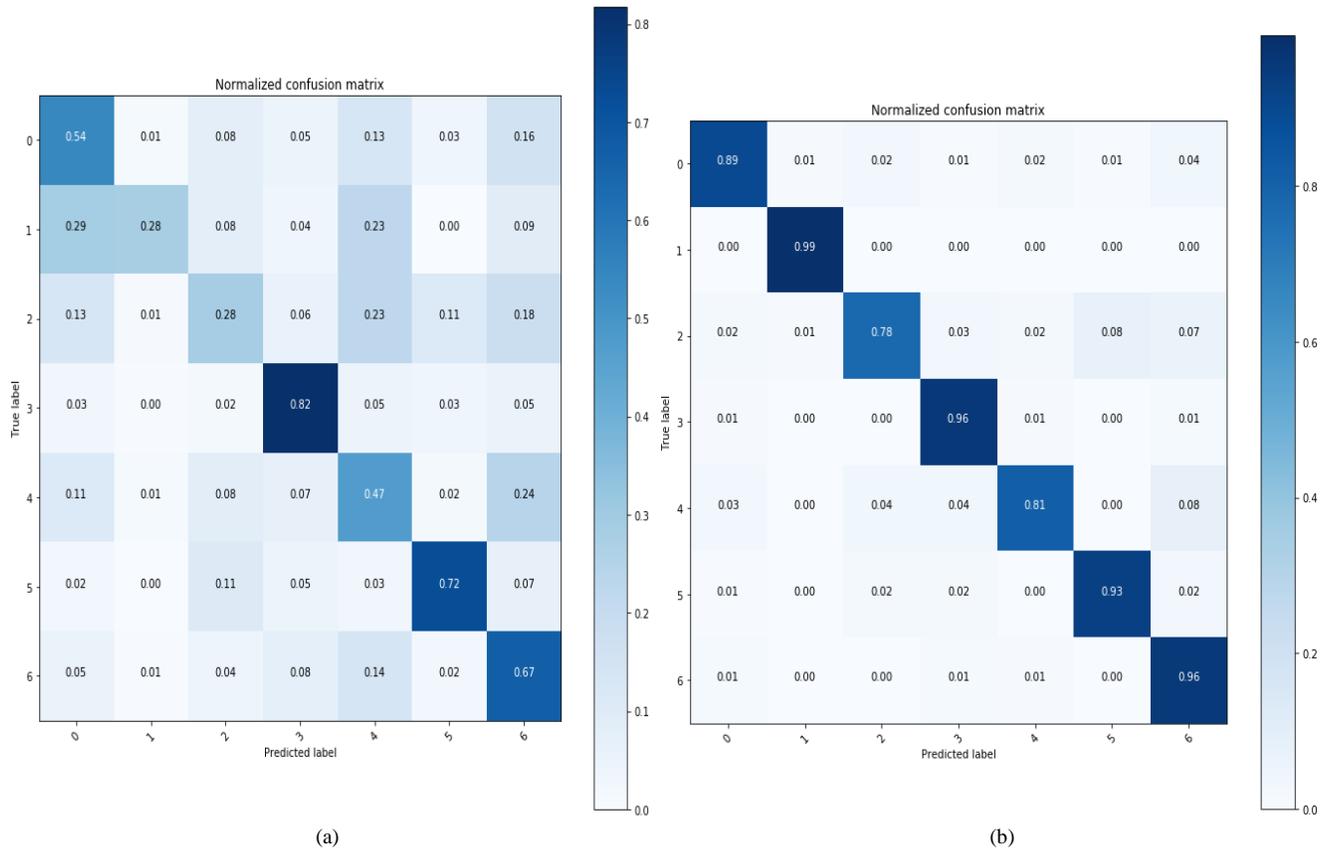


Fig. 17. The Confusion Matrix. (a) Original FER2013 Dataset and (b) Modified Balanced FER2013 Dataset.

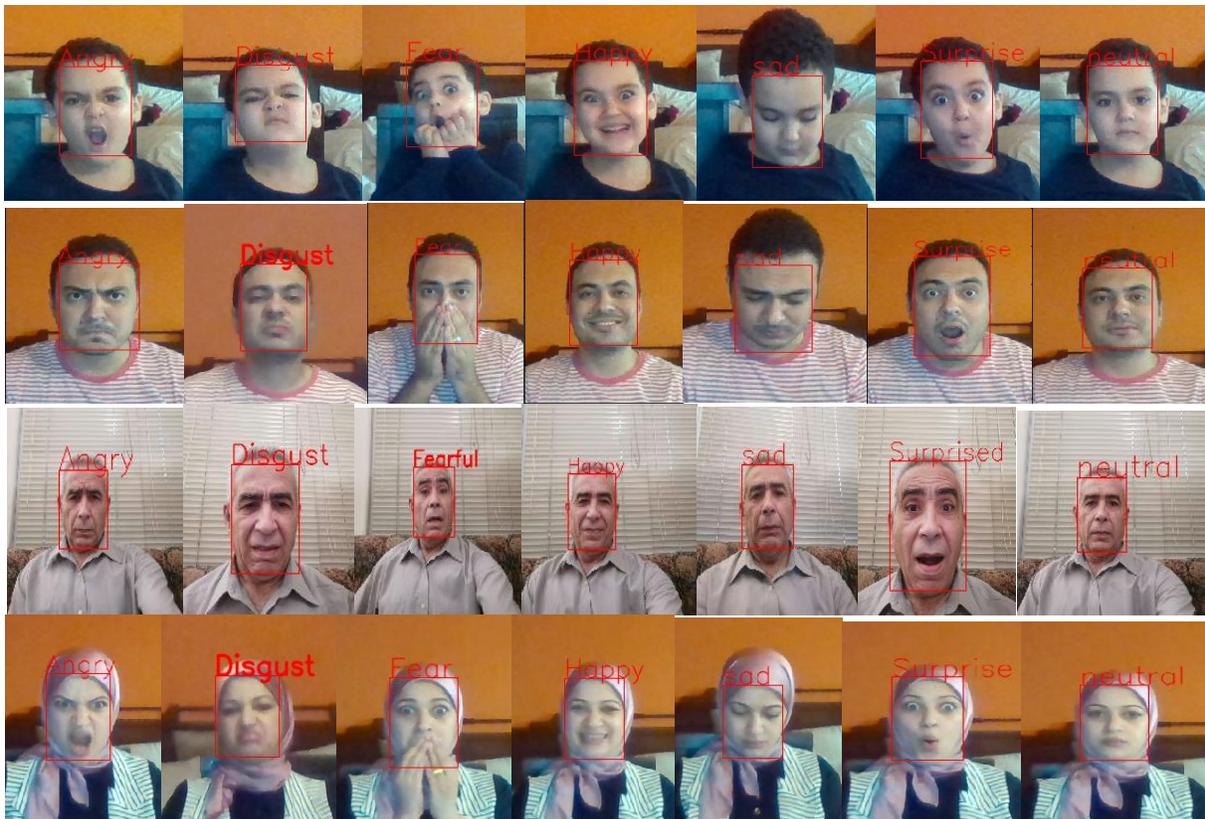


Fig. 18. Real Time Face Expression Recognition for different Subjects Showing the Seven Principal Expressions.

TABLE II FRAME PER SECOND OF THE PROPOSED METHODOLOGIES USING CPU/GPU BACKEND

The Proposed Methodology	Face Detection	FPS
1	Haar cascade classifier	7.41
2	Haar cascade classifier	23.12
3	deep learning-based approach	30.30
4	deep learning-based approach	51.43

As shown in Table II, machine learning-based methodology which used Haar cascaded classifier in face detection phase has been tested on CPU and GPU and the frame per second (FPS) has been compared. Using NVIDIA GPU has improved inference speed by up to 312.01%.

Deep learning-based methodology which used deep learning in face detection phase has also been tested on CPU and GPU and FPS has been compared. FPS throughput rate is improved by over 169.74% with OpenCV's DNN module and an NVIDIA GPU.

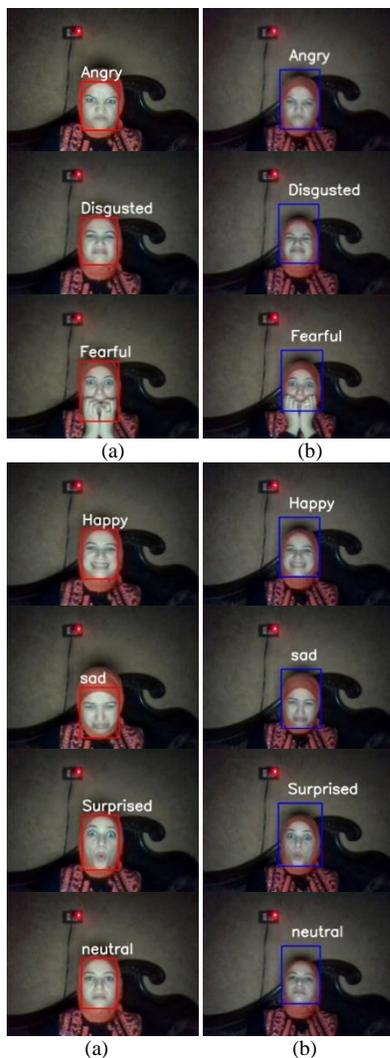


Fig. 19. The Output of Real Time FER System. (a) Deep Learning-based Approach and (b) Feature-based Approach.

Fig. 19 displays the output of FER system for real time video sequence using deep learning and Haar cascaded classifier in face detection phase, respectively.

V. CONCLUSION AND FUTURE WORK

A deep neural network model based upon VGG-Face pre-trained framework for face expression recognition has been implemented. An efficient stable Cycle Generative Adversarial Network, (CycleGAN), is designed and used to produce a balance relabeled FER2013 dataset. The designed model produced an optimal performance and a state-of-the-art overall recognition rate of 91.76% in only 0.44 second using Alien ware laptop with Nvidia GeForce GTX 1070 GPU. A GPU-based framework has been designed for face expression recognition in real time video stream. OpenCV library has been compiled from source to achieve the best exploitation of hardware resources (GPU). The framework involves two main stages, face detection and face expression recognition. The face detection phase has been realized through machine learning-based approach (Haar cascaded classifier) and deep learning-based approach. Face expression recognition has been accomplished using transfer learning approach. Both Methodologies have been tested with CPU and GPU as backend. The performance is evaluated through FPS of the whole process. Deep learning has been assessed to be faster and more accurate.

Future directions will focus on applying the model for real time face expression recognition in video sequences or online learning platform. The simplicity, the high recognition rate and the speed of the facial expression classification model make it appropriate for implementing a productive and profitable computer vision machine.

REFERENCES

- [1] S. Yousefi, M. P. Nguyen, N. Kehtarnavaz, and Y. Cao, "Facial expression recognition based on diffeomorphic matching," Proc. - Int. Conf. Image Process., 2010, ICIP pp 4549–4552. <https://doi.org/10.1109/ICIP.2010.5650670>.
- [2] C. Tsangouri, W. Li, Z. Zhu, F. Abtahi, and T. Ro, "An Interactive Facial-Expression Training Platform for Individuals with Autism Spectrum Disorder," in: 2016 IEEE MIT Undergraduate Research Technology Conference (URTC), 2016, pp. 1–4, <https://doi.org/10.1109/URTC.2016.8284067>.
- [3] Z. Ying and G. Zhang, "Facial expression recognition based on NMF and SVM," Proc. - 2009 Int. Forum Inf. Technol. Appl. IFITA 2009, 3, 2009, pp 612–615. <https://doi.org/10.1109/IFITA.2009.279>.
- [4] Y. Luo, C. M. Wu, and Y. Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM," Optik (Stuttg), 124, 2013, pp 2767–2770. <https://doi.org/10.1016/j.ijleo.2012.08.040>.
- [5] C. Shan, S. Gong, and P. W. McOwan, "Robust facial expression recognition using local binary patterns," Proc. - Int. Conf. Image Process., ICIP 2, 2005, pp 370–373. <https://doi.org/10.1109/ICIP.2005.1530069>.
- [6] A. A. Nashat, "Facial expression recognition using best tree RD-LGP encoded features and HMM," Int. J. Wavelets, Multiresolution Inf. Process, 16, 2018, <https://doi.org/10.1142/S0219691318500479>.
- [7] Y. Wang, Y. Li, Y. Song, and X. Rong, "Facial expression recognition based on random forest and convolutional neural network," Inf. 10, 2029, pp 1–16. <https://doi.org/10.3390/info10120375>.
- [8] E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, "Training deep networks for facial expression recognition with crowd-sourced label distribution," ICMI 2016 - Proc. 18th ACM Int. Conf. Multimodal Interact, 2016, pp 279–283. <https://doi.org/10.1145/2993148.2993165>.

- [9] C. Pramerdorfer and M. Kampel, "Facial Expression Recognition using Convolutional Neural Networks," State of the Art., 2016.
- [10] X. Wang, J. Huang, J. Zhu, M. Yang, and F. Yang, "Facial expression recognition with deep learning," ACM Int. Conf. Proceeding Ser, 2018. <https://doi.org/10.1145/3240876.3240908>.
- [11] A. Chen, H. Xing, and F. Wang, "A Facial Expression Recognition Method Using Deep Convolutional Neural Networks Based on Edge Computing," IEEE Access, 8, 2020, pp 49741–49751. <https://doi.org/10.1109/ACCESS.2020.2980060>.
- [12] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, "Learning social relation traits from face images," Proc. IEEE Int. Conf. Comput. Vis., 2015 International Conference on Computer Vision, ICCV 2015, pp 3631–3639. <https://doi.org/10.1109/ICCV.2015.414>.
- [13] fer2013 | Kaggle [WWW Document], n.d. URL <https://www.kaggle.com/deadskull7/fer2013> (accessed 11.3.20).
- [14] Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," Neural Networks, 64, 2015, pp 59–63. <https://doi.org/10.1016/j.neunet.2014.09.005>.
- [15] G. C. Porusniuc, F. Leon, R. Timofte, and C. Miron, "Convolutional neural networks architectures for facial expression recognition," 2019 7th E-Health Bioeng. Conf. EHB 2019. <https://doi.org/10.1109/EHB47216.2019.8969930>.
- [16] J. Y. Zhu, T. Park, P. Isola, and A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," Proc. IEEE Int. Conf. Comput. Vis., October 2017, pp 2242–2251. <https://doi.org/10.1109/ICCV.2017.244>.
- [17] J. Brownlee, "Generative Adversarial Networks with Python," Mach. Learn. Mastery 2019, pp 1–654.
- [18] A. M. Bukar and H. Ugail, "Convnet features for age estimation," Proc. Int. Conf. Comput. Graph. Vis. Comput. Vis. Image Process, Big Data Anal. Data Min. Comput. Intell. - Part Multi Conf. Comput. Sci. Info 2017, pp 94–102.
- [19] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," IEEE Signal Process. Lett., 23, 2016, pp 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>.

A Deep Learning Approach Combining CNN and Bi-LSTM with SVM Classifier for Arabic Sentiment Analysis

Omar Alharbi

Computer Department
Jazan University, Jazan, Saudi Arabia

Abstract—Deep learning models have recently been proven to be successful in various natural language processing tasks, including sentiment analysis. Conventionally, a deep learning model's architecture includes a feature extraction layer followed by a fully connected layer used to train the model parameters and classification task. In this paper, we employ a deep learning model with modified architecture that combines Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM) for feature extraction, with Support Vector Machine (SVM) for Arabic sentiment classification. In particular, we use a linear SVM classifier that utilizes the embedded vectors obtained from CNN and Bi-LSTM for polarity classification of Arabic reviews. The proposed method was tested on three publicly available datasets. The results show that the method achieved superior performance than the two baseline algorithms of CNN and SVM in all datasets.

Keywords—Sentiment analysis; Arabic sentiment analysis; deep learning approach; convolutional neural network CNN; bidirectional long short-term memory Bi-LSTM; support vector machine; SVM

I. INTRODUCTION

Sentiment analysis (SA) is one of the most active research areas in natural language processing (NLP). The SA task aims to equip a machine with the ability to categorize people opinions into positive or negative based on the sentiment expressed in the texts. SA is technically a challenging task, as human language is complex and diverse. Nevertheless, research on SA has achieved considerable progress, especially for the English language, while Arabic SA is developing slowly despite increasing Arabic language usage on the Internet [1]. That can be attributed to the Arabic language's complex morphological nature and diverse dialects [2]. Additionally, technical issues such as scarce linguistic resources and limited linguistic tools of Arabic increased the level of complexity [3]. The Arabic language is one of the most common languages that is used by more than 400 million individuals to daily communication. Arabic is formally written in a form called Modern Standard Arabic (MSA), which is understood all over the Arabic world. Despite this fact, internet users usually tend to use dialectal words alongside MSA to write reviews, tweets, or comments. Dialectal words result from behaviors such as replacing characters and changing the pronunciation or the style of writing of nouns, verbs, and

pronouns of MSA [4]. Thus, there are no standard rules that can handle the morphological or syntactic aspects of reviews written in MSA with the presence of dialects.

One of the prominent approaches employed in SA is machine learning (ML). In this context, SA is formalized as a classification problem which is addressed by using algorithms like Naive Bayes (NB), Support Vector Machine (SVM) and Maximum Entropy (MaxEnt). In fact, ML has proved its efficiency and competence; therefore, in the past two decades, it mostly dominates the SA task. However, a fundamental problem that would decay the performance of ML is the text representation through which the features are created to be fed to the learning process. For text representation, the most used model in literature is the bag-of-words (BOW) with unigram, bigram, or part of speech (POS). By using BOW, words are often weighted by their presence (binary scheme) or frequency like in term frequency (TF) or term frequency-inverse document frequency (TF-IDF). This model is relatively straightforward and important; however, it can be problematic and may degrade ML-based sentiment classification performance when it comes to a large number of features or the need for semantic information [5-8].

Over the past decade, the deep learning approach and word vector representations have attained an increasing interest from NLP researchers as they can successfully handle the limitations of traditional ML methods at NLP tasks, including SA (Kim 2014). Deep learning is an extended approach of ML and a subset of the neural network. It has a structure composed of multiple hidden layers, which enable it to automatically discover semantic representations of texts from data without feature engineering [9]. This approach has improved the state-of-the-art in many SA tasks, including sentiment classification, opinion extraction, and fine-grained sentiment analysis [10]. Along with deep learning, word vector representations, also known as word embeddings, has emerged as a powerful features representation model for the classification task. Word embeddings is a modern approach to learn real-valued low-dimensional vector space representations for a text [11]. It aims to encode continuous semantic similarities between words based on their distributional properties in a large corpus [12]. Word embeddings are typically extracted by using neural networks as the underlying predictive model [13]. One of the most used methods for word embedding is Word2Vec that was proposed by [11]. Word2Vec produces useful word representations learned by a 2-layer neural network based on

¹ <https://en.unesco.org/commemorations/worldarabiclanguageaday>

the model that has been proposed in [14]. Word2vec is one of the key methods, which has led to the state-of-the-art results achieved by the deep learning approach on SA. Word2vec can be implemented using two main learning models; they are continuous bag-of-words (CBOW) and Skip-gram.

Researchers have proposed many variants of deep learning models to address the SA of English, which have shown promising efficacy. For instance, in [15], the authors use Recursive Neural Tensor Network (RNTN) along with word embeddings to address fine-grained sentiment classification for phrase level. In another work [16], the authors apply an ensemble model that includes Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) networks on top of pre-trained word vectors for tweets sentiment classification. The authors in [17] develop an adaptive Recursive Neural Network (RNN) that uses more than one composition functions and adaptively selects them depending on the linguistic tags and the combined vectors for tweets sentiment classification. Unlike English, there have not been many studies that address the Arabic language in SA based on a deep learning approach and word embeddings, except a few such as [18-22], which will be discussed with other studies in the following section. Accordingly, there is still room for exploring different deep learning algorithms and investigating their performance in addressing Arabic SA, especially with the challenges that the Arabic language is still imposing.

In this work, we propose a new method that combines deep learning models with SVM for improving the SA of Arabic reviews. Conventionally, a deep learning model's architecture includes features extraction layers (e.g. CNN or RNN) followed by a fully connected layer used for training the model parameters and classification task. However, in this work, we decided to modify the structure and explore the impact of this structure on sentiment classification performance. The structure is altered by replacing the fully connected network with a linear SVM classifier that trains the embedded vectors generated by the combined deep learning models. This is justified by SVM's remarkable improvement in recent research in SA, including Arabic SA, such as in [20, 23, 24]. The idea of combining the neural network and SVM was initially proposed in [25] to improve multilayer perceptron classifiers. In this respect, many researchers presented various models for different classification tasks including image classification [26, 27], visual recognition [28], intrusion detection [29], object categorization [30], speech recognition [31], and sentiment analysis [32]. Although the proposed method is inspired by these studies, it is different as we use a linear SVM as a replacement for the fully connected layer instead of replacing only the top layer's activation function (i.e. Softmax or Logistic). In particular, this research presents a combination of CNN and Bidirectional-LSTM with linear SVM to improve the sentiment classification performance on Arabic reviews at the document level. To the best of our knowledge, this is the first attempt to use such a combination for Arabic SA. In this work, we used three publically available datasets for evaluation: LABR [33], OMCCA [34], and the dataset presented in [35].

The remaining sections of this paper are organized as follows: Section 2 discusses the related work in Arabic SA. In Section 3, we present the proposed method. Section 4 describes

the experimental details used to evaluate the performance of the proposed method. In Section 5, we report and discuss the obtained results. Section 6 presents the conclusion and future work.

II. RELATED WORK

Researchers have presented several methods and models to tackle Arabic SA based on three major approaches, namely, machine learning (ML) [36, 37], semantic orientation [38, 39], and hybrid approach [40]. In this research, we focus on the studies that employ the ML approach, specifically deep learning methods, which have been increasingly utilized in the past decade. These methods have presented remarkable improvements in the SA field, especially for the English language, such as in [41-45]. On the other hand, only a few studies have attempted to utilize the deep learning approach for Arabic SA. For instance, in [46], the authors use an ensemble model of CNN and LSTM proposed by [16] to predict the sentiment of Arabic tweets for the sentence level. The model is trained on top pre-trained word vectors developed in [47]. They evaluated the model on ASTD dataset [48] and achieved an F1-score of around 64% and an accuracy of around 65%. Another work applies CNN and LSTM to sentiment analysis of Arabic tweets described in [49]. They designed a system to identify the sentiment's class and intensity as a score between 0 and 1. The authors used word and document embedding vectors to represent the tweets. They translated the tweets into English to benefit from the available preprocessing tools. As they highlighted, the step of the translation led to degrading the overall performance. Although the system includes some preprocessing steps, it excludes some other important normalizing processes such as removing diacritics, punctuations and repeating characters.

In [50], the authors investigate LSTM and Bi-LSTM models' performance on Saudi dialectal tweets. After some basic normalization processes, they use the CBOW model to represent words with vectors. As reported, Bi-LSTM outperformed LSTM and SVM with an accuracy of 94%. The study of [19] explores different deep learning models, including a combination of LSTM and CNN, to predict tweets' sentiment. They also investigate the use of CBOW and Skip-gram models of available pre-trained word vectors introduced in [51]. The experiments show that the highest results were obtained by using a two-layer LSTM model. Barhoumi et al. [52] utilized Bi-LSTM and CNN to evaluate different types of Arabic-specific embeddings. They suggested using available Arabic NLP tools to address the effect of the agglutinate and morphological rich specificity of Arabic on word embedding quality. However, these NLP tools such as lemmatization, light stemming, and stemming have been mostly developed to MSA texts.

The work in [53] reports efforts to detect the sentiment of tweets of time series data in the domains of stock exchange and sport. They use a deep neural network model with eight fully connected hidden layers in which each layer has 100 neurons. They compared the model against KNN, NB, and decision tree classifiers. Their model showed superiority with an F1-score of around 91% and accuracy of around 88%. Nevertheless, the authors did not mention the kind of feature representation used

in this work. The work in [54] explores different architectures for Arabic sentiment classification using Deep Belief Networks (DBF), Deep Auto Encoders (DAE), and Recursive Auto Encoder (RAE). The experimental results show that RAE obtained the best performance due to its capability to handle the sentence's context and order of parsing. Yet, the authors focus on sentence-level sentiment classification, and they used the traditional BOW to represent texts. In [55], the authors implemented LSTM to handle Arabic sentiment classification. The model was implemented along with character/word embedding features. In their experimental results, LSTM has shown the ability to improve the sentiment classification and achieved an accuracy of around 82%. The same author in another work [56] found that SVM work better than RNN on the same dataset, but the execution time of RNN was faster. However, both studies focus on aspect-based sentiment analysis of Arabic reviews.

This current work introduces a combination of CNN with Bi-LSTM and a linear SVM classifier for Arabic sentiment classification. Unlike the studies mentioned above, the proposed model does not follow the conventional CNN and Bi-LSTM. Rather, the fully connected layers would be replaced with a linear SVM algorithm to perform the classification task.

III. PROPOSED METHOD

The proposed method for sentiment classification of Arabic reviews starts with preparing data by applying some data preprocessing techniques. Then, the word vector representation is built using pre-trained word embeddings trained on other large Arabic text corpora. After that, we combine CNN and Bi-LSTM for generating the final features representation to be passed to a linear SVM classifier.

A. Data Preprocessing

The preprocessing stage consists of removing noise from data and normalization. The process of removing noise from data includes removing repeated letters, diacritics, punctuations, numerals, English words, and elongation. After that, a normalization process was applied to particular letters, for example, the letters (ا, ا, ا) were converted to (ا), the letters (ي, ي) were converted to (ي), the letter (ة) was converted to (ة), and finally, the letter (ة) was converted to (ة). It is worth mentioning that we implemented these particular preprocessing steps to be compatible with those used in the pre-trained word embeddings model we intend to use. More details about the pre-trained word embeddings model in the next section.

B. Word Vector Representations

To generate representations for the reviews, we adopted the word vectors approach. Word vectors (also known as word embeddings) are vectors of real numbers representing the distribution of words or phrases in a given text [14]. This approach's key benefit is that high-quality word embeddings can be learned efficiently, mainly if a much larger corpus of text were used for training [57]. However, the adopted dataset's size is relatively small in our case, which might lead to low-quality word embeddings. Therefore, we decided to use a pre-trained word embeddings model that trained on other large Arabic texts. One of the most common Arabic pre-trained word embeddings is Aravec [47]. This pre-trained model was

trained on texts from multiple sources such as Twitter, Web pages, and Wikipedia with tokens of more than 3,300,000,000. This model was implemented by the well-known word2vec technique described in [11] based on CBOW and Skip-gram architectures.

C. Convolutional Neural Network Layer

We first build a CNN layer to extract the features of reviews. The inputs to the CNN are matrices of word embeddings X that reflect the word vector representations of the reviews. Each review is mapped to a matrix of size $s \times d$, where s is the number of words in the review and d is the dimension of the embedding space. All reviews must be zero-padded to have the same matrix dimension $X \in \mathbb{R}^{s' \times d}$. Then, in order to compute a new feature map, inputs are convolved with a filter matrix $w \in \mathbb{R}^{h \times d}$, where h is the size of the convolution. The convolution is computed as in Equation 1:

$$c_i = f(\sum_{j,k} w_{j,k} (X_{[i:i+h-1]_{j,k}}) + b) \quad (1)$$

Where $b \in \mathbb{R}$ is a bias term and $f(x)$ is a nonlinear activation function. The output is the result of element-wise product between an input matrix X and a filter matrix w , which then be summed as a single value in a feature map $c = [c_1, c_2, \dots, c_n]$, where $c \in \mathbb{R}^{n-h+1}$. To obtain rich features representation the model can use multiple filters of different length that can work in parallel. Then, the convolved results are passed to the pooling layer to reduce the representation dimensionality by identifying the most important features. We employed the max-pooling technique which returns the maximum k values from the feature map $\hat{c} = \max(c)$. This technique has the ability to force the network to capture the most useful local features produced by the convolutional layer [58]. A typical architecture of a CNN layer is illustrated in Fig. 1. The output of this layer is fed into the next layer that is Bi-LSTM, based on the assumption that is more informative representation will be obtained, more details in the next section.

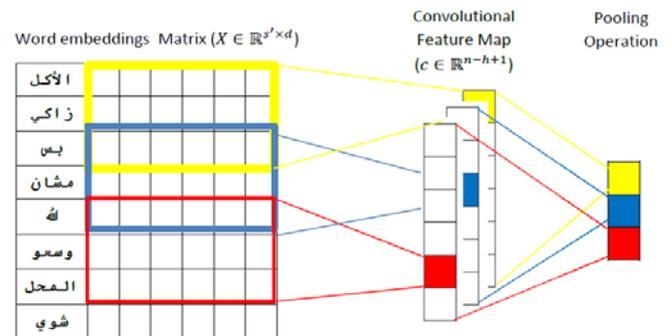


Fig. 1. A typical CNN Layer Architecture, the Colored Boxes Represent the Size of a Filter Spatially.

D. Bidirectional-Long Short-Term Memory (Bi-LSTM) Layer

As mentioned above the input into this layer is the final feature map obtained from CNN. First, let us present a little background about LSTM to understand the Bi-LSTM. LSTM is a type of recurrent neural network RNN, which was developed by [59]. The main objective of LSTM is to avoid vanishing gradient problem that encounters RNN, where the

gradient that is propagated back through the network either decays or grows exponentially over time [60]. In other words, RNN is not an effective model for understanding the context behind an input with long-term dependencies. On the other hand, LSTM with its complicated dynamics that consists of several so-called memory blocks, can effectively resolve this issue and make a prediction based on time series data. Each memory block is composed of an input gate i , a forget gate f , an output gate o , and a cell state c . These elements are used to compute the hidden state h by the following composite function in Equation 2:

$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \\
 c_t &= f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \\
 h_t &= o_t \tanh(c_t)
 \end{aligned} \tag{2}$$

Where x_t is the input vector to the memory cell at time t , σ is the sigmoid function, W is the weight matrices, and b is the bias.

However, one downside of LSTM inherited from RNN is that it does not consider the whole context because the sentence is only read in a forward direction [16]. In this case, the context information provided by the future words will be dismissed, resulting in low classification performance. An extension of LSTM called bidirectional LSTM (Bi-LSTM) [61] was proposed to avoid this problem. In this work, we adopted this method to obtain more informative features based on the features given by CNN. Bi-LSTM simultaneously trains two separate LSTMs of opposite directions (one for forward and one for backward) on the input sequence and then merge the outputs. More formally, as explained above, the input vectors x_t are fed one at a time t into the LSTM, making use of all the available input information up to the current time frame t_c to predict y_{t_c} . However, applying the bidirectional network's notion will enable the network to use the input information coming up later than t_c by delaying the output by a certain number M time frames up to x_{t+M} to predict y_{t_c} . The structure of Bi-LSTM is illustrated in Fig. 2 that shows the basic structure of folded Bi-LSTM which computes the forward hidden sequence \vec{h} , the backward hidden sequence \overleftarrow{h} , and the output sequence y by iterating the backward and forward layers. This layer also if followed by a max-pooling layer as outlined in the previous section.

After this layer, we have a layer that concatenates the list of outputs from the previous layers into a single vector.

Conventionally, this vector is passed to a fully connected layer that applies weights over the generated features $\alpha(w * x + b)$ to predict the class of a given input, where $w \in \mathbb{R}^{m \times m}$ is the weights matrix, and α is the activation function. However, in this work, we will replace these fully connected layers with a linear SVM to perform the prediction process, more details in the next section.

E. Support Vector Machine (SVM) Classifier

As mentioned earlier, we present a model that combines CNN and Bi-LSTM with SVM, where CNN and Bi-LSTM used for feature vectors generation and SVM for sentiment classification. SVM was employed as it has proven to be an effective algorithm in many NLP tasks, including SA. The key idea behind combining these heterogeneous methods is to use each method's advantages to improve the Arabic language's sentiment classification. In general, SVM is a machine learning algorithm for binary classification problems introduced by [63]. The SVM classifier aims to find an optimal hyperplane that classifies given data points into one of two classes by maximizing the margin. More formally, a linear SVM will receive a concatenated vector of features x_i associated with its labels y_i to be processed by a hyperplane $f(w, x) = w^T \cdot x_i + b$ to compute the coefficients of the hyperplane as feature weights, where w is the normal vector to the hyperplane and also known as the weight vector and b is the bias.

The architecture of the proposed model is illustrated in Fig. 3. Obviously, our model's architecture is essentially the same as the basic CNN and Bi-LSTM with a difference in using the SVM algorithm for classification. Further clarity, CNN or Bi-LSTM typically use a fully connected network with different activation functions (i.e. logistic) for learning and classification, as explained earlier. However, in this work, we replace the fully connected layer with a linear SVM which use a margin-based loss function for classification.

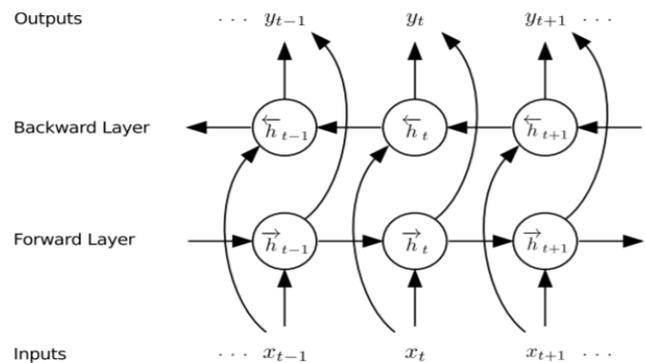


Fig. 2. Architecture of Bi-LSTM (Source: [62]).

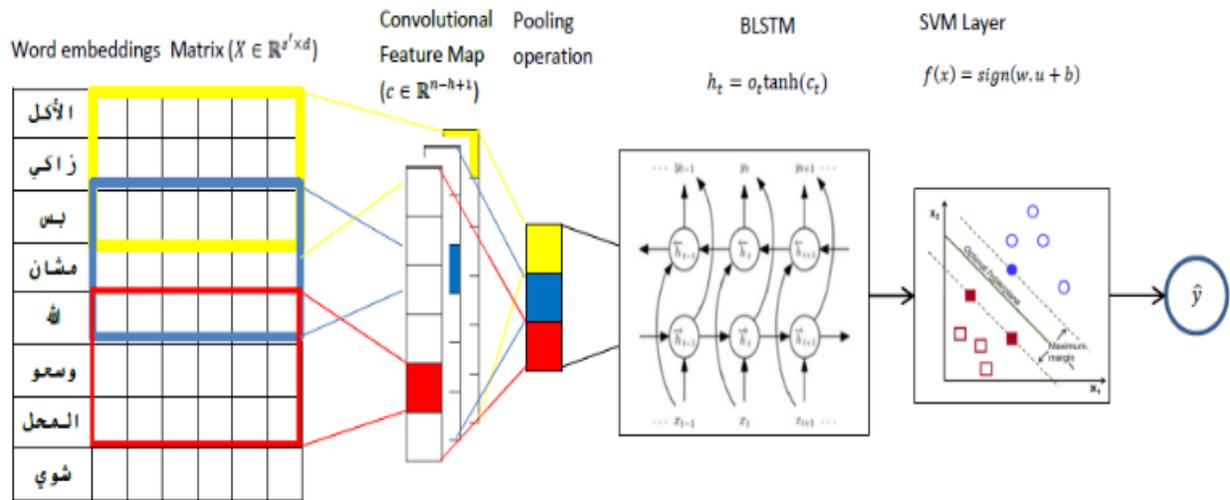


Fig. 3. Architecture of the Proposed Model.

IV. EXPERIMENTS

This section presents the datasets, models settings, and evaluation metrics. For evaluation, we developed some experiments to compare our model’s performance versus baseline models. Particularly, the author used the individual classifiers of CNN and SVM as baseline models. Additionally, we compared the proposed model results with some of the state-of-the-art deep learning models proposed by other related studies. The experiments also explore the best pre-trained word embeddings architecture for the proposed model. For this purpose, each performance experiment is carried out with CBOW and Skip-gram. To perform all the experiments in this work, we used the Keras library. More details about the experiments are provided in the following subsections.

A. Datasets

We used three publicly available datasets, namely, LABR, OMCCA, and the dataset presented in [35]. These datasets vary in size and writing form, where MSA and different Arabic dialects were detected. For example, LABR contains over 63000 reviews written mostly in MSA with the presence of dialectal phrases. On the other hand, On the other hand, OMCAA consists of over 28000 reviews written mostly in Jordanian and Saudi Arabic dialects. The last dataset presented 2400 reviews written in Jordanian dialect with the presence of MSA. All the datasets are annotated on the document level, and we considered only two polarity classes, which are positive and negative. The training/validation/testing split is set to 70%, 15%, 15%, respectively, in our experiments. Due to the limited capabilities of the device used to perform the experiments, we considered only the randomly balanced versions of these datasets. Table I presents more details about datasets.

TABLE I. CHARACTERISTICS OF DATASETS USED IN OUR EXPERIMENTS

Dataset	Balanced	Review Type
LABR	16444	Books
OMCCA	4990	Products
[35]	2400	products

B. Evaluation Metrics

The performance is quantified using the following evaluation metrics: F1-score, Precision, and Recall; see Equations 3, 4 and 5.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - score = 2 \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

where TP, TN, FP FN indicate a true positive, true negative, false positive, false negative, respectively.

C. Hyper-parameters Settings

This section describes the hyper-parameters settings for each model used in the experiments. The process of selecting the optimal hyper-parameters is a challenging task, and it varies based on the characteristics of the dataset. To this end, we performed several trials to choose the hyper-parameters with which the classifiers may yield the best performance results. As a result, the hyper-parameters in Table II have been assigned to the layers of CNN and Bi-LSTM. It is worth mentioning that all the data points have been zero-padded before being passed to the deep learning models, so all the input vectors have the same length. In the case of the CNN is a baseline model, the word vectors will be fed to a fully connected layer of size 64 nodes. The dense hidden layer is trained and regularized using a Relu function and l₂-norm method, respectively. The weights are then passed to an output layer with a sigmoid function to give the final classification probability. We also applied a Dropout layer after each model and early stopping strategy to reduce the over-fitting problem. Then, we added a max-pooling layer after each layer to keep the most important features. During the networks’ training process, Adam algorithm) is employed to perform the optimization with a learning rate of 0.001, and binary cross-entropy function for loss minimizing. A linear SVM classifier with its default parameters in Keras library has been employed for the classification process. Regarding generating word

embeddings based on Aravec, after many trials, we decided to apply the skip-gram and CBOW models built on Web pages with tokens of 132,750,000 and embedding dimension of 300. For words that are not contained in the pre-trained model, the embedding is set to a vector of zeros.

TABLE II. HYPER-PARAMETERS OF CNN AND BI-LSTM

Parameter	CNN	Bi-LSTM
Filters	250	-
Kernel size	4	-
Strides	3	-
Output dimensions	-	150
Activation	Relu	Tanh
Dropout	0.5	0.5
Max Pooling	2	2
Epoch size	30	30
Batch size	30	30
Loss function	binary_crossentropy	binary_crossentropy
Optimizer	Adam	Adam
L2 regularization	0.01	0.05

V. RESULTS

Results of the proposed model against the baseline models on the datasets are summarized in Table III. As it can be seen, our model with skip-gram and CBOW outperforms the other models in all datasets. The results show a significant improvement in the F1-score of the proposed model in all datasets, where the highest is achieved in the third dataset with around 8% compared to CNN and SVM when skip-gram is used and around 7% when CBOW is applied. Although the

proposed model outperforms SVM and CNN with CBOW on OMCCA, the highest results are obtained using skip-gram with an improvement of 7% and 8% compared to CNN and SVM, respectively. It is also noticed that the proposed model show improvement on the LABR dataset when skip-gram is used with around 6% and 7% compared to CNN and SVM, respectively, and around 5% for both models in the case of skip-gram.

Additionally, the proposed model performs very well in terms of the recall and precision in all dataset with a bit of trade-off, where the recall is the highest compared to all models used with 91.6% on the third dataset. On the other hand, the highest precision value is around 91% and achieved by CNN on the OMCCA dataset. Based on the two-word embeddings models, CNN achieved classification performance close to SVM with slight superiority to the former. This might indicate that the representation captured by both models is not enough to identify the latent connections between words, and a richer representation is required for better performance.

Finally, after the proposed model has been proven to be effective in Arabic sentiment classification for all datasets used in this work, we compare our results with other studies' results that employed state-of-the-art deep learning models. We choose work that used LABR dataset as it is the most common dataset in the Arabic SA literature. To guarantee a fair comparison, we only consider those who experimented on balanced class labels and used Aravec pre-trained model. In this sense, the work in [64] reported a significant improvement in Arabic sentiment classification based on a combination of Bi-LSTM and CNN on different datasets, including LABR. However, our proposed model outperforms their model, where they achieved an F1-score of 76.9%, which is 10% less than the highest results of our model on LABR.

TABLE III. RESULTS OF A COMPARISON BETWEEN THE PROPOSED AND BASELINE MODELS ON THE USED DATASETS WITH SKIP-GRAM AND CBOW EMBEDDINGS MODELS

Dataset	Model	Representation	Recall	Precision	F1-score	
LABR	CNN	Skip-gram	73.05	89.24	80.34	
		CBOW	76.68	82.68	79.56	
	SVM	TF-IDF	74.46	86.49	79.08	
		Proposed Model	Skip-gram	85.92	86.36	86.15
			CBOW	82.0	86.31	84.10
OMCAA	CNN	Skip-gram	74.70	91.30	82.17	
		CBOW	76.28	87.72	81.60	
	SVM	TF-IDF	83.26	79.06	81.11	
		Proposed Model	Skip-gram	90.0	89.28	89.64
			CBOW	88.35	84.29	86.27
[35]	CNN	Skip-gram	74.16	90.81	81.65	
		CBOW	77.27	85.53	81.19	
	SVM	TF-IDF	75.83	88.77	81.58	
		Proposed Model	Skip-gram	91.66	86.61	89.06
			CBOW	88.33	86.88	87.60

VI. CONCLUSIONS AND FUTURE WORK

This paper presented a model that uses a linear SVM classifier on top of a combination of CNN and Bi-LSTM for Arabic sentiment classification. Unlike the conventional architecture of the deep learning model, the proposed model was tailored by replacing the fully connected layer with an SVM classifier, which receives the embedded vectors extracted by CNN and Bi-LSTM. The experimentation has shown the effectiveness of the proposed model with a significant improvement over the baseline models. Furthermore, we showed that the proposed model outperforms one of the state-of-the-art deep learning models with a considerable improvement. Yet, there is room for improvement as the proposed model does not concern with some issues that might affect the performance (e.g. negation handling). Further work is also required to explore the effectiveness of the proposed model with deeper layers and diverse architectures on different Arabic benchmark datasets. Additionally, to extend the work so that it can be applied to other Arabic SA tasks such as aspect-based sentiment analysis.

REFERENCES

- [1] M. Al-Ayyoub, et al. A Comprehensive Survey of Arabic Sentiment Analysis. *Information processing & management*. 56(2), pp. 320-342, 2019.
- [2] I. A. Al - Sughayer and Al - Kharashi, I. A. Arabic Morphological Analysis Techniques: A Comprehensive Survey. *Journal of the American Society for Information Science and Technology*. 55(3), pp. 189-213, 2004.
- [3] S. R. El-Beltagy and Ali, A. "Open Issues in the Sentiment Analysis of Arabic Social Media: A Case Study," in 9th International Conference on Innovations in Information Technology. IEEE. Abu Dhabi, pp. 215-220, 2013.
- [4] K. Shaalan, Bakr, H., and Ziedan, I. "Transferring Egyptian Colloquial Dialect into Modern Standard Arabic," in International Conference on Recent Advances in Natural Language Processing (RANLP-2007), Borovets, Bulgaria. pp. 525-529, 2007.
- [5] A. L. Maas, et al. "Learning Word Vectors for Sentiment Analysis," in Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies. Association for Computational Linguistics. Portland, Oregon, USA, pp. 142-150, 2011.
- [6] Y. Kim. "Convolutional Neural Networks for Sentence Classification," in Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics. Doha, Qatar, pp. 1746-1751, 2014.
- [7] D. Tang, et al. "Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics. Baltimore, Maryland, pp. 1555-1565, 2014.
- [8] F. Xu, Zhang, X., and Yang, A. Investigation on the Chinese Text Sentiment Analysis Based on Convolutional Neural Networks in Deep Learning. *CMC-Computers, Materials & Continua*. 58(3), pp. 697-709, 2019.
- [9] D. Tang, Qin, B., and Liu, T. Deep Learning for Sentiment Analysis: Successful Approaches and Future Challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 5(6), pp. 292-303, 2015.
- [10] L. Deng and Liu, Y. *Deep Learning in Natural Language Processing* (1st ed). Springer. Singapore, 2018, DOI: 10.1007/978-981-10-5209-5.
- [11] T. Mikolov, et al. "Efficient Estimation of Word Representations in Vector Space," in 1st International Conference on Learning Representations. Arizona, USA, pp. 2013.
- [12] J. Turian, et al. "A Preliminary Evaluation of Word Representations for Named-Entity Recognition," in NIPS Workshop on Grammar Induction, Representation of Language and Language Learning. pp. 1-8, 2009.
- [13] O. Levy and Goldberg, Y. "Neural Word Embedding as Implicit Matrix Factorization," in Advances in neural information processing systems. pp. 2177-2185, 2014.
- [14] Y. Bengio, et al. A Neural Probabilistic Language Model. *Journal of machine learning research*. 3(Feb), pp. 1137-1155, 2003.
- [15] R. Socher, et al. "Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank," in Proceedings of the 2013 conference on empirical methods in natural language processing. pp. 1631-1642, 2013.
- [16] M. Cliche. "Bb_Twtr at Semeval-2017 Task 4: Twitter Sentiment Analysis with Cnns and Lstms," in Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017). Association for Computational Linguistics. Vancouver, Canada, pp. 573-580, 2017.
- [17] L. Dong, et al. "Adaptive Recursive Neural Network for Target-Dependent Twitter Sentiment Classification," in Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 2: Short papers). pp. 49-54, 2014.
- [18] A. Dahou, et al. "Word Embeddings and Convolutional Neural Network for Arabic Sentiment Classification," in Proceedings of coling 2016, the 26th international conference on computational linguistics: Technical papers. pp. 2418-2427, 2016.
- [19] S. Al-Azani and El-Alfy, E.-S. M. "Hybrid Deep Learning for Sentiment Polarity Determination of Arabic Microblogs," in International Conference on Neural Information Processing. Springer. pp. 491-500, 2017.
- [20] A. M. Alayba, et al. "Arabic Language Sentiment Analysis on Health Services," in 2017 1st International Workshop on Arabic Script Analysis and Recognition (ASAR). IEEE. pp. 114-118, 2017.
- [21] R. Baly, et al. A Sentiment Treebank and Morphologically Enriched Recursive Deep Models for Effective Sentiment Analysis in Arabic. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*. 16(4), pp. 23, 2017.
- [22] A. M. Alayba, et al. "Improving Sentiment Analysis in Arabic Using Word Representation," in 2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR). IEEE. pp. 13-18, 2018.
- [23] A. Shoukry and Rafea, A. "Sentence-Level Arabic Sentiment Analysis," in the International Conference on Collaboration Technologies and Systems. IEEE. Colorado, pp. 546-550, 2012.
- [24] O. Al-Harbi. Classifying Sentiment of Dialectal Arabic Reviews: A Semi-Supervised Approach. *International Arab Journal of Information Technology*. 16(6), pp. 995-1002, 2019.
- [25] J. A. Suykens and Vandewalle, J. Training Multilayer Perceptron Classifiers Based on a Modified Support Vector Method. *IEEE Transactions on Neural Networks*. 10(4), pp. 907-911, 1999.
- [26] Y. Tang. "Deep Learning Using Linear Support Vector Machines," in International Conference on Machine Learning 2013: Challenges in Representation Learning Workshop. Georgia, USA, pp. 2013.
- [27] A. F. Agarap. An Architecture Combining Convolutional Neural Network (Cnn) and Support Vector Machine (Svm) for Image Classification. arXiv preprint arXiv:1712.03541. pp., 2017.
- [28] J. Nagi, et al. "Convolutional Neural Support Vector Machines: Hybrid Visual Pattern Classifiers for Multi-Robot Systems," in 2012 11th International Conference on Machine Learning and Applications. IEEE. pp. 27-32, 2012.
- [29] A. F. M. Agarap. "A Neural Network Architecture Combining Gated Recurrent Unit (Gru) and Support Vector Machine (Svm) for Intrusion Detection in Network Traffic Data," in Proceedings of the 2018 10th International Conference on Machine Learning and Computing. ACM. pp. 26-30, 2018.
- [30] F.-J. Huang and LeCun, Y. "Large-Scale Learning with Svm and Convolutional Nets for Generic Object Categorization," in Proceeding. Computer Vision and Pattern Recognition Conference (CVPR'06). pp. 2006.
- [31] A. Alalshakmubarak and Smith, L. S. "A Novel Approach Combining Recurrent Neural Network and Support Vector Machines for Time Series Classification," in Proceeding 9th International Conference on Innovations in Information Technology (IIT). IEEE. pp. 42-47, 2013.

- [32] M. S. Akhtar, et al. "A Hybrid Deep Learning Architecture for Sentiment Analysis," in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. pp. 482-493, 2016.
- [33] M. Nabil, Aly, M., and Atiya, A. Labr: A Large Scale Arabic Sentiment Analysis Benchmark. arXiv preprint arXiv:1411.6718. pp., 2014.
- [34] A. Y. Al-Obaidi and Samawi, V. W. "Opinion Mining: Analysis of Comments Written in Arabic Colloquial," in Proceedings of the World Congress on Engineering and Computer Science. pp. 2016.
- [35] O. Al-Harbi. Using Objective Words in the Reviews to Improve the Colloquial Arabic Sentiment Analysis. International Journal on Natural Language Computing. 6(3), pp. 1-14, 2017.
- [36] T. Khalil, et al. "Which Configuration Works Best? An Experimental Study on Supervised Arabic Twitter Sentiment Analysis," in the First International Conference on Arabic Computational Linguistics. IEEE. Cairo, pp. 86-93, 2015.
- [37] A. Aliane, et al. "A Genetic Algorithm Feature Selection Based Approach for Arabic Sentiment Classification," in the 13th International Conference of Computer Systems and Applications. IEEE. Agadir, pp. 1-6, 2016.
- [38] I. Guellil and Azouaou, F. "Arabic Dialect Identification with an Unsupervised Learning (Based on a Lexicon). Application Case: Algerian Dialect," in 2016 IEEE Intl Conference on Computational Science and Engineering (CSE) and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES). IEEE. pp. 724-731, 2016.
- [39] S. Tartir and Abdul-Nabi, I. Semantic Sentiment Analysis in Arabic Social Media. Journal of King Saud University-Computer and Information Sciences. 29(2), pp. 229-233, 2017.
- [40] A. Shoukry and Rafea, A. "A Hybrid Approach for Sentiment Classification of Egyptian Dialect Tweets," in the First International Conference on Arabic Computational Linguistics. IEEE. Cairo, pp. 78-85, 2015.
- [41] S. Zhou, Chen, Q., and Wang, X. "Active Deep Networks for Semi-Supervised Sentiment Classification," in Proceedings of the 23rd International Conference on Computational Linguistics: Posters. Association for Computational Linguistics. pp. 1515-1523, 2010.
- [42] X. Glorot, Bordes, A., and Bengio, Y. "Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach," in Proceedings of the 28th international conference on machine learning (ICML-11). pp. 513-520, 2011.
- [43] C. Dos Santos and Gatti, M. "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts," in Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. pp. 69-78, 2014.
- [44] D. Tang, Qin, B., and Liu, T. "Aspect Level Sentiment Classification with Deep Memory Network," in Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics. Austin, Texas, pp. 214-224, 2016.
- [45] X. Wang, Jiang, W., and Luo, Z. "Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts," in Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers. pp. 2428-2437, 2016.
- [46] M. Heikal, Torki, M., and El-Makky, N. Sentiment Analysis of Arabic Tweets Using Deep Learning. Procedia Computer Science. 142, pp. 114-122, 2018.
- [47] A. B. Soliman, Eissa, K., and El-Beltagy, S. R. Aravec: A Set of Arabic Word Embedding Models for Use in Arabic Nlp. Procedia Computer Science. 117, pp. 256-265, 2017.
- [48] M. Nabil, Aly, M., and Atiya, A. "Astd: Arabic Sentiment Tweets Dataset," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. pp. 2515-2519, 2015.
- [49] M. Abdullah, Hadzikadicy, M., and Shaikhz, S. "Sedat: Sentiment and Emotion Detection in Arabic Text Using Cnn-Lstm Deep Learning," in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE. pp. 835-840, 2018.
- [50] R. M. Alahmary, Al-Dossari, H. Z., and Emam, A. Z. "Sentiment Analysis of Saudi Dialect Using Deep Learning Techniques," in 2019 International Conference on Electronics, Information, and Communication (ICEIC). IEEE. pp. 1-6, 2019.
- [51] A. A. Altowayan and Tao, L. "Word Embeddings for Arabic Sentiment Analysis," in 2016 IEEE International Conference on Big Data (Big Data). IEEE. pp. 3820-3825, 2016.
- [52] A. Barhoumi, et al. "An Empirical Evaluation of Arabic-Specific Embeddings for Sentiment Analysis," in International Conference on Arabic Language Processing. Springer. pp. 34-48, 2019.
- [53] N. Abdelhade, Soliman, T. H. A., and Ibrahim, H. M. "Detecting Twitter Users' Opinions of Arabic Comments During Various Time Episodes Via Deep Neural Network," in International Conference on Advanced Intelligent Systems and Informatics. Springer. pp. 232-246, 2017.
- [54] A. Al Sallab, et al. "Deep Learning Models for Sentiment Analysis in Arabic," in Proceedings of the second workshop on Arabic natural language processing. Association for Computational Linguistics. Beijing, China, pp. 9-17, 2015.
- [55] M. Al-Smadi, et al. Using Long Short-Term Memory Deep Neural Networks for Aspect-Based Sentiment Analysis of Arabic Reviews. International Journal of Machine Learning and Cybernetics. 10(8), pp. 2163-2175, 2019.
- [56] M. Al-Smadi, et al. Deep Recurrent Neural Network Vs. Support Vector Machine for Aspect-Based Sentiment Analysis of Arabic Hotels' Reviews. Journal of computational science. 27, pp. 386-393, 2018.
- [57] J. Brownlee. Deep Learning for Natural Language Processing: Develop Deep Learning Models for Your Natural Language Problems Machine Learning Mastery. 2017,
- [58] R. Collobert, et al. Natural Language Processing (Almost) from Scratch. Journal of machine learning research. 12(Aug), pp. 2493-2537, 2011.
- [59] S. Hochreiter and Schmidhuber, J. Long Short-Term Memory. Neural computation. 9(8), pp. 1735-1780, 1997.
- [60] M. Sundermeyer, Schlüter, R., and Ney, H. "Lstm Neural Networks for Language Modeling," in Thirteenth annual conference of the international speech communication association. International Speech Communication Association (ISCA). Portland, Oregon, pp. 2012.
- [61] M. Schuster and Paliwal, K. K. Bidirectional Recurrent Neural Networks. IEEE Transactions on Signal Processing. 45(11), pp. 2673-2681, 1997.
- [62] A. Graves, Jaitly, N., and Mohamed, A.-r. "Hybrid Speech Recognition with Deep Bidirectional Lstm," in 2013 IEEE workshop on automatic speech recognition and understanding. IEEE. pp. 273-278, 2013.
- [63] C. Cortes and Vapnik, V. Support-Vector Networks. Machine Learning. 20(3), pp. 273-297, 1995.
- [64] K. A. Kwaik, et al. "Lstm-Cnn Deep Learning Model for Sentiment Analysis of Dialectal Arabic," in International Conference on Arabic Language Processing. Springer. pp. 108-121, 2019.

Emotional Intelligence Robotics to Motivate Interaction in e-Learning: An Algorithm

Dalia khairy¹, Salem Alkhalaf², M. F. Areed³, Mohamed A. Amasha⁴, Rania A. Abougalala⁵

Department of Computer Teacher Preparation, Damietta University, Damietta, Egypt^{1,4,5}

Department of Computer Science, Qassim University, Alrass, Saudi Arabia²

Department of Computer Science, Damietta University, Damietta, Egypt³

Abstract—The development of emotional intelligence robotics in the learning environment plays valuable support for social interaction among students. Emotional intelligence robots should be scalable to recognize emotions, appear empathetic in learning situations, and enrich the confidence with students for active interaction. This paper presents some related issues about integrating emotional intelligence robotics in E-learning such as its role and outcomes to motivate interaction during education and discover the main aspects of the emotional intelligence between humans and robots. This paper aims to determine the design requirements of emotional robots. Besides, this paper proposed a framework of educational Robotics with Emotional Intelligence in Learning (EREIL). EREIL consists of three main units; student emotions discovery, student emotions representation, and EREIL-Student Communication (RSC). In addition, it introduces a perception of EREIL working. In the future, this paper tries to merge more sensor devices and machine learning algorithms to integrate face analysis with speech recognition. Besides, it can add a persuasion unit in the EREIL robot to convince students with better learning choices to their abilities.

Keywords—Robotics; emotional intelligence; interaction; e-learning; motivation; robot with emotional intelligence; machine learning algorithms; face analysis; speech recognition

I. INTRODUCTION

Educational Robotics(ER) has become increasingly common to handle possible future adaptation issues in E-learning. ER presents a deeply exciting learning environment in college, encouraging collaboration among learners, and realizing the generation of innovative experience in a reflexive method instructed by the teacher[1]. Recently, Teachers have often been struggling with keeping students' motivation and social interaction to the E-learning process. Student motivation is the backbone of the E-learning process because it affects student performance. This reflects the fact that ER can help to overcome the special educational needs of students such as Autism Spectrum Disorders(ASD)[2].

Furthermore, ER is an innovative field for students to build E-learning experiences [3][4]. It is related to both the observation and simple handling of a robot as well as more demanding tasks as the students are involved in planning, problem-solving, and decision making in relation to the robot's behavior, with the aim of developing creative thinking, the highest level of thinking [5]. ER for students with special educational needs includes the use of social robots for

teaching various skills such as cognitive function, communication, and collaborative play[6][7].

Emotions have a complex structure to include specific social-cognitive dimensions to analyze things from the perspective of others and sharing other people's empathy[8][9]. Besides, emotion detection is a significant issue for special education, in two essential directions. The first regards students with special needs themselves who appear to have a deficiency in social perception and cognition[10]. The second direction regards children with insufficient mastering of social skills do not communicate and relate to others efficiently while they experience problems in playing, working, and learning with peers, which may well result in a certain degree of isolation [11].

Recently, emotional intelligence robots contribute to enhancing collaborative knowledge and individual and corporate self-efficacy[12]. In turn, they play a valuable role in promoting academic accomplishment, and cognitive-motivational features of learning in a student with intellectual disability increased significantly through the influence of robotics activities [13]. Furthermore, emotional intelligence robots enhance memory and attention for students with insufficient social experiences[14][15]. Besides, emotional intelligence robots encourage the processing of the complexity of human emotional expressions[16]. Thus, the field of Affective Computing requires to be updated and contributed with researches and experiences of emotional intelligence robots. Additionally, some practices in this field support improving creativity and sequential thinking for students[17].

The rest of the paper is organized as follows Section 2, highlights the emotional intelligence of humans and robots. In Section 3 the investigation of gender effects in students' perception is presented. Sections 4 and 5 look into age and educational robotics with educational intelligence and outcome of educational robotics with emotional intelligence in learning environment, respectively. Sections 6 and 7 display a framework of educational Robotics with Emotional Intelligence in Learning(EREIL); requirements, components, and an example of EREIL working. Section 8 concludes the paper proposing and future work to improve EREIL.

II. EMOTIONAL INTELLIGENCE FOR HUMANS AND ROBOTS

Robotics encourages integrative knowledge because of its inherent variation. Consequently, Robotics merges the learning among computational aspects, sensor devices, and

expanding rapidly to emotional intelligence. Hence, it has the potential to excite and support talent by promoting technically-minded students. Robotics has particular value in this learning domain because it depends on computational thinking and has a powerful experimental domain for sensing, measuring, designing behavior, and emotional intelligence [18][19].

Emotional intelligence enriches the quality of the learning climate and improves academic performance. Besides, Emotional intelligence provides more knowledge to solve learning obstacles, support better learning organization, enhance personal growth, and handle stress through learning environments[20][21]. Emotional intelligence enhances richness with students and provides further versatility, adjustability, and skill at working in groups in the learning environment[22].

Emotional robots have played a vital role in the future of artificial intelligence when we focus on relational agents. Furthermore, relational agents present human social experiences, such as hand signs, position changes, and facial feelings, spoken words that sent fun, empathy, culture, and kindness[23][24]. Besides, emotional intelligence is an essential part of human intelligence and cognitive performance in the learning environment. Hence, emotional intelligence promotes the human connection with students and influencing their emotional progress[25].

The communication between students and an emotional robot influences the student's understanding of an emotional robot. Also, it concerns the student's performance during the interaction. Thus, students cooperate with an emotional robot that generates more significant advances in trust and works excellently on the assignment taught by an emotional robot. Besides, emotional robot converges on how to perceives and promotes emotions in particular communicative conditions. Furthermore, an emotional robot in learning environment affects to the emergence of team dynamics methods like; communication, collaboration, team administration[26].

The researches prove that the variation of the stimuli has a great role in human-robot interactions. The stimuli like; text, audio, video, and image. Text stimulus led to even higher perceived emotional intelligence of the student than the video, audio, and image stimuli, while the variation in the stimuli-type does not affect the perceived emotional intelligence of the robot. Furthermore, the high emotional intelligence robots and humans as more trustworthy than the low emotional intelligence robots and humans. Besides, the content of the verbal interaction is essential for students' perceptions of emotional intelligence in both students and robots and not the tone of voice, posture, gesturing, or other bodily dynamics. The emotional robots in the learning environment might have to be endowed at least primary emotional intelligence capabilities. These capabilities enable them to focus on the students' emotional state, establish and maintain students' trust and build working relationships with them [27].

III. STUDENTS' PERCEPTION OF GENDER EFFECTS IN DESIGNING ROBOTICS

The emotional intelligence of students integrates gender parameters when making decisions and formed expectations in social interaction. In turn, that affects designing educational robotics[28]. Also, students have a positive attitude in their awareness, behavior, and perception towards the educational robotics have the same gender. That because the thought based on the same gender educational robotics is more believable, reliable, interactive, and engaging. Although the gender features offered in appearance more than objective parameters. For example, female students prefer educational robotics with pink lips, with a female name, and a female voice. Versus male students prefer educational robotics with gray lips, with a male name, and a male voice. As a result, this makes them feel more comfortable in social interaction, more desirable, more accepting, and more perfectly capable of performing a service task[28][29].

The female has high emotional intelligence than the male in human interactions. And the same goes for educational robotics. Consequently, students may lose their trust in male educational robotics. That indicates the essential for enhancing tools of male educational robotics to acquire them high emotional intelligence. Thus, disappearing gender-based differences in emotional intelligence [27].

Various strategies are adopted to facilitate optimal levels of perceived emotional intelligence and trust in educational robotics. For example, students will previously make decisions when discovering the gender of educational robotics regarding emotional intelligence and social interaction. As a solution, the design of educational robotics presenting emotional intelligence would be in a female voice. That is due to male educational robotics might disappoint expectations. However, for performing tasks that include sensitive social situations a male voice of educational robotics might be superior to reduce undesired results towards potential errors [30].

IV. AGE AND EDUCATIONAL ROBOTICS WITH EDUCATIONAL INTELLIGENCE

Age is relevant to expanded learning and practice experience. More youthful students manage to be limited exposure to while more experienced students seem to be extra excited with educational robotics with emotional intelligence [31]. This new trend should support various platforms, application interfaces to encourage a creative learning environment. Thus, these platforms seem to be quite promising for young students because they can reduce the age entrance for sharing in educational robotics with emotional intelligence and programming skills [32]. The integrated technologies and hybrid systems can facilitate engaging educational robotics with emotional intelligence for both young and older students. In turn, that can decrease cognitive content for students by engaging a new trend of learning through learning curriculum simultaneously [33].

V. OUTCOMES OF EDUCATIONAL ROBOTICS WITH EMOTIONAL INTELLIGENCE IN E-LEARNING ENVIRONMENT

The outcomes can be details as follows [34]:

- ER with EI assists to build an innovative learning environment and creates a valuable and attractive learning style.
- ER with EI integrates physical and mental experience for communication with the learning situations and their instruments.
- ER with EI enables students to discover by performing, managing conceptual theories, and integrating knowledge.
- ER with EI may encourage motivation for studying in courses that are commonly perceived by students as boring and not so exciting.
- ER with EI promotes knowledge spirit for both children and adult students who are further enthusiastic about robotics evolution.
- ER with EI adopted a facilitator learning environment for both students and professors. Consequently, professors prove high investment and concrete results in the use of EI in the studying environment.
- ER with EI provides more inclusive engagement in learning activities to retain students' concentration in the long-term and makes the professors' teaching further satisfying.

- ER with EI enriches the sufficient and motivation of the teaching method. Besides, it overcomes the boredom of traditional education methods.

VI. A FRAMEWORK OF EDUCATIONAL ROBOTICS WITH EMOTIONAL INTELLIGENCE IN LEARNING (EREIL)

This section describes the proposed framework called EREIL. EREIL aims to motivate interaction among students in the learning environment. EREIL improves to assist students in promoting their social connections during the daily performance in the learning environment. This section divides into two main subparts: the design requirements of EREIL and the components of EREIL.

A. The Requirements of EREIL

The requirements of EREIL can be divided into three types; needs determination & analysis, design, and technology.

1) *Needs determination and analysis requirements:* These requirements analyze the steps to integrate EREIL in e-learning environment as seen in Fig. 1.

2) *Design requirements:* These requirements regarding design principles of EREIL, the requirements explained in Fig. 2.

3) *Other requirements:* These requirements can be represented into three requirements. First, encouraging innovative teaching methods: It includes adapting teamwork and critical thinking. Second, developing programmable systems: It includes focusing on enriching observation, analysis, supporting modeling and simulation. Third, adopting project-based learning: It includes learning-based problem solving and adapting cultivating collaboration skills.

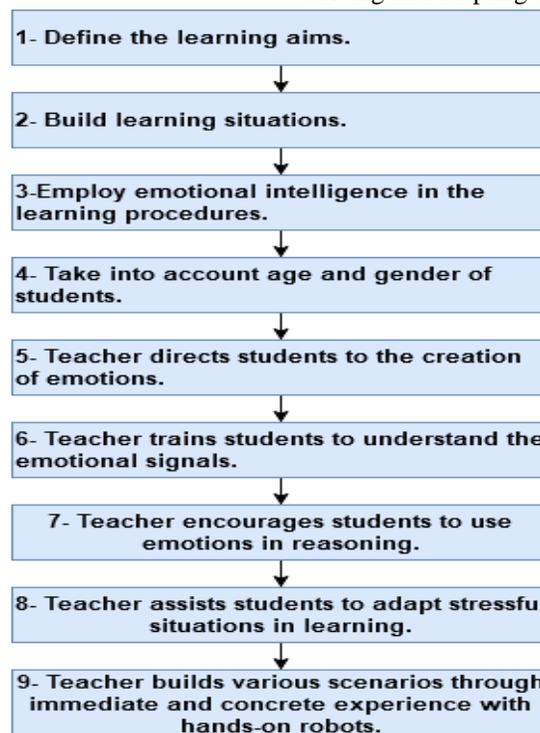


Fig. 1. Needs Determination and Analysis Requirements of Emotional Intelligence Robots.

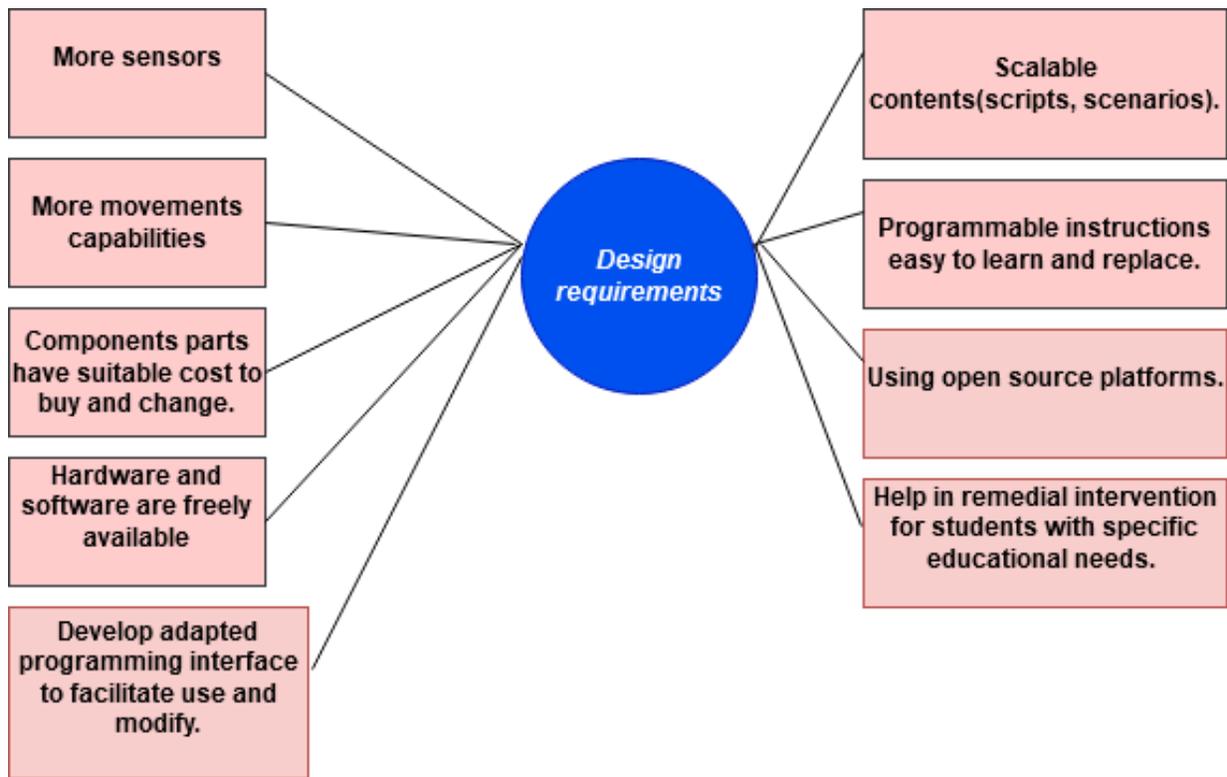


Fig. 2. Design Requirements of Emotional Intelligence Robots.

B. The Components of EREIL

EREIL consists of three main units; student emotions discovery, student emotions representation, and EREIL-Student Communication (RSC) (see Fig. 3). It is clear that the two units responsible for discovering, implementing, and processing, and representing emotions. Furthermore, the third unit enhances the student practice and supplies a running visible interface to the EREIL robot.

EREIL can implement the following essential characteristics:

- It recognizes the emotions of several students.

- It determines the common sentiment of a collection of students.
- It shows student determined sentiments according to the emotional moods of the students.
- It helps students supporting the learning schedule.
- It stores learning tasks to tell students to perform them at a determined duration.
- It convinces students by providing them reasons to confirm suggested learning activities according to their learning profile and individual abilities.

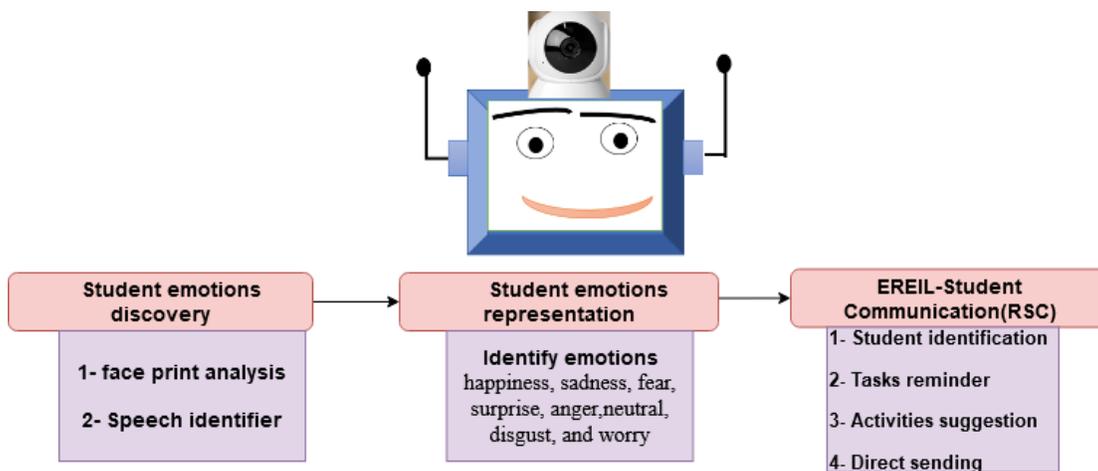


Fig. 3. Components of EREIL Robot.

1) *Student emotions discovery*: The discovery process of students' emotions plays an essential role in social interaction in the learning environment. Thus, the implementation of emotions discovery helps students to practice sharing, interaction, and collaboration. Also, it can be useful to recommend an adaptive response or recommend modifications for a better response.

EREIL is a proposed framework to support sensing devices, recognize and represent some factors of student performance. EREIL supports the student's body signals (i.e., emotions, speech level, gestures, etc.) to recognize these communications performed in the learning environment. EREIL discovers student's body signals through contacting eyes, gestures, face analyzing, and voice recognition. Through this discovery, EREIL can determine student identification. Besides, it is important to increase sensors that support EREIL to obtain students identification. Although sensor devices are not sufficient in EREIL, it is necessary to support artificial intelligence (AI) fields. AI allows to analysis and training of the collected raw data from sensor devices. With AI potentials, we can use face identification, speech recognizing, interpret natural language, and emotional discovery.

a) *EREIL for student's determination and sentiment analysis*: EREIL determines the student identification and other students who can share the learning activities with him. It aims to select suitable machine learning techniques that can contribute to students' face recognition. This process should be agile and easy to decrease the training time and not confuse the student. Furthermore, the student face recognition process depends on converting face to vectors with special determinates. Every face has specific landmarks that can be measured, such as the distance between the eyes, the length of the jaw line, the shape of the eyebrows, and more. Together, they create a unique faceprint that can later be compared with other students until a match is found, thus recognize student.

EREIL allows classifying emotions through a deep learning machine to analyze students' sentiments. EREIL classifies students' emotions to happiness, sadness, fear, surprise, anger, neutral, disgust, and worry. Besides, EREIL can use the web service to train and examine students' emotions because this procedure can be a difficult and high cost if it computes on EREIL.

b) *Emotional web assistance for EREIL*: As mentioned, to reduce the overload of computational analysis on EREIL for determining students' identifications and emotions, EREIL depends on the webserver. This web assistance supports EREIL to recognize students' faces, analyze emotions, and recognize the sharing emotions among a group of students. EREIL encourages the students to optimize the decision-making in communication with other students and enhance taking decisions regarding cooperating and sharing learning activities.

Nevertheless, if EREIL does not know the student identification. EREIL displays a collection of questions to promote communication with the student for discovering the student's name, face, and identification. Then, EREIL creates a conversation with the student to recognize more information

about him and stores the new data. Furthermore, EREIL separately recognizes students in groups. In other words, EREIL recognizes each student alone and takes assistance from a web server for analyzing face. Then, if EREIL discovers another unknown student, the same process will repeat. Also, recognizing the emotions of a students' group allows EREIL to understand what the sensitive mood of the group is. Consequently, EREIL tells the student to perform alone if this emotion is further away from his interests.

2) *Student emotions representation*: EREIL contributes to understanding emotions for students. And that is in turn influences on changing students' reactions. As a result, it is essential for EREIL to represent expressions of its own emotional situation. Thus, these representations would be directly benefited the interaction with students. So, EREIL uses sensors devices to recognize the learning environment and use this knowledge to express its emotional situation.

EREIL can express emotions by sensors of its face like contact eyes to start and manage communication. Looking at student's eyes during speech attracts attention and integrates interaction during the speech.

Face tracking technology can assist in face detection. While the student moves his eyes towards learning activities, EREIL can follow this moving to start and maintain eye contact. Thus in turn encourages the student's attention and promotes interaction. Besides, this information enriches EREIL to start of dialog at a suitable time, greet the student with a friendly smile. Furthermore, EREIL keeps eyebrows and mouth to express different emotions. Also, eye contact contributes to measure engagement between students and EREIL, better understand students' demands, and produce more appropriate information.

3) *EREIL- Student Communication (RSC)*: The student's communication is promoted with EREIL using RSC. RSC is the unit is responsible for enhancing the student practice in learning activities and supplying a running visible interface to the EREIL robot. RSC is a supporter for student to increase learning motivation through remembered daily learning assignments and proposing activities. Besides, RSC explains new activities to expand the learning experiences of students and recommends with external learning sources to encourage rounded learning by integrated the practical exercises with academic learning.

RSC consists of four subunits: 1) student identification, 2) tasks reminder, 3) activities suggestion, 4) direct sending.

- **Student identification**: This subunit stores the personal data and learning progress information of each student. It RSC aims to develop the students' learning aspects. Hence, storage the students' learning activities, interests, preferences, and the right time to perform them. However, RSC needs the personal information; name, address, age, parents, closer friends, learning style, and preferred subjects to create recommendations and advice related to the personal environment.

- **Activities suggestion:** This subunit decides and recommends activities to the student based on personal data and interests. It aims to keep students in active mind and continuous research to reach rounded learning, e.g., use external video, audio, documentation films, reading reference books to connect academic learning with practical needs, engage other students in learning gamification. Besides, if closer students of the student are comprised a group on RSC, in turn RSC will recommend activities and learning games suitable to shared interests and in appropriate time to all of them.
- **Tasks reminder:** This subunit displays the activities, lessons, assignments of daily tasks to students. EREIL keep these learning tasks with date and time to tell each student about its schedule.
- **Direct sending:** This subunit presents a means to send information directly to EREIL through its touchscreen. When a student wants to send information to EREIL. Besides, EREIL uses direct sender to create an interactive situation among students and respond to students' decisions. On the other hand, a student can send a message through the visual interactive screen to

use EREIL sensors devices to catch information. It is essential to teach EREIL, by keeping interaction with the students.

Fig. 4 shows Algorithm 1 for how EREIL discovers the student's face. EREIL greets the student when EREIL discovers the student's face. Besides, it explains the EREIL behavior with unknown student. EREIL does not detect the student's face. EREIL will ask the student about his name, capture a photo, and stores it.

Besides, Fig. 5 shows Algorithm 2 for how EREIL reminds the student of the scheduled tasks. For example, EREIL reminds the student of the Science Lab time at 1:00 PM by displaying on the touch screen and wait for the student's response.

Furthermore, Fig. 6 shows Algorithm 3 for the activity suggestion from EREIL to the student. For example, EREIL suggests activity "play volley ball by displaying on the touch screen and wait for the student's response". In addition, Fig. 7 presents Algorithm 4 for direct sending between EREIL and student.

Algorithm 1: How EREIL discovers the student's face

Step1: Start

Step2: EREIL senses the existence of someone.

Step3: EREIL opens the camera to discover a student.

Step4: EREIL compares the captured photo of the student with stored photos of the students.

Step5: If the captured photo matches any of the stored photos **Then**

Appear a greeting message on a touchscreen

" **Welcome** + Student's name"

Step6: Else EREIL appears a message on a touchscreen

"Who are you?"

Step7: The student writes his name on a touchscreen.

Step8: EREIL takes a photo to store it and to detect the student later.

Step9: End

Fig. 4. Algorithm1 for Discovering the Student's Face.

Algorithm 2: How EREIL reminds the student of the scheduled tasks

Step1: Start

Step2: EREIL watches time.

Step3: EREIL keeps in touch with scheduled tasks.

Step4: EREIL compares between the current time and time of scheduled tasks.

Step5: If the current time matches any of the scheduled tasks **Then** EREIL alarms a student by displaying on a touchscreen.

Step6: Appear a message " Hi + student name". It is 1:00 PM. You have to go to Science Lab.

Step7: EREIL presents two choices on a touch screen.

Step8: If the student accepts to go the task **Then** Student will click Let's go button on a touch screen.

Fig. 5. Algorithm 2 for Tasks Reminder using EREIL.

Algorithm 3: How EREIL suggests activities to the student.

Step1: Start

Step2: EREIL watches time.

Step3: EREIL keeps in touch with activities suggestion.

Step4: EREIL compares between the current time and time of activities suggestions.

Step5: If the current time matches any of the activities suggestions **Then** EREIL alarms a student by displaying on a touchscreen.

Step6: Appear a message " Hi + student name". It is 3:00 PM. Let's play volleyball.

Step7: EREIL presents two choices on a touch screen.

Step8: If the student accepts to practice activity **Then** Student will click Let's go button on a touch screen.

Step9: Else student will click Another suggestion button on a touch screen.

Step10: If the student clicks Another suggestion button **Then** go to Step 6 with changing activity.

Step11: Else go to Step 12.

Fig. 6. Algorithm 3 for Activities Suggestion using EREIL.

Algorithm 4: Direct sending between EREIL and the student.

Step1: Start

Step2: Student requests EREIL to send a message.

Step3: EREIL appears touch screen for direct sending.

Step4: Student writes the request on EREIL touch screen.

Ex.: Student writes "suggest me an external resources about the human digestive system".

Step5: EREIL appears a message "Wait a few seconds. I will display some videos".

Step6: Student watches to videos.

Fig. 7. Algorithm 4 for Direct Sending using EREIL.

VII. PERCEPTION OF EREIL WORKING: AN EXAMPLE

To explain the performance of EREIL, this section displays an example that illustrates the various processes made by EREIL as an emotional robot recommender in the learning environment.

The EREIL will communicate with students by recognizing them, identifying their emotions, and recommending actions and activities in the learning environment.

The perception of EREIL performance as seen in Fig. 8:

1) EREIL starts performance when seeing a student or group of students.

2) EREIL starts a dialogue with the student and ask him some question to detect the identification. This case when EREIL does not recognize the student.

3) Besides, EREIL captures the student's face, analyzes the coordinates of the face (x, y), and encoding the photos to recognize identification.

Note: The camera of EREIL will locate in the center of EREIL's face to easily capture the student's face. Furthermore, EREIL periodically captures students' faces and stores them with learning progress, personal profiles, interests, activities, and schedule.

4) If EREIL recognizes the student's face. Then EREIL will greet the student with his name.

5) When EREIL knows the student, it will discover his emotion. If the EREIL identifies more than one student in the same group of students in the capturing, the manner is

returned for each face, beginning the computation of engagement emotion.

6) As soon as the students' emotion is detected, EREIL will show empathy with the student. Also, it can move his eyes, mouth, and eyebrows. Besides, EREIL can suggest some learning activities and control its sensors devices to suitable for student emotion.

7) After that, EREIL reminds some tasks from the learning schedule. Then, EREIL suggests the interested students and closest to the student.

8) Furthermore, EREIL can suggest some activities and present some external learning sources as, books, video, virtual labs, and serious games.

Note: In step7 and step8, EREIL begins an active dialog with students, detects their emotions, and chooses the appropriate activities and tasks.

9) During the previous steps, EREIL can receive direct messages through the direct sender subunit and send responses to students.

10) If EREIL finds a student alone or a student refuses the suggested activity, EREIL can present joyful messages as an attempt to engage him in the learning environment again. Besides, EREIL finds a way to personalize interactions with students. EREIL can stores valuable information gathered during their previous interactions, such as their favorite learning activities, preferred external sources, closest students, and academic learning progress.

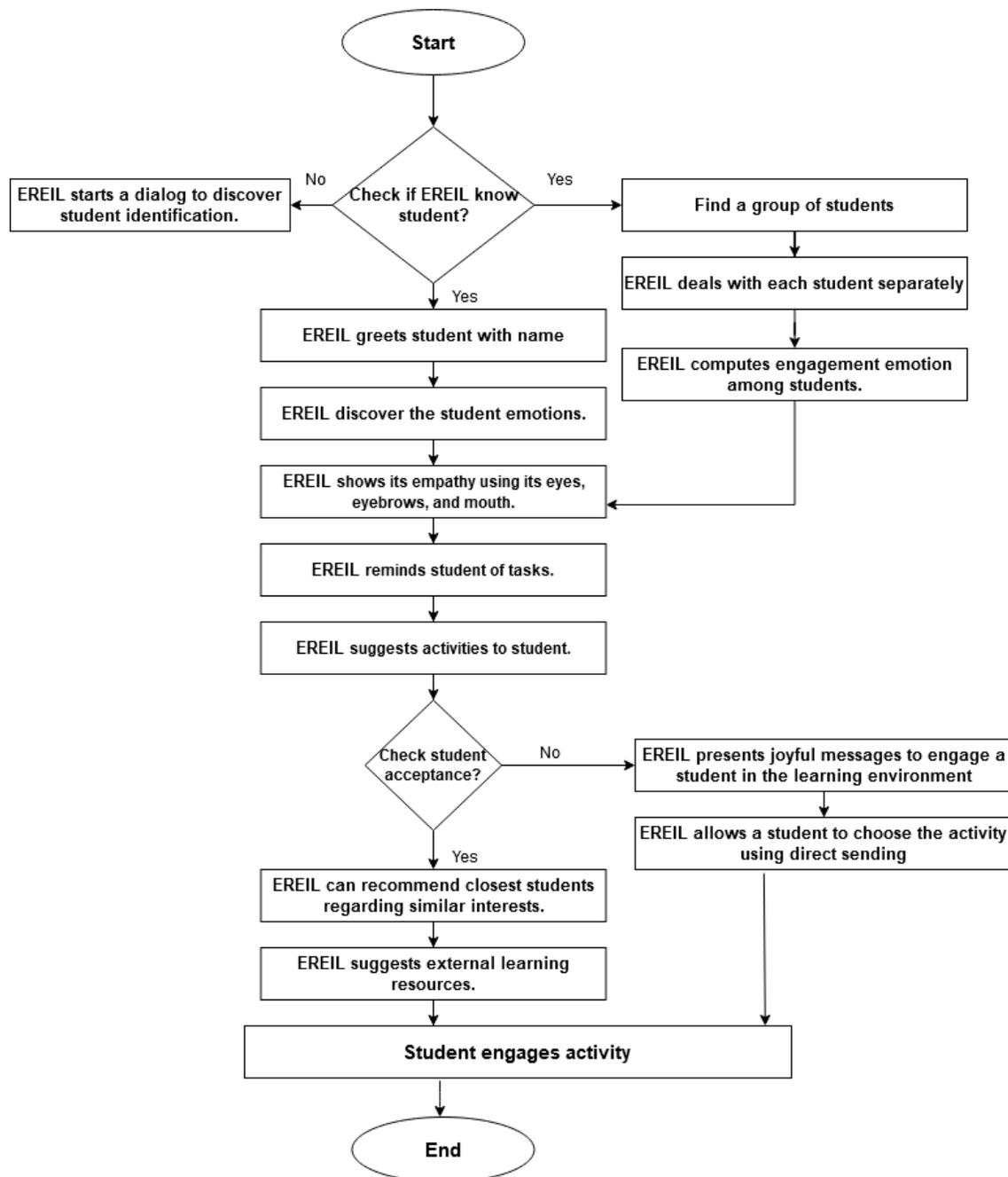


Fig. 8. Flowchart of an Example of EREIL Working.

VIII. CONCLUSION AND FUTURE WORK

This paper discussed educational robotics with emotional intelligence as an assist to motivate interaction in the learning environment. These robots can overcome the special educational needs of students such as autism spectrum disorders. Also, emotional intelligence robots can help students who have a loss in social communication and perception, mental impairment of spoken and non-spoken communication, presence of personal isolated interests, and students who reach the point of social alienation. It presented a proposed framework EREIL, as a student supporter for social interaction in the learning environment. Furthermore, it

presented the main requirements to build emotional intelligence robots as requirements of needs determination, designing, and technology. Besides, it presents proposed EREIL three main units: student emotions discovery, student emotions representation, and EREIL-Student Communication (RSC). The first two units are responsible for discovering, implementing, and processing, and representing emotions. The third unit enhances the student practice and supplies a running visible interface to the EREIL robot. Hence, the third unit is divided into four subunits: student identification, tasks reminder, activities suggestion, and direct sending. Additionally, it presented an example of EREIL working.

The existing work agrees with the following related works. The author in [35] introduced a robot that can understand and express emotions in voice, gesture, and gait using a controller trained only on voice. Besides, [5] proposed a novel algorithm for emotional intelligence robots to recognize human emotions from daily-life gestures. They used machine-learning techniques to automatically infer human emotions with high accuracy. Moreover, [8] showed that the importance of emotional intelligence has a significant effect on employee retention and performance, whereas artificial intelligence plays a significant moderating role in employee performance. Additionally, [13] presented human-robot interaction and how to recognize human emotional states. Also, it explained emotion levels from users are detected through vision and speech inputs.

As future work, we hope to execute a practical case study of the EREIL robot and enhance discovering emotions. It can be implemented by adding more sensor devices and machine learning algorithms to integrate face analysis with speech recognition. Another aspect that can be presented is to store the expert teachers' opinions of students for tracking their instructions in learning activities and their opinions in students' learning progress. Besides, it can add a persuasion unit in the EREIL robot to convince students with better learning choices to their abilities. Additionally, the persuasion unit can provide students with suggestions related to their learning.

REFERENCES

- [1] Khairy, D., Abougalala, R. A., Areed, M. F., & Atawy, S. M. (2020). Educational Robotics Based On Artificial Intelligence And Context-Awareness Technology: A Framework. *Journal of Theoretical and Applied Information Technology*, 98(13), pp. 2227- 2239.
- [2] Amo, D., Fox, P., Fonseca, D., & Poyatos, C. (2021). Systematic Review on Which Analytics and Learning Methodologies Are Applied in Primary and Secondary Education in the Learning of Robotics Sensors. *Sensors*, 21(1), 153.
- [3] Misirli, A., & Komis, V. (2014). Robotics and programming concepts in Early Childhood Education: A conceptual framework for designing educational scenarios. In *Research on e-Learning and ICT in Education* (pp. 99-118). Springer, New York, NY.
- [4] Alimisis, D. (2013). Educational robotics: Open questions and new challenges. *Themes in Science and Technology Education*, 6(1), 63-71.
- [5] Loghmani, M. R., Rovetta, S., & Venture, G. (2017, May). Emotional intelligence in robots: Recognizing human emotions from daily-life gestures. In *2017 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1677-1684). IEEE.
- [6] Krägeloh, C. U., Bharatharaj, J., Sasthan Kutty, S. K., Nirmala, P. R., & Huang, L. (2019). Questionnaires to measure acceptability of social robots: a critical review. *Robotics*, 8(4), 88.
- [7] Hughes-Roberts, T., Brown, D., Standen, P., Desideri, L., Negrini, M., Rouame, A., ... & Hasson, C. (2019). Examining engagement and achievement in learners with individual needs through robotic-based teaching sessions. *British Journal of Educational Technology*, 50(5), 2736-2750.
- [8] Prentice, C., Dominique Lopes, S., & Wang, X. (2020). Emotional intelligence or artificial intelligence—an employee perspective. *Journal of Hospitality Marketing & Management*, 29(4), 377-403.
- [9] Ornaghi, V., Brockmeier, J., & Grazzani, I. (2014). Enhancing social cognition by training children in emotion understanding: A primary school study. *Journal of Experimental Child Psychology*, 119, 26-39.
- [10] Damiano, L., Dumouchel, P., & Lehmann, H. (2015). Artificial empathy: An interdisciplinary investigation.
- [11] Ziouzos, D., Ioannou, M., Ioanna, T., Bratitsis, T., & Dasygenis, M. (2020). Emotional Intelligence and Educational Robotics: The Development of the EI-EDUROBOT. *European Journal of Engineering Research and Science*.
- [12] Kanda, T., Shimada, M., & Koizumi, S. (2012, March). Children learning with a social robot. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 351-358). IEEE.
- [13] Erol, B. A., Majumdar, A., Benavidez, P., Rad, P., Choo, K. K. R., & Jamshidi, M. (2019). Toward artificial emotional intelligence for cooperative social human-machine interaction. *IEEE Transactions on Computational Social Systems*, 7(1), 234-246.
- [14] Fridin, M., & Yaakobi, Y. (2011, May). Educational robot for children with ADHD/ADD. In *Architectural Design, In. Conf on Computational Vision and Robotics*.
- [15] Caci, B., D'Amico, A., & Chiazzese, G. (2013). Robotics and virtual worlds: An experiential learning lab. In *Biologically Inspired Cognitive Architectures 2012* (pp. 83-87). Springer, Berlin, Heidelberg.
- [16] Hudson, S. (2020). Artificial intelligence, Cognitive Robotics and Human Psychology. DO-10.13140/RG. 2.2. 20153.52323.
- [17] Vlachos, E., & Schärfe, H. (2012, October). Android emotions revealed. In *International Conference on Social Robotics* (pp. 56-65). Springer, Berlin, Heidelberg.
- [18] Siciliano, B., & Khatib, O. (Eds.). (2016). *Springer handbook of robotics*. Springer.
- [19] Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education*, 58(3), 978-988.
- [20] Salminen, M., Hamari, J., & Ravaja, N. (2021). Empathizing with the End User: Effect of Empathy and Emotional Intelligence on Ideation. *Creativity Research Journal*, 1-11.
- [21] Krakovsky, M. (2018). Artificial (emotional) intelligence. *Communications of the ACM*, 61(4), 18-19.
- [22] Kim, Y., & Baylor, A. L. (2016). Based design of pedagogical agent roles: A review, progress, and recommendations. *International Journal of Artificial Intelligence in Education*, 26(1), 160-169.
- [23] Yonck, R. (2020). *Heart of the machine: Our future in a world of artificial emotional intelligence*. Arcade.
- [24] Jeon, M. (2017). Emotions and affect in human factors and human-computer interaction: Taxonomy, theories, approaches, and methods. In *Emotions and affect in human factors and human-computer interaction* (pp. 3-26). Academic Press.
- [25] Veletsianos, G., & Russell, G. S. (2014). Pedagogical agents. In *Handbook of research on educational communications and technology* (pp. 759-769). Springer, New York, NY.
- [26] Fan, L., Scheutz, M., Lohani, M., McCoy, M., & Stokes, C. (2017, August). Do we need emotionally intelligent artificial agents? First results of human perceptions of emotional intelligence in humans compared to robots. In *International Conference on Intelligent Virtual Agents* (pp. 129-141). Springer, Cham.
- [27] Chita-Tegmark, M., Lohani, M., & Scheutz, M. (2019, March). Gender effects in perceptions of robots and humans with varying emotional intelligence. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 230-238). IEEE.
- [28] Tay, B., Jung, Y., & Park, T. (2014). When stereotypes meet robots: the double-edge sword of robot gender and personality in human-robot interaction. *Computers in Human Behavior*, 38, 75-84.
- [29] Devillers, L. (2020). Social and emotional robots: useful artificial intelligence in the absence of consciousness. In *Healthcare and Artificial Intelligence* (pp. 261-267). Springer, Cham.
- [30] Lohani, M., Stokes, C. K., Oden, K. B., Frazier, S. J., Landers, K. J., Craven, P. L., ... & Macannuco, D. J. (2017, March). The impact of non-technical skills on trust and stress. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 191-192).

- [31] Sapounidis, T., & Alimisis, D. (2020). Educational Robotics for STEM: A Review of Technologies and Some Educational Considerations. Science and Mathematics Education for 21st Century Citizens: Challenges and Ways Forward; Nova Science Publishers: Hauppauge, NY, USA, 167-190.
- [32] Sapounidis, T., Stamelos, I., & Demetriadis, S. (2016). Tangible user interfaces for programming and education: A new field for innovation and entrepreneurship. *Advances in Digital Education and Lifelong Learning*, 2, 271-295.
- [33] Sapounidis, T., & Demetriadis, S. (2016, November). Educational robots driven by tangible programming languages: A review on the field. In *International Conference EduRobotics 2016* (pp. 205-214). Springer, Cham.
- [34] D'Amico, A., Guastella, D., & Chella, A. (2020). A Playful Experiential Learning System With Educational Robotics. *Frontiers in Robotics and AI*, 7, 33.
- [35] Lim, A., & Okuno, H. G. (2014). The mei robot: towards using motherese to develop multimodal emotional intelligence. *IEEE Transactions on Autonomous Mental Development*, 6(2), 126-138.

AUTHORS' PROFILE

Mohamed A. Amasha, was born in Damietta Egypt, in 1970. He is professor in the application of computer. He worked as a vice dean the faculty of computer and information Damietta university. He is also a manager of E-portal at Damietta University. He works as a lecturer at Damietta University, Faculty of specific education, Computers Dept., With over 11 years of experience in Qassim university KSA, He has published in A.U.C in Egypt, IJACSA, IJORCS, Springer, IEEE.

Marwa F. Areed, is now a lecturer at Damietta University, Faculty of Engineering, Computers & Systems Dept., with over 13 years of experience in Mansoura University Faculty of Engineering from 2000 till 2017 and over three years in Taibah University, Computer Science Collage, Medina, KSA. She received M.Sc. (2003) and Ph.D. (2009) in Computer Engineering, both from Faculty of Engineering, Mansoura University. Her scientific interest includes Data Security in clouds, image compression, digital watermarking & steganography, Cloud Computing, Clustering Algorithms, Sentiment analysis and wireless networks. She has published in INFO2008 in Cairo, (International journal of computer Application 2012, international Journal of Engineering & technology 2014, Simulation Modeling Practice &Theory 2018 and finally in Journal of the International Measurement Confederation (IMEKO).

Dalia Khairy was born in Egypt, Damietta, in 1985. She is a lecturer in the application of computer in education. He works as lecturer, in Damietta University, Egypt.

Raian A. Abougalala, was born in Egypt, Damietta, in 1982. She is a lecture in the application of computer in education. She works as lecture, in Damietta University, Egypt. She has published in Springer.

Salem Alkhalaf , is an Associate Professor in Computer department, College of Science and Arts in Al-Rass, Qassim University. He graduated Bachelor of Education degree in Computer Since from the Department of Computer, Teachers College (Riyadh) in 2003, Also he graduated with Honors degree. And he graduated the Master of ICT from Griffith University (Brisbane, Australia) in 2008. And, he graduated the PhD in.

Building Research Productivity Framework in Higher Education Institution

Ahmad Sanmorino¹

Faculty of Engineering
Universitas Sriwijaya, Palembang, Indonesia

Ermatita^{2*}, Samsuryadi³, Dian Palupi Rini⁴

Faculty of Computer Science
Universitas Sriwijaya, Palembang, Indonesia

Abstract—The purpose of this study is to build a framework for improving research productivity in higher education institutions. The research begins by collecting data and defining candidate variables. The next process is to determine the selected variable from the candidate variable. Variable selection is carried out in three stages, univariate selection, feature importance, and correlation matrix. After the variable selection stage, eight input variables and one target variable were obtained. The eight input variables are Article (C), Conference (CO), Grant (GT), Research Grantee (RG), Rank (R), Degree (D), IPR, and Citation (C). The target variable is Research Productivity (RP). This selected variable is used to build the framework. The next step is to test the framework that has been built. The testing process involves four data mining classifiers, Support Vector Machine, Decision Tree, K-Nearest Neighbor, and Naïve Bayes. The classification results are tested using confusion matrix-based testing, accuracy, precision, sensitivity, and f-measure. The testing results show the proposed framework is able to obtain high accuracy scores for each classification algorithm. It means the proposed framework is relevant to use.

Keywords—Framework; research productivity; variable selection; data mining classifier

I. INTRODUCTION

Lecturers are the main research actors in a higher education institution. Lecturers are required to conduct research which is one of the three main functions, besides teaching and serving the community. The research achievement target is in accordance with the research scheme chosen by the lecturer. Research results are the targets achieved by researchers from a research activity at the end of the period. Research does not only talk about the quantity of research productivity but also shows the quality of research in a higher education institution [1]. Therefore, the increase in research productivity, both quantity and quality must be measured, in order to know the extent of research progress in a higher education institution [2].

The increasing research productivity is strongly influenced by the environment and the involvement of stakeholders who have an interest in research [3]. This involvement is better known as collaboration. Research collaborations are carried out between one researcher or a group of researchers with other researchers. Each researcher comes from the same or different disciplines, or even different universities [4]. On a wide scale, research collaboration happens between countries, because distance is not a problem now [5]. In recent years, data mining-based knowledge management has been used as

the best approach to achieve the goals of an organization with a focus on knowledge creation [6]. One mechanism to increase research productivity is to use a knowledge-sharing approach that involves the role of academics in higher education [7]. The results of this study indicate that the involvement of academics in higher education in research productivity has a variance of 22.6 percent. This shows that the character of academics such as education degree, academic rank, and experience has a considerable influence on research productivity.

In higher education institutions, the data mining approach is the right solution for the analysis of very large research data. Through a data mining approach, researchers know which variables are significant in research productivity. These variables are then used as constructs to build a mechanism for increasing research productivity. The mechanism for increasing research productivity is formulated in the form of a model or framework. The framework development process starts from the preprocessing stage, by selecting the variables to be used. The role of the data mining approach in this case is as a tool for analysis or testing of the framework that has been built. Tests are carried out to determine the performance of the proposed framework. The analysis and testing process involves several data mining algorithms. Furthermore, a comparison of the test results using several data mining algorithms is carried out in order to obtain the best results. The results of this test also show the framework's performance from various points of view, because each data mining algorithm used in testing has different characteristics and approaches. Next is a discussion on research related to research productivity in higher education institutions, followed by an explanation of materials and methods, then results - discussion, and conclusion.

II. RELATED WORK

The framework is defined as mutually supporting parts to achieve a goal. The framework is analogous to a skeleton in the human body that is interconnected, mutually supportive, influencing one another. The framework has a clear direction of achievement, usually illustrated by an arrow to a point. Many researchers have developed frameworks for various needs. In the previous study, researchers built a research productivity framework by combining knowledge sharing and gamification-based variables [8][9]. Another example of developing a framework using knowledge sharing in a higher education environment has been carried out by some researchers [7]. Research productivity is used to determine the

*Corresponding Author

position of higher education institutions on a national and international scale. A mechanism is needed to optimize research productivity. Sample data were taken from tutors to professors in Malaysia with a ratio of 50:30:20 for senior lecturers: assoc. professor: professor.

Through the proposed knowledge management framework (KS), researchers have succeeded in proving that the role of academics, which 12 constructs, has a positive effect on research productivity. The 12 constructs used are commitment, social network, management support, social media, attitude, subjective norm, intention, and behavior, perceived behavior control, facilitating conditions, trust, and research productivity. The results showed that academic productivity has a variance of 22.6 percent. This suggests the academic behavior of KS has a large impact on research productivity. The academic attitude, academic commitment, trust, and social network explain the variance of 36.4 percent. Management support has a variance of 38.7 percent for subjective norms while facilitating conditions and social media have a variance of 26.5 percent for perceived behavioral control. Academics KS intention and KS behavior explain the variance of 62.1 and 47.1 percent, respectively.

The framework is composed of variables that are related to each other. The variable selection process starts with the selection of features from the dataset that has been collected. There are several studies and publications related to research productivity (Table I).

TABLE I. RELATED STUDY WITH RESEARCH PRODUCTIVITY

Author	Variable Selection Mechanism	Algorithm
Henry <i>et al.</i> [10]	Chi-Square, Nagelkerke R Square	Logistic Regression
Ramli <i>et al.</i> [11]	Not mentioned	Logistic Regression, Decision Tree, Artificial Neural Network, SVM
Nazri <i>et al.</i> [12]	Spearman Rho Correlation	Decision Tree, PART, J-48, C4.5
Wichian <i>et al.</i> [13]	Chi-Square, Cronbach Alpha, R-Square	Neural Network Analysis (Back Propagation)
Sanmorino <i>et al.</i>	Chi-Square, Extra Tree, Pearson Correlation Co.	SVM, Decision Tree, K-NN, Naïve Bayes

There are several studies related to optimization of research productivity [14]. One of them discusses the gap in the number of professors against other academics, students, or faculty members. In other words, students and faculty members need to be involved in research. The proposed model increases research productivity in higher education institutions. The idea of this model is to involve students and faculty members in intensive research through a curriculum design that focuses on research, which enables students and faculty members to participate in research projects sponsored by the industrial world.

Apart from research, the performance of a lecturer is measured based on the quality of teaching and service to the community. Research related to teacher performance has been conducted [15]. Through this research, several factors associated with teacher performance were tested. The factors that influence teacher performance are currently unclear, so

testing is needed to determine these factors. After the various factors are known, they are used to improve the quality of teacher performance in schools.

Researchers propose data mining-based classification and association models, such as decision trees, rule induction, K-NN, and Naïve Bayes to evaluate teacher performance in providing educational services in schools. Some of the attributes used in the test are teacher name, course, class, workspace, training, number of training, and several questions related to teacher performance in schools. The next step is the measurement of accuracy for the data mining method used. In addition to the use of data mining as previously stated, a data mining classifier is also used for various problem solutions such as performance prediction [16][17], performance improvement [18], or decision support system analysis which has been carried out by several researchers [19][20].

III. MATERIAL AND METHOD

The dataset in this research has been collected by the Ministry of Research and Technology Republic of Indonesia through the Science and Technology Index (SINTA) platform. SINTA was launched and has been actively used by academics since 2017. SINTA provides access to citations and scientific expertise in Indonesia. On its official website, SINTA is referred to as an information system used to measure the performance of researchers, including lecturers, and scientific journals in Indonesia. Apart from that, SINTA is a web based platform which is very easy to use. Another reason is because SINTA as an online database accommodates research data from lecturers from all over Indonesia, which is needed to carry out this research. The SINTA platform is equipped with a rating system for researchers and journals in Indonesia [21].

The framework testing process will use a data mining approach. In this study, data mining algorithms were used to measure the performance of the conceptual framework. Another goal is to find patterns and relationships between variables in the dataset. To accommodate the testing stage, this study applied the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology [22]. There are six stages in CRISP-DM [23], shown in Fig. 1.

A. Business Understanding

Business understanding is the first stage in CRISP-DM. At this stage, knowledge of business objects is required, an understanding of the scope of the problem, and how to obtain data. Activities undertaken in the business understanding stage include: (1) clearly defining goals and specifications, (2) translate goals and specifications, and (3) determine the boundaries of data mining problems. The next step is to prepare an initial strategy to achieve the goals.

B. Data Understanding

The data understanding stage begins with data collection, identifying data types, qualitative or quantitative, and measurement levels such as nominal, ordinal, binary, and interval [24]. At this stage an understanding of the dataset is needed, to determine properties such as variables or attributes used in modeling.

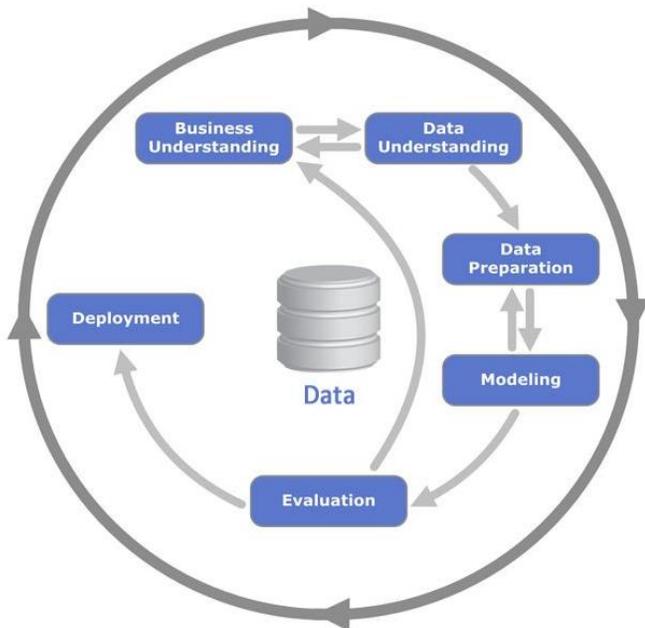


Fig. 1. CRISP-DM Methodology.

C. Data Preparation and Modeling

This stage begins with the identification of the variables used to build the framework. This process focuses on identifying significant variables toward the target variable and removing irrelevant or less important variables from the dataset. Irrelevant variables have a negative impact on the overall model performance. The details of the data preparation and modeling stages are shown in Fig. 2.

Variable selection is one of the core concepts which greatly affect the performance of the data mining model. Some of the advantages obtained by doing variable selection are: (a) reducing overfitting, (b) reducing training time, and most importantly, (c) increasing accuracy. There are three stages of variable selection carried out in this study: (a) univariate selection, (b) feature importance, and (c) correlation matrix.

In the univariate selection stage, the Chi-Square statistical test is used. Chi-Square is used to test the relationship between two variables. In other words, Chi-Square is used to measure how strong the relationship between variables [25][26]. In this study, the relationship tested is between the input variables and the target variables. Variables with a significant relationship value are used for the constructs of the framework. The character of Chi-Square always has a positive value. The formula for Chi-Square is:

$$X_c^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

Where, c = degrees of freedom, O = observed value(s), and E = expected value(s). If data from two variables are given, the observed number (O) and the expected number (E) are obtained. Chi-Square measures the deviation between the expected number E and the observed number O .

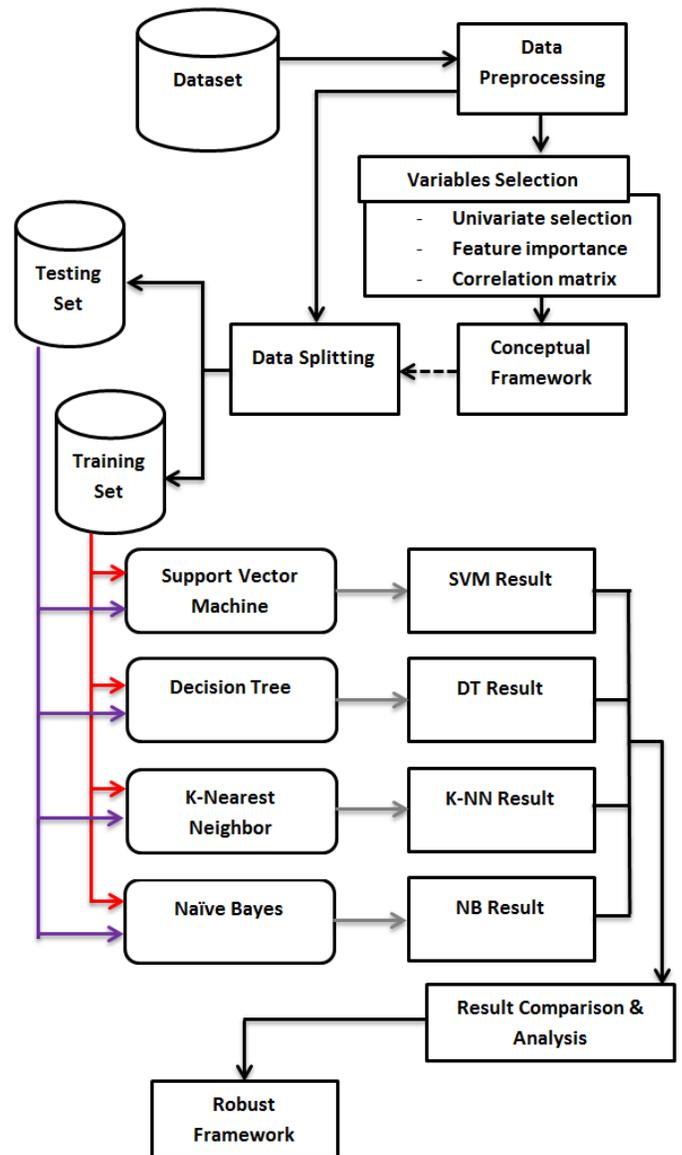


Fig. 2. Data Preprocessing and Modeling.

After the univariate selection stage, it is followed by the feature importance stage. Feature importance is similar to information gain, which extracts the information level (weight) of a feature or variable [27][28]. The results of the selection using feature importance show the score for each variable. The higher the score of a variable, the more relevant or important it is to the target variable. Feature importance uses a Tree-based classifier. In this study, the Extra Tree Classifier used to extract the important variables from the prepared dataset [29]. The correlation matrix shows the correlation between input variables with other input variables or input variables to the target variable. Correlation can be positive if an increase in the Input variable has an impact on an increase in the target variable, or conversely, an increase in the input variable decreases the target variable. Unlike univariate selection (Chi-Square), the correlation matrix can be negative. The correlation matrix test is usually visualized with a heat map. The heat map shows the variables most

related to the target variable and vice versa. After obtaining the relevant variables, the next steps are broken down into two stages: (a) building a conceptual framework, (b) dividing the sample. The sample will be divided into two parts, with a ratio of 70:30, 70 percent for training, and 30 percent for validation. Training and testing data are used as modeling input. This modeling stage is a test for the conceptual framework. In this testing phase, four data mining algorithms are used, Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbor, and Naïve Bayes. This testing phase shows the framework's performance from various points of view because each data mining algorithm used in testing has different characteristics and approaches. The next step is to compare the test results to get the best results.

D. Evaluation and Deployment

The confusion matrix is used to determine the best model. By looking at the confusion matrix value, the accuracy of each model is known. Classification is included in supervised learning, which is a predictive model where the prediction results are discrete. The way to measure the performance of the classification model is to compare the actual value with the predicted value. The confusion matrix is a performance measurement for machine learning classification problems, where the output is two or more classes [30]. The Confusion Matrix is a table with four different combinations of predicted and actual values [31]. There are four terms that represent the results of the classification process in the confusion matrix, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Based on the Confusion Matrix, the formula for accuracy is obtained:

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (2)$$

Accuracy shows how accurate the model is in classifying correctly.

$$Precision = \frac{(TP)}{(TP+FP)} \quad (3)$$

Precision shows the accuracy between the actual data and the prediction results displayed by the model.

$$Sensitivity = \frac{TP}{(TP+FN)} \quad (4)$$

Recall or sensitivity shows the success of the model in retrieving information.

$$F - Measure = \frac{(2*Recall*Precision)}{(Recall+Precision)} \quad (5)$$

F-Measure (f1-score) shows the weighted average comparison of precision and recall [32]. Accuracy is appropriate to use as a reference for the performance of the classification method if the dataset has a very symmetric amount of FN and FP data. However, if the numbers are not symmetric, it is suggested to use the F-Measure as a reference.

IV. RESULT AND DISCUSSION

At an early stage, candidates for the variables are defined as shown in Table II.

The next step is an analysis of the candidate variables. The analysis focuses on the relevance and ease of obtaining data

for each variable. Based on the analysis, there are several variables that cannot be used: (1) Working hours, the obstacles faced are difficulties in getting information about working hours, (2) Marital status, because this information is personal, so researcher prefers not to use it, (3) SINTA's score, is optional because the calculation of the score comes from the number of articles and the number of citations, which the variables have been determined, (4) Experience, there is no valid data yet for research experience. After defining the variables, the next step is variable selection.

TABLE II. CANDIDATE VARIABLE

Variable Name	Variable Description	Measurement Level
Degree (D)	Lecturer education degree	Nominal (Master, Doctor)
Gender (G)	Lecturer's gender	Binary(Male, Female)
Working hours (WH)	Type of working hours	Ordinal (Part time, Full Time)
Nationality (N)	Nationality	Nominal (Indonesia, Non Indonesia)
Rank (R)	Lecturer's rank	Ordinal (Lecturer, Assist Prof, Assoc Prof, Full Prof)
Marital Status (MS)	Lecturer's marital status	Nominal (Single, Married, Widowed)
Conference (CO)	The total number of attended conferences	Ordinal (Never, Ever, Often)
Article (A)	The total number of published articles on Scopus	Ordinal (None, Very Few, Few, Enough, Much)
Citation (C)	The total number of citations for the published articles on Scopus	Ordinal (None, Few, Many, Very Much)
Intellectual Property Rights (IPR)	The total number of IPR registered	Ordinal (None, Few, Many)
Experience (E)	Research experience	Ordinal (Inexperienced, Short Time, Long Enough, Very Experienced)
Research grantee (RG)	Lecturers who receive research grants	Ordinal (Yes, No)
Grant (GT)	The total number of grants obtained	Ordinal (None, Few, Many, Very Much)
SINTA's score (SS)	Lecturer's SINTA score	Ordinal (Low, Medium, High, Very High).
Research Productivity (RP)	Target variable	Binary (Fulfilled, Not Fulfilled)

A. Univariate Selection

Univariate selection is used to select the variable with the strongest relationship toward the target variable. Chi-Square statistical testing shows the results of the selection in order, as shown in Table III.

The test results show that Article (A) is in the first rank. This shows Article (A) has the strongest relationship toward the target variable, followed by Citation (C), Conference (CO), Grant (GT), and others. The results of this test also show the number of articles and the number of citations,

which play an important role in measuring the research performance of a lecturer. Then, for the two lowest ranks, it turns out that Gender (G) has the weakest relationship toward the target variable. In other words, Gender (G) does not have a significant effect on research productivity. Nationality (N) is in the lowest rank, because all lecturers are from Indonesia. This variable does not make a significant difference to the target variable. Variables with a score below 1 are not used in building the proposed framework, so only eight input variables and one target variable remain.

TABLE III. UNIVARIATE SELECTION

No	Variable Name	Chi-Square Score
1	Article (A)	76.533603
2	Citation (C)	47.256279
3	Conference (CO)	45.680553
4	Grant (GT)	22.205091
5	IPR	5.002761
6	Rank (R)	4.027538
7	Research grantee (RG)	2.538671
8	Degree (D)	1.129551
9	Gender (G)	0.118249
10	Nationality (N)	0.000000

B. Feature Importance

Through the feature importance stage, it is possible to know the importance of each variable. The higher the score of a variable, the more relevant or important it is to the target variable. The results of feature importance using the Extra Tree Classifier are shown in Fig. 3.

The selection of feature importance shows Article (A) is in the first rank, with the value 0.3447, followed by Conference (CO) and Citation (C). This measurement shows Article (A) is the variable most relevant to the target variable. Overall, the test results using feature importance are not different from the univariate selection, where the two lowest ranks are Research Grantee (RG) and Nationality (N). These two variables are the least relevant to the target variable. There is a difference in the bottom two variables between univariate selection and feature importance. As a solution, the third step was carried out, the correlation matrix.

C. Correlation Matrix

Correlation can be positive if an increase in the Input variable has an impact on an increase in the target variable, or conversely, an increase in the input variable decreases the target variable. The correlation matrix results using heat maps are shown in Fig. 4 and Fig. 5. Heat maps showing the correlation between input variables with other input variables or input variables for the target variable. Like the two previous

steps, Nationality (N) and Gender (G) have poor correlation with other variables. Even Nationality does not have a correlation (zero correlation) with other variables. Gender (G) still has a correlation. Although the correlation to the target variable is the lowest when compared to others. The correlation of Gender (G) to the target variable is 0.046. Even Gender (G) has a negative correlation with Article (A), Research Grantee (RG), and Grant (GT). For other variables, the correlation to the target variable is still > 0.2 (Table IV).

Article (A) has the highest correlation to the target variable, 0.77, followed by Citation (C) of 0.67. The average score of Article (A) on other input variables is very high, thus increasing its correlation to the target variable. The significant difference compared to the previous stage is that Degree (D) and IPR have a low correlation score. This happens because the correlation of Degree (D) and IPR for other input variables is very low so that it affects their correlation to the target variable. However, it is still fair to use as a construct for the proposed framework.

After getting the input and target variables, the next step are to build the framework. Fig. 6 shows the conceptual framework.

Framework consists of eight input variables and one target variable. The eight input variables are Article (C), Conference (CO), Grant (GT), Research Grantee (RG), Rank (R), Degree (D), IPR, and Citation (C). The target variable is Research Productivity (RP). The framework that has been built must be tested first. The data mining approach was chosen as a testing tool because it is in accordance with the characteristics of the dataset that has been prepared. In this testing phase, four data mining algorithms are used, Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (K-NN), and Naive Bayes (NB). The classification results using data mining algorithms tested using confusion matrix-based measurement. The test results using the confusion matrix-based measurement (accuracy, precision, sensitivity, f1-measure) are shown in Table V.

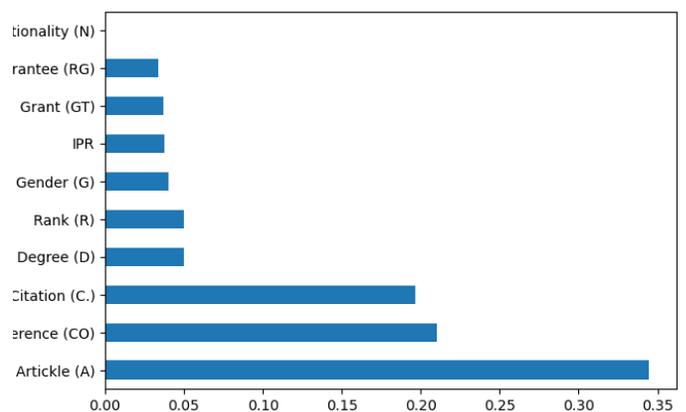


Fig. 3. Feature Importance.



Fig. 4. Correlation Matrix (Include Gender and Nationality).

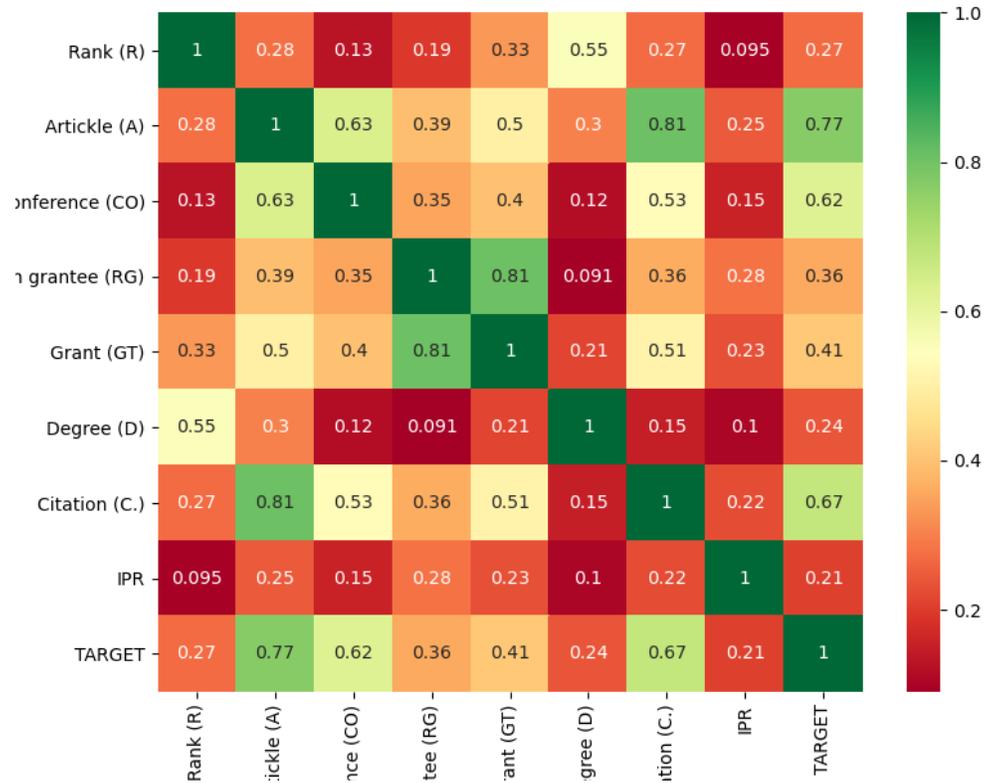


Fig. 5. Correlation Matrix (Exclude Gender and Nationality).

TABLE IV. THE CORRELATION OF INPUT VARIABLES TOWARD TARGET VARIABLE

No	Variable Name	Correlation Score
1	Article (A)	0.77
2	Citation (C)	0.67
3	Conference (CO)	0.62
4	Grant (GT)	0.41
5	Research grantee (RG)	0.36
6	Rank (R)	0.27
7	Degree (D)	0.24
8	IPR	0.21

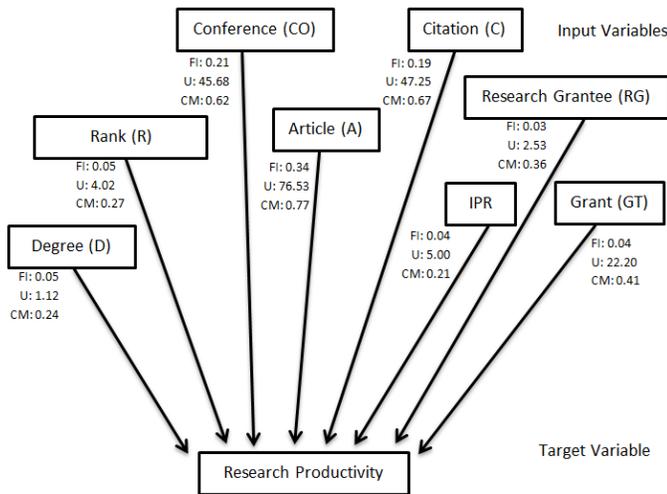


Fig. 6. Conceptual Framework.

TABLE V. ACCURACY, PRECISION, SENSITIVITY, F-MEASURE AND MISCLASSIFICATION RATE FOR 4 ALGORITHMS

Classifier	Accuracy	Precision	Sensitivity	f1-Score	Misclass. rate
SVM	78.26%	85.29%	77.27%	76.67%	21.74%
DT	86.95%	87.12%	87.12%	86.95%	13.05%
K-NN	95.65%	96.15%	95.45%	95.61%	4.35%
NB	86.95%	90.00%	86.36%	86.54%	13.05%

The testing result shows the proposed framework is able to obtain high accuracy scores for each classification algorithm. The highest accuracy score on the K-NN classification algorithm is 95.65 percent, followed by Decision Tree and Naïve Bayes, each with 86.95 percent; the last is Support Vector Machine at 78.26 per cent. Just like the accuracy score, for the measurement of precision, sensitivity, F-Measure, the K-NN algorithm is also the highest, with the lowest misclassification rate, only 4.35 percent. The results of confusion matrix-based testing prove that the proposed framework is relevant to use, with high accuracy scores and little misclassification rate. When compared with the results of other related research tests regarding research productivity, the position of the results of this test is (Table VI).

TABLE VI. THE COMPARISON OF VARIABLES, ALGORITHM, AND ACCURACY OF CLASSIFICATION TESTING RESULTS

Author Name	Variables Used	Algorithm Used	Accuracy
Henry <i>et al.</i> [10]	Age Cohort, Highest Qualification, Cluster, Lecturer Track, Achievement, Job Policy, Monthly Income, Research Leadership, Research Supervision	Logistic Regression	78.2%
Ramli <i>et al.</i> [11]	Age, Gender, Marital Status, Qualification, Experience, Position, Division, Citation, Article, Conference, and Target (Status of Research Performance)	Logistic Regression	80.31%
		Decision Tree	83.40%
		Artificial Neural Network	82.24%
		Support Vector Machine	80.47%
Nazri <i>et al.</i> [12]	Age, Designation, No. Research Grant, Gender, Performance Score, Marital Status, Working Status, Amount of Grant, Department, Administrative Post, No. PhD Student, Faculty, Invitation as Keynote Speaker, Article (Index)	Decision Tree	70.30%
		PART	75.00%
		J-48	75.30%
		C4.5	70.20%
Wichian <i>et al.</i> [13]	Age, Academic Position, Thinking, Research Mind, Volition - Control, Meeting of International, Research Skill - Techniques, Research Fund, Research Management, Communication, Networking and Teamwork, Institutional Policy, Library Expenditure, Computing Facility	Neural Network (Back Propagation)	90.72%
Sanmorino <i>et al.</i> (this study)	Article, Conference, Grant, Research Grantee, Rank, Degree, IPR, Citation, and Target (Research Productivity)	Support Vector Machine	78.26%
		Decision Tree	86.95%
		K-Nearest Neighbors	95.65%
		Naïve Bayes	86.95%

There are differences in the combination of algorithms, variables and the number of datasets used that affect the performance of the classification algorithm, but this study has proven that the framework designed based on the variable selection has a relevant good accuracy score. Researchers cannot say that the results of this test are better than other related studies. To prove the test results of a study are better than other studies, the same scenario must be used, in the sense of where to collect the data, the number of datasets, the mechanism for selecting variables, the number of variables must all be the same, because they can affect the test results.

V. CONCLUSION

The framework development process starts from collecting datasets and determining candidate variables. The next process is to determine the selected variable from the candidate

variable. Variable selection is carried out in three stages, univariate selection, feature importance, and correlation matrix. After the variable selection stage, eight input variables and one target variable were obtained. The eight input variables are Article (C), Conference (CO), Grant (GT), Research Grantee (RG), Rank (R), Degree (D), IPR, and Citation (C). The target variable is Research Productivity (RP). This selected variable is used to build the framework. The next step is to test the framework that has been built. The testing process involves four data mining classifiers. The classification results are tested using confusion matrix-based testing, accuracy, precision, sensitivity, and f1-measure. The testing results show the proposed framework is able to obtain high accuracy scores for each classification algorithm. It means the proposed framework is relevant to use. There are several things recommended for future work, such as increasing the number of datasets, using other variables relevant to research productivity, such as research collaboration, teamwork, or research facilities in a higher education institution.

ACKNOWLEDGMENT

The first author is a doctoral student at the Faculty of Engineering, Universitas Sriwijaya. The authors would like to thank Universitas Sriwijaya for their support in carrying out this research.

REFERENCES

- [1] C. N. Tan, "Improving Research Productivity through Knowledge Sharing: The Perspective of Malaysian Institutions," no. October, pp. 701–712, 2015.
- [2] G. Abramo, C. A. D. Angelo, G. Abramo, C. Andrea, and D'Angelo, "How do you define and measure research productivity? Scientometrics," no. November, 2014, doi: 10.1007/s11192-014-1269-8.
- [3] G. Li et al., "Enhancing research publications and advancing scientific writing in health research collaborations: Sharing lessons learnt from the trenches," *J. Multidiscip. Healthc.*, vol. 11, pp. 245–254, 2018, doi: 10.2147/JMDH.S152681.
- [4] Z. Zuo and K. Zhao, "The more multidisciplinary the better ? – The prevalence and interdisciplinarity of research collaborations in multidisciplinary institutions," *J. Informetr.*, vol. 12, no. 3, pp. 736–756, 2018, doi: 10.1016/j.joi.2018.06.006.
- [5] D. Press, "How to set-up a long-distance mentoring program : a framework and case description of mentorship in HIV clinical trials," no. January 2013, 2014, doi: 10.2147/JMDH.S39731.
- [6] C. Sassenberg, C. Weber, and M. Fathi, "A Data Mining based Knowledge Management Approach for the Semiconductor Industry," no. July, 2009, doi: 10.1109/EIT.2009.5189587.
- [7] M. A. Fauzi, C. T. Nya-Ling, R. Thursamy, and A. O. Ojo, "Knowledge sharing: Role of academics towards research productivity in higher learning institutions," *VINE J. Inf. Knowl. Manag. Syst.*, vol. 49, no. 1, pp. 136–159, 2019, doi: 10.1108/VJIKMS-09-2018-0074.
- [8] A. Sanmorino, Ermatita, and Samsuryadi, "The preliminary results of the kms model with additional elements of gamification to optimize research output in a higher education institution," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 5, pp. 554–559, 2019.
- [9] A. Sanmorino, Ermatita, Samsuryadi, and D. P. Rini, "A Robust Framework using Gamification to Increase Scientific Publication Productivity," *Proc. - 2nd Int. Conf. Informatics, Multimedia, Cyber, Inf. Syst. ICIMCIS 2020*, pp. 29–33, 2020, doi: 10.1109/ICIMCIS51567.2020.9354319.
- [10] C. Henry, N. A. Md Ghani, U. M. A. Hamid, and A. N. Bakar, "Factors contributing towards research productivity in higher education," *Int. J. Eval. Res. Educ.*, vol. 9, no. 1, pp. 203–211, 2020, doi: 10.11591/ijere.v9i1.20420.
- [11] N. A. Ramli, N. H. M. Nor, and S. S. M. Khairi, "Prediction of research performance by academics in local universities using data mining approach," vol. 040021, no. August, 2019, doi: 10.1063/1.5121100.
- [12] M. Z. A. Nazri, R. A. Ghani, S. Abdullah, M. Ayu, and R. N. Samsiah, "Predicting Academic Publication Performance using Decision Tree .," no. 2, pp. 180–185, 2019.
- [13] S. Na Wichian, S. Wongwanich, and S. Bowarnkitiwong, "Factors affecting research productivity of faculty members in government universities: LISREL and Neural Network analyses," *Kasetsart J. - Soc. Sci.*, vol. 30, no. 1, pp. 67–78, 2009.
- [14] P. S. Aithal, "Study on Research Productivity in World Top Business Schools," *Int. J. Eng. Res. Mod. Educ. ISSN 2455 - 4200*, vol. I, no. I, pp. 629–644, 2016.
- [15] R. K. Hemaidd and A. M. E.- Halees, "Improving Teacher Performance using Data Mining," *Ijarcece*, vol. 4, no. 2, pp. 407–412, 2015, doi: 10.17148/ijarcece.2015.4292.
- [16] V. Vijayalakshmi, K. Panimalar, and S. Janarthanan, "Predicting the performance of instructors using Machine learning algorithms," no. December, 2020.
- [17] Q. A. Al-Radaideh and E. Al Nagi, "Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance," *Int. J. Adv. Comput. Sci. Appl.*, vol. 3, no. 2, p. 8, 2012, [Online]. Available: www.ijacsa.thesai.org 144.
- [18] Q. Zhang, W. Hu, Z. Liu, and J. Tan, "TBM performance prediction with Bayesian optimization and automated machine learning," *Tunn. Undergr. Sp. Technol.*, vol. 103, no. June, p. 103493, 2020, doi: 10.1016/j.tust.2020.103493.
- [19] B. Wah, S. Huat, N. Huselina, and M. Husain, "Expert Systems with Applications Using data mining to improve assessment of credit worthiness via credit scoring models," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13274–13283, 2011, doi: 10.1016/j.eswa.2011.04.147.
- [20] H. Jantan, N. M. Yusoff, and M. R. Noh, "Towards Applying Support Vector Machine Algorithm in Employee Achievement Classification," pp. 12–21, 2014.
- [21] L. Lukman et al., "Case Study Proposal of the S-score for measuring the performance of researchers , institutions , and journals in Indonesia," vol. 5, no. 2, pp. 135–141, 2018.
- [22] C. Industry et al., "Data Science Process," no. 1, pp. 19–37, 2019, doi: 10.1016/B978-0-12-814761-0.00002-2.
- [23] K. Jensen, "IBM SPSS Modeler CRISP-DM Guide," 2016.
- [24] G. M. Robinson, *Statistics, Overview, Second Edition.*, vol. 13. Elsevier, 2020.
- [25] K. Molugaram and G. S. Rao, *Chi-Square Distribution*. 2017.
- [26] F. Girosi and G. King, "Model Selection," *Demogr. Forecast.*, pp. 94–123, 2018, doi: 10.2307/j.ctv301hd6.12.
- [27] I. Kareva and G. Karev, "Replicator dynamics and the principle of minimal information gain," *Model. Evol. Heterog. Popul.*, pp. 129–154, 2020, doi: 10.1016/b978-0-12-814368-1.00008-4.
- [28] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. Bin Bin Idris, A. M. Bamhdi, and R. Budiarto, "CICIDS-2017 Dataset Feature Analysis with Information Gain for Anomaly Detection," *IEEE Access*, vol. 8, pp. 132911–132921, 2020, doi: 10.1109/ACCESS.2020.3009843.
- [29] K. Kaur and S. K. Mittal, "Classification of mammography image with CNN-RNN based semantic features and extra tree classifier approach using LSTM," *Mater. Today Proc.*, no. xxxx, 2020, doi: 10.1016/j.matpr.2020.09.619.
- [30] F. Evaluating, P. To, and P. Model, "Model Evaluation," 2015, doi: 10.1016/B978-0-12-801460-8.00008-2.
- [31] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification : A measure driven view," *Inf. Sci. (Ny)*, no. xxxx, 2019, doi: 10.1016/j.ins.2019.06.064.
- [32] L. Derczynski, "Complementarity , F-score , and NLP Evaluation," pp. 261–266, 2013.

Method for Determination of Support Length of Daubechies Basis Function for Wavelet MRA based Moving Characteristic Estimation

Kohei Arai

Faculty of Science and Engineering
Saga University, Saga City
Japan

Abstract—Method for determination of support length of Daubechies basis function for wavelet Multi Resolution Analysis: MRA based moving characteristic estimation is proposed. The method based on root mean square difference between original and reconstructed image from just a low frequency component from MRA. Also, one of applications of the method for detection and tracking of moving targets, typhoon and boundary between warm and cold current observed from satellite remote sensing images is shown. It is found that the proposed method allows to detect and to track for moving targets of a typhoon and a boundary in the time series of GOES images.

Keywords—Multi-dimensional wavelet transformation; multi resolution analysis: MRA; moving target detection; support length; typhoon movement; boundary detection and tracking

I. INTRODUCTION

Moving characteristics such as velocity and vibration of moving objects are being analyzed by taking advantage of the characteristics of time-frequency analysis by wavelet analysis [1]. Especially, the estimation of the moving vector of the moving object with the multi-resolution analysis based on the Daubechies basis function with the finite length support (compact support), which is widely used for the detection of the moving object because of its short support length Has been proposed [2].

A wavelet is a spatial frequency filter, and although it is desirable that its support has a finite length, in many cases the support length is infinite [3]. The Daubechies basis function is a finite length support, and its support length is short, so it is applied to moving object detection. However, there is no report that mentions how to select the optimum support length suitable for the moving speed.

From the viewpoint of making the complexity of the model that expresses the phenomenon appropriate, studies have been conducted on the optimization of base functions based on the information theory evaluation criterion Minimum Description Length: MDL [4].

It is designed to select the basis function that minimizes the code length for data representation. After that, there was a proposal for optimization based on Akaike's Information Criteria: AIC [5]. This optimization method deals with the entire basis function and is suitable for constructing filters such as noise removal. This paper proposes a method to select the

optimum support length for the moving speed of a moving object. The validity of the proposed method has been confirmed by taking typhoons and tides with different moving speeds estimated from time-series data of geostationary meteorological satellite images as examples.

Methods related to wavelet analysis, in particular, detection of moving objects [6] and estimation of movement parameters [7] based on continuous and discrete wavelet transforms and multiresolution analysis [8] have been proposed [9,10]. Also proposed are a method that considers tracking of a moving object using a Kalman filter [11] and a method that considers the selection of wavelet parameters that match the motion of the analysis target [12]. Method for support length determination of base function of wavelet for edge and line detection as well as moving object and change detections is proposed [12].

On the other hand, method for car in dangerous action detection by means of wavelet Multi-Resolution Analysis based on appropriate support length of base function is proposed [13]. Meanwhile, prediction method of El Nino Southern Oscillation event by means of wavelet-based data compression with appropriate support length of base function is proposed [14]. Bright band height assignment with precipitation radar data based on Multi-Resolution Analysis: MRA of wavelet analysis is also proposed [15].

Applying a basis function to multidimensional time series data and transforming it into the time-frequency domain is called Discrete Wavelet Transformation (DWT), and its inverse is called Inverse Discrete Wavelet Transformation (IDWT). A method for estimating the movement vector of a moving object from time-series image information using the Daubechies basis function (complete orthonormal system) that can be completely restored by these is proposed. At this time, it is considered that the method for determining the support length of the basis function has not been determined.

This depends on the frame rate of the moving image to be acquired and the moving characteristics (change rate) of the moving object. Although the optimal support length can be determined if the movement characteristics are known, it is customary that the movement characteristics to be obtained are not given in advance. Therefore, a method for determining the optimum support length by some method is required.

Here, the author proposes an experimental method for determining the optimal support length. That is, some support length candidates that are considered to be appropriate are selected (for example, 2, 4, 8, 16 etc.), DWT based on the Daubechies basis function of those support lengths is applied to only one stage, and then low frequency is applied. Reconstruction (IDWT) is performed using only the components, and the root mean square error (RMSE) of the reconstructed data and the original data is evaluated, and this is the method that maximizes the maximum support length.

Since the energy (power spectrum) of the analysis target data is invariant, when this RMSE is large, it means that the high frequency components could be extracted efficiently. Therefore, it means that the change (movement) component was effectively extracted by DWT / IDWT. The author shows the validity of this proposed method by introducing some examples.

As an application example of the proposed method, a relatively steeply moving typhoon and a relatively slowly changing tide are taken from Geostationary Operational Environmental Satellite (GOES) visible band images acquired on the hourly basis. Determine the optimal support length of Daubechies basis functions for moving analysis of.

The DWT based on this basis function is applied to the original image to separate the original image into low-frequency and high-frequency components, the low-frequency components are removed and reconstructed to extract edges, and the extracted edge time-series data is obtained.

The method of estimating the movement vector is tried based on this. As a result, it was confirmed that the desired edge could be extracted by DWT and IDWT based on the support length basis function determined based on the proposed method, and it was confirmed that it was effective for the detection and movement tracking of typhoons and tides.

Section 2 describes related research works. Section 3 details the method of constructing basis functions and the definition of their support length. Section 4 proposes the proposed method for analyzing the moving characteristics of moving objects, and Section 5 introduces the above-mentioned case to show that the optimum support length exists and how to obtain it. Finally, the conclusion is stated.

II. RELATED RESEARCH WORKS

As for wavelet analysis related research works, bright band height assignment with precipitation radar data based on Multi-Resolution Analysis: MRA of wavelet analysis is conducted [16]. Also, improvement of secret image invisibility in circulation image with Dyadic wavelet based data hiding with run-length coding is achieved [17].

Extraction of line features from multifidus muscle of CT scanned images with morphological filter together with wavelet multi resolution analysis is made [18]. On the other hand, CO₂ concentration change detection in time and space domains by means of wavelet analysis of MRA: Multi Resolution Analysis is conducted [19].

Polarimetric SAR image classification with high frequency component derived from wavelet multi resolution analysis: MRA is proposed [20]. Meanwhile, embedded object detection with radar echo data by means of wavelet analysis of MRA: Multi Resolution Analysis is conducted [21].

Maximum entropy method based blind separation by means of neural network with feature enhancements by wavelet analysis is proposed [22]. On the other hand, method for support length determination of base function of wavelet for edge and line detection as well as moving object and change detections is proposed [23].

Human gait gender classification using 2D discrete wavelet transforms energy is conducted [24]. Meanwhile, visualization of 3D object shape complexity with wavelet descriptor and its application to image retrievals is attempted together with visualization of 3D object shape complexity with wavelet descriptor and its application to image retrievals [25].

Wavelet based image retrieval method is proposed [26]. Meanwhile, DP matching based image retrieval method with wavelet Multi Resolution Analysis: MRA which is robust against magnification of image size is conducted [27]. On the other hand, wavelet based change detection for four dimensional assimilation data in space and time domains is attempted [28].

Method for image source separation by means of independent component analysis: ICA, Maximum Entropy Method: MEM, and wavelet based method is attempted [29]. On the other hand, identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation is conducted [30].

Method for object motion characteristics estimation based on wavelet Multi-Resolution Analysis: MRA is proposed [31]. Also, identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation is attempted [32].

Comparative study on discrimination methods for identifying dangerous red tide species based on wavelet utilized classification methods is conducted [33]. On the other hand, identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation is conducted [34].

Method for car in dangerous action detection by means of wavelet Multi-Resolution Analysis based on appropriate support length of base function is proposed [35]. Meanwhile, improvement of automated detection method for clustered micro calcification base on wavelet transformation and support vector machine is conducted [36].

Prediction method of El Nino Southern Oscillation: ENSO event by means of wavelet based data compression with appropriate support length of base function is proposed [37]. Meanwhile, method for data hiding based on Legall 5/2 (Cohen-Daubechies-Feauveau: CDF 5/3) wavelet with data compression and random scanning of secret imagery data is proposed [38].

Image retrieval and classification methods based on Euclidian distance between normalized features including wavelet descriptor are proposed [39]. On the other hand, gender classification method based on gait energy motion derived from silhouettes through wavelet analysis of human gait moving pictures is proposed [40].

Multifidus muscle volume estimation based on three dimensional wavelet Multi Resolution Analysis: MRA with buttocks Computer Tomography: CT images is conducted [41]. In a meantime, method for object motion characteristics estimation based on wavelet Multi resolution Analysis: MRA is also proposed [42].

Gender classification method based on gait energy motion derived from silhouette through wavelet analysis of human gait moving pictures is proposed [43]. On the other hand, method for object motion characteristics estimation based on wavelet Multi resolution Analysis: MRA is proposed [44].

Comparative study of feature extraction components for several wavelet transformations for ornamental plants is conducted [45]. Meanwhile, human gait gender classification using 3D discrete wavelet transformation feature extraction is attempted [46].

Image retrieval method utilizing texture information derived from Discrete Wavelet Transformation: DWT together with color information is proposed [47]. Meanwhile, noble method for data hiding using steganography Discrete Wavelet Transformation: DWT and cryptography triple Data Encryption Standard: DES is proposed [48].

Phytoplankton discrimination method with wavelet descriptor based shape feature extraction from microscopic images is proposed [49]. Also, change detection method with multi-temporal satellite images based on wavelet decomposition and tiling is proposed [50]. On the other hand, circle feature extraction from remote sensing satellite images based on polar coordinate representation of wavelets is conducted [51].

III. DISCRETE WAVELET TRANSFORMATION WITH BI-ORTHOGONAL DAUBECHIES BASIS FUNCTION

A. Discrete Wavelet Transformation

The Discrete Wavelet Transformation: DWT of a given discrete scalar signal $\eta = (\eta_1, \eta_2, \dots, \eta_n)^T$ is described as $C_n \eta$ by a square matrix C_n composed of a sequence $\{p_k\}$ and a sequence $\{q_k\}$. p_i is for low-frequency components, coefficient q_i is for high-frequency components, C_n divides f into low-frequency components and high-frequency components.

The given scalar time series is divided into low frequency components and high frequency components by DWT. At this time, for example, the DWT for eight elements having a support length of 2 becomes Eq. (1), and when the support length is 4, it becomes Eq. (2).

$$C_8^{[2]} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \end{bmatrix} = \begin{bmatrix} p_0 & p_1 & \dots & & & & & \\ q_0 & q_1 & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & p_0 & p_1 & & \\ & & & & q_0 & q_1 & & \\ & & & & & & & \\ & & & & & & & \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \end{bmatrix} = \begin{bmatrix} p_0\eta_1 + p_1\eta_2 \\ q_0\eta_1 + q_1\eta_2 \\ p_0\eta_3 + p_1\eta_4 \\ q_0\eta_3 + q_1\eta_4 \\ p_0\eta_5 + p_1\eta_6 \\ q_0\eta_5 + q_1\eta_6 \\ p_0\eta_7 + p_1\eta_8 \\ q_0\eta_7 + q_1\eta_8 \end{bmatrix} \quad (1)$$

$$C_8^{[4]} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \end{bmatrix} = \begin{bmatrix} p_0 & p_1p_2 & p_3 & \dots & & & & \\ q_0 & q_1q_2 & q_3 & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & p_0 & p_1p_2 & p_3 \\ & & & & & & q_0 & q_1q_2 & q_3 \\ & & & & & & & & \\ & & & & & & & & \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \end{bmatrix} =$$

$$\begin{bmatrix} p_0\eta_1 + p_1\eta_2 + p_2\eta_1 + p_3\eta_2 \\ q_0\eta_1 + q_1\eta_2 + q_2\eta_1 + q_3\eta_2 \\ p_0\eta_3 + p_1\eta_4 + p_2\eta_3 + p_3\eta_4 \\ q_0\eta_3 + q_1\eta_4 + q_2\eta_3 + q_3\eta_4 \\ p_0\eta_5 + p_1\eta_6 + p_2\eta_5 + p_3\eta_6 \\ q_0\eta_5 + q_1\eta_6 + q_2\eta_5 + q_3\eta_6 \\ p_0\eta_7 + p_1\eta_8 + p_2\eta_7 + p_3\eta_8 \\ q_0\eta_7 + q_1\eta_8 + q_2\eta_7 + q_3\eta_8 \end{bmatrix} \quad (2)$$

The coefficients of base functions with different That is, the support length is different even if the order is the same. This is called a two-dimensional DWT. Similar to the two-dimensional DWT, the three-dimensional DWT is realized by performing filter processing on each dimension as shown in Fig. 1. Furthermore, multivariate multidimensional DWT is performed by filtering the dimension of each variable.

B. Support Length

Daubechies support lengths are obtained as solutions of simultaneous equations as shown in the following,

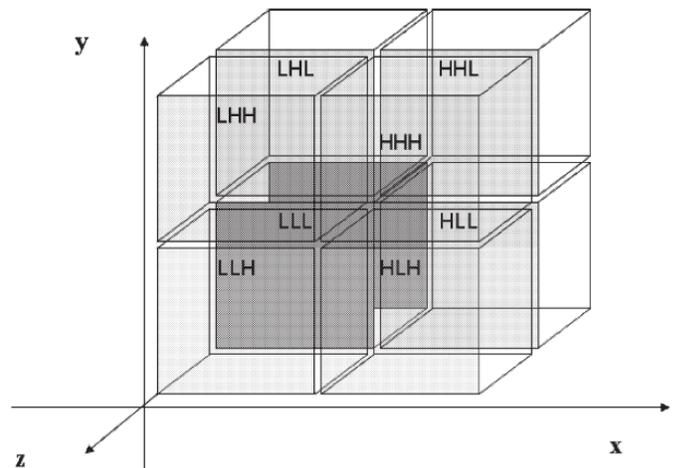


Fig. 1. 3D Discrete Wavelet Transformation (3D DWT).

When the base function is Daubechies and the support length is 2, the coefficient is obtained by solving the simultaneous equations of Eq. (3). When the support length is 4, the coefficient is obtained so as to satisfy Eq. (4). Further, in the case of an arbitrary support length (Sup), the coefficient of the basis function can be obtained by solving Eq. (5) simultaneously.

$$\begin{aligned} C_n^{[2]T} C_n^{[2]} &= I_n \\ p_0 + p_1 &= \sqrt{2}, \\ q_0 = p_1, q_1 = -p_1, 0^0 q_0 + 1^0 q_1 &= 0 \end{aligned} \quad (3)$$

$$\begin{aligned} C_n^{[4]T} C_n^{[4]} &= I_n \\ p_0 + p_1 + p_2 + p_3 &= \sqrt{2}, \\ q_0 = p_3, q_1 = -p_2, q_2 = p_1, q_3 = -p_0, \quad 0^0 q_0 + 1^0 q_1 &= 0 \\ 0^0 q_0 + 1^0 q_1 + 2^0 q_2 + 3^0 q_3 &= 0 \\ 0^1 q_0 + 1^1 q_1 + 2^1 q_2 + 3^1 q_3 &= 0 \end{aligned} \quad (4)$$

$$\begin{aligned} C_n^{[sup]T} C_n^{[sup]} &= I_n \\ \sum_{j=0}^{sup-1} p_j &= \sqrt{2}, \\ q_j &= (-1)^j p_{(sup-1)-j}, j = 0,1,2, \dots, (sup - 1) \\ \sum_{j=0}^{sup-1} j^r q_j &= 0, r = 0,1,2, \dots, (\frac{sup}{2} - 1) \end{aligned} \quad (5)$$

This is called a two-dimensional wavelet transform. The three-dimensional wavelet transform is realized by performing a filtering process on each dimension, like the two-dimensional wavelet transform. Furthermore, the multivariate multidimensional wavelet transform is performed by filtering the dimension of each variable.

C. Multi Dimensional Discrete Wavelet Transformation

By applying the DWT based on the basis function obtained by the above method to the two-dimensional image data, it is possible to obtain an arbitrary frequency component included in the image. In addition, in the case of time-series image data of moving images, it is possible to investigate any frequency component at any time by applying 3D DWT.

Detection of a moving object included in moving image data, tracking, and analysis of movement characteristics are performed by the following procedure. First, the 2D image of each frame number of the moving image is filtered in the horizontal direction of the target 2D image, and divided into two regions of low frequency component (L component) and high frequency component (H component).

Next, filter processing is performed in the vertical direction of each of these two areas. By the above processing, the two-dimensional discrete signal is divided into four (LL component, LH component, HL component, HH component). This is the two-dimensional DWT. Here, if f is the one-dimensional scalar time-series data and F is the spectrum data after DWT, DWT is expressed by Eq. (6).

$$F=Cnf \quad (6)$$

Since the discrete wavelet transform based on Daubechies basis function is considered,

$$CnCn^T=I \quad (7)$$

Therefore, the original time series data can be completely restored by reconstructing after division (reverse transformation after transformation). Also, the 3D DWT for $f_{x,y,z}$ of 3D data is as follows,

$$F=[Cn[Cm[CIf_{x,y,z}]^T]^T]^T \quad (8)$$

The change can be extracted by applying IDWT only to the high frequency components after DWT. For example, if DWT is applied to a two-dimensional image and IDWT is performed with the elements of the LL and LH or LL and HL components set to 0, the two-dimensional reconstructed image extracts vertical or horizontal change (moving) components. It has been done.

The three-dimensional reconstructed image is also obtained by applying the three-dimensional DWT of Fig. 1 to the three-dimensional moving image data for the change component in the time axis direction and performing IDWT with the LLL and LHL or LLL and HLL component elements set to 0. Means that the time change of the vertical or horizontal change (movement) component is extracted.

D. Moving Characteristic Estimation

The moving direction and moving speed are obtained through the above process. At this time, the DWT based on the Daubechies basis function can accurately extract the spatiotemporal changes of the moving object. At that time, it is necessary to appropriately select the basis function to be used and its support length. The basis function may be a perfect orthonormal system, but the support length depends on the movement characteristics to be extracted and the frame rate in moving image acquisition, so the best choice is desired.

E. Method for Optimum Support Length Determination

The Daubechies basis function is selected because it is necessary to be able to completely restore by wavelet transformation and inverse transformation for the movement tracking analysis of the object. The optimum support length differs depending on the movement characteristics of the object and the image acquisition conditions. If the low frequency component is extracted by applying DWT to the multidimensional time series data with time axis and the Root Mean Square Error: RMSE between it and the original image is small, it can be said that the edge component (high frequency component) included in the original image is small. If this difference is large, it means that the edge was extracted most efficiently.

Therefore, in order to determine the optimal support length for edge extraction, DWTs with different support lengths are applied one stage, and the RMSEs of the resulting low-frequency component image and the original image are examined, and this is the optimal support length with a large support length. Propose a method of lengthening. Attention should be paid to the effect of noise enhancement associated with edge extraction.

Wavelet analysis has the function of a filter bank, and it is possible to select the optimum part of the filter bank by adjusting the support length. It is necessary to optimize the support length to prevent the noise from being overemphasized.

This support length depends on the moving speed of the object and the observation speed, and if the observation conditions are the same, there is a support length suitable for the moving speed of the object. That is, qualitatively, it can be said that a basis function with a long support length is suitable for analyzing a phenomenon in which an object moves fast, and conversely, a short support length is suitable for an object moving slowly. Since temporal and spatial changes are generally irrelevant, the author first proposes a method for determining the optimal support length in the image space, and then determining the optimal support length on the time axis.

IV. EXPERIMENT

Here, typhoon and tide detection in GOES images and movement tracking analysis are illustrated as an example of change characteristic analysis by DWT based on Daubechies basis function for change phenomenon of time series data of 2D image.

A. GOES Data used in the Experiment

The author selected the following typical typhoons with fast moving speeds and typical tides off Sanriku with relatively slow moving speeds, and obtained the support lengths of the optimum basis functions for each moving speed.

1) Typhoon: IR1 (10.2-11.2 μm) data of the meteorological satellite GOES-9 acquired on an hourly basis from October 17, 2004, 19:00 (JST) to October 19, 02:00 (JST).

2) Tide: IR1 (10.2-11.2 μm) data of the meteorological satellite GOES-9 acquired on an hourly basis from 00:00 (JST) on October 13, 2008 to 07:00 (JST) on October 14th, 2008.

These images have an image size of 640 x 480 pixels and an image file format PGM (Portable Gray Map). From the data, the area around the typhoon and off Sanriku was extracted as an image size of 256 x 265 pixels and used for the experiment. Some GOES images of typhoons and tides are shown in Fig. 2(a) and (b), respectively.

B. Optimal Support Length Determination Method for Edge Extraction based on RMSE between Reconstructed Image and Original Image by DWT / IDWT

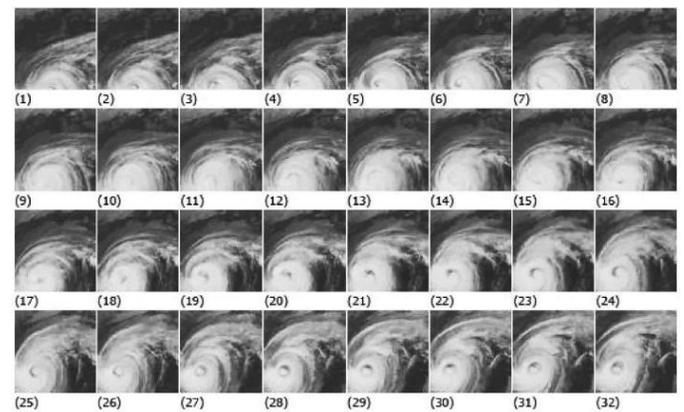
DWT based on Daubechies basis functions with support lengths of 2, 4, 8 and 16 was applied, IDWT was performed using only the LL component, and the RMSE with the original image was evaluated. The reconstructed image and RMSE are shown in Fig. 3 (a) and (b). In the figure, each data indicates the RMSE in each frame.

In the case of the typhoon (a), the RMSE is large up to the first 3 frames and from the last 3 frames, and it can be seen that the time change is sharp. It is also clear that the RMSE of the tide of (b) is low overall, the RMSE is large from the beginning up to about 8 frames, and there is a sharp temporal change with the movement of clouds. It can be seen that the time change is extremely low in the subsequent frames, and there is little movement of the tide.

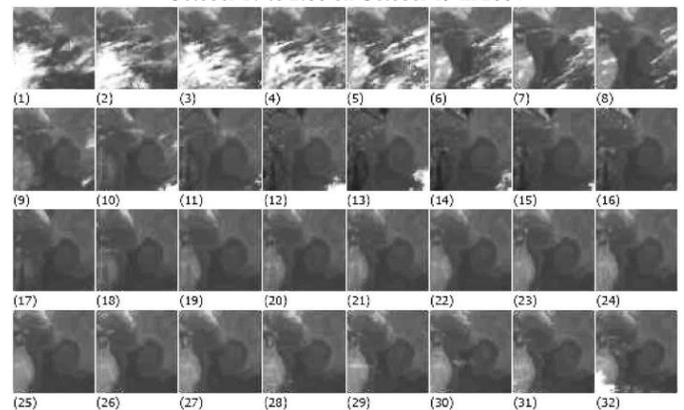
The average and standard deviation of RMSE are shown in Table I(a) and (b). From these, it can be seen that the optimum support length is 8 for each of the typhoons and the tides off Sanriku, since the flatness of the RMSE is the largest.

C. Verification of the Proposed Method by the Number of Extracted Edge Pixels

A threshold of 240 DN (maximum pixel value); the extracted edge image obtained by binarization using 255 DN) is shown in Fig. 4(a) and (b). Table II shows the average and standard deviation of the number of edge pixels extracted by DWT / IDWT with the support length changed to 2, 4, 8 and 16.

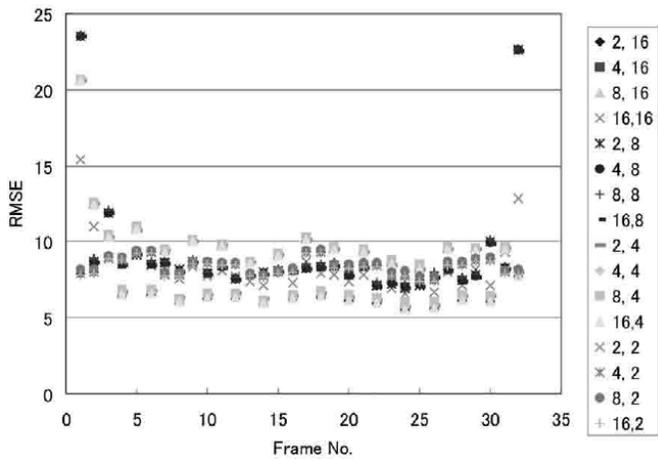


(a) GOES IR-1 imagery data of the typhoon acquired during from 19:00 on October 17 to 2:00 on October 19 in 2004

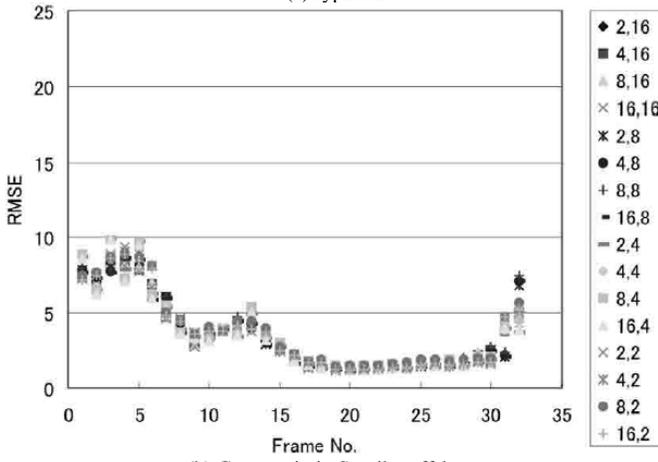


(b) GOES IR-1 imagery data of current rip in Sanriku offshore acquired during from 0: 00 on October 13 to 7: 00 on October 14 in 2004

Fig. 2. The Time Series of Original GOES Imagery Data (32 Frames) of Japanese Vicinity.



(a)Typhoon



(b) Current rip in Sanriku offshore.

Fig. 3. RMSE between the Original and Reconstructed Images without LL Component by using Discrete Wavelet Transformation based on Daubechies Mother Wavelet Function with the different Support Length.

TABLE I. AVE.: AVERAGE AND STD.: STANDARD DEVIATION OF RMSE: ROOT MEAN SQUARE ERROR BETWEEN THE ORIGINAL IMAGE AND RECONSTRUCTED IMAGE WITHOUT LL COMPONENT FROM THE TRANSFORMED IMAGE WITH DISCRETE WAVELET TRANSFORMATION BASED ON DAUBECHIES MOTHER WAVELET FUNCTION WITH SUPPORT LENGTH OF 2, 4, 8 AND 16

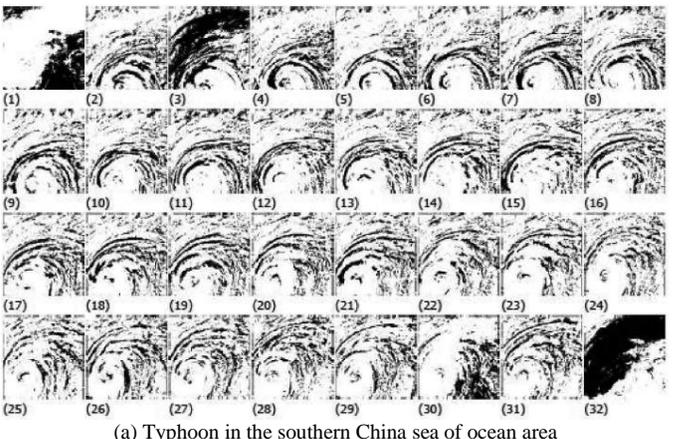
(A) TYPHOON

SupportL.	2	4	8	16	
Ave.	RMSE	8.612	8.629	8.950	8.513
Std.	RMSE	2.376	2.759	2.989	2.547

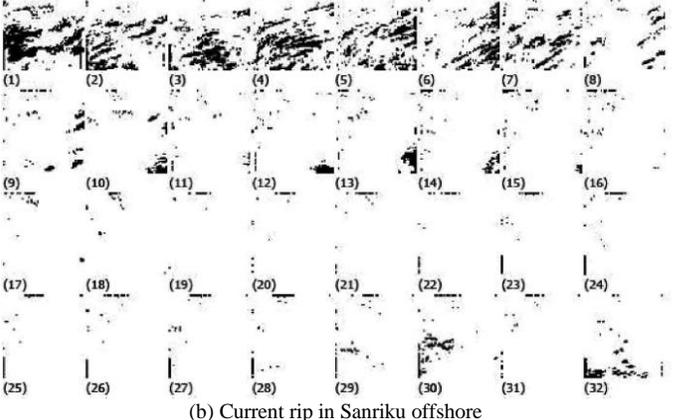
(B) CURRENT RIP IN SANRIKU OFFSHORE

SupportL.	2	4	8	16	
Ave.	RMSE	3.574	3.594	3.741	3.487
Std.	RMSE	2.525	2.443	2.411	2.444

From Table II(a), the optimum support length for edge extraction and movement analysis for the tide off Sanriku is considered to be 2 (when the support length is 2, the number of flat edge pixels is 0.5% compared to when the support length is 8.



(a) Typhoon in the southern China sea of ocean area



(b) Current rip in Sanriku offshore

Fig. 4. Example of Detected Edge Images of(a) Typhoon and (b) Current Rip in Sanriku offshore Obtained from Binarized Images with the Threshold of 240 Digital Number: DN (255 in Maximum DN)of Reconstructed Images without LL Component after the Discrete Wavelet Transformation with Daubechies Mother Wavelet Function with Support Length of 8.

TABLE II. AVE.: AVERAGE AND STD.: STANDARD DEVIATION OF RMSE: ROOT MEAN SQUARE ERROR BETWEEN THE ORIGINAL IMAGE AND RECONSTRUCTED IMAGE WITHOUT LLL COMPONENT FROM THE TRANSFORMED IMAGE WITH DISCRETE WAVELET TRANSFORMATION BASED ON DAUBECHIES MOTHER WAVELET FUNCTION WITH SUPPORT LENGTH OF 2, 4, 8 AND 16

Support Length.	2	4	8	16	
Ave.	RMSE	14.292	13.229	10.860	20.474
Std.	RMSE	8.671	7.713	6.747	30.109

This is because, as can be seen from the standard deviation, the edges accompanying the movement of clouds included in the images of frame numbers 1-5 are large and the standard deviation is raised. Therefore, taking this into consideration, it was confirmed that the DWT / IDWT based on basis functions with a support length of 8 is optimal for spatiotemporal edge extraction. In addition, the effect of noise increase due to edge enhancement can be seen from Fig. 4(a), (b) and Table II. When the edge component is large like a typhoon image, the influence is large. The standard deviation in the table reflects the noise component, but when the support length is optimum, the noise component, that is, the standard deviation, is small.

D. Optimal Support Length for Moving Vector Detection along the Time Axis

As mentioned above, temporal changes and spatial changes are originally different. Therefore, it is better to determine the optimal support length in the time axis direction separately from spatial edge extraction. The method is as follows:

1) Determine the optimal support length for extracting the spatial change based on the RMSE described above by two-dimensional DWT / IDWT.

2) Next, find the optimum support length for detecting the movement vector in the time axis direction. Therefore, first, a two-dimensional DWT based on the optimal support length basis function found in (1) is applied to the original image to generate a time-series DWT transformed image data set.

3) After that, one-dimensional DWT is applied to the time-series data in which the transformed image is arranged in the time axis direction for each pixel, and IDWT is performed using only the L component.

4) After that, the RMSE between the reconstructed image and the original image is evaluated to determine the optimal support length for extracting changes in the time axis direction. At this time, the changes in the RMSE time axis direction for each support length were compared with the RMSE time average.

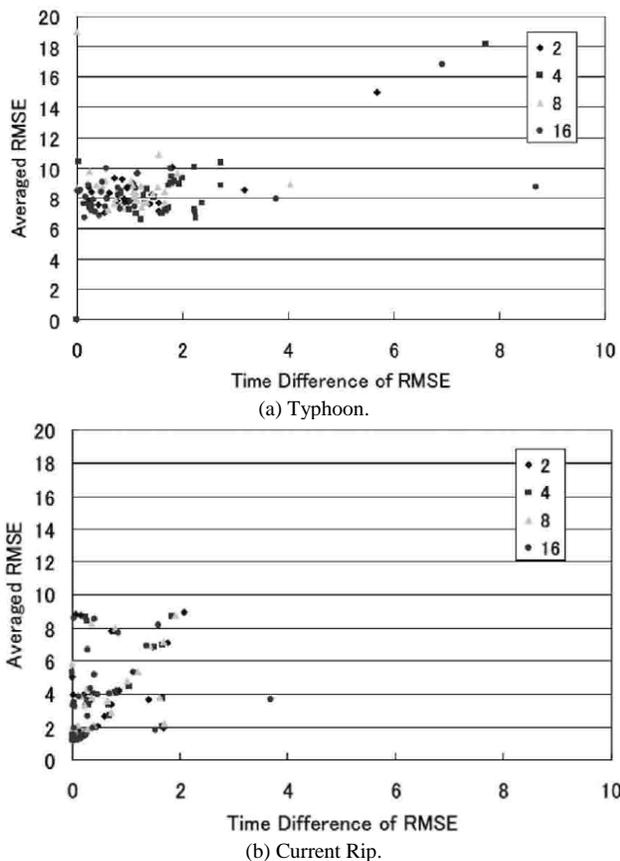


Fig. 5. Most Appropriate Support Length of Daubechies Mother Wavelet Function for Detection of Moving Targets from Time Series of GOES Imagery Data.

As shown in Fig. 5, it can be seen that, as a general tendency, the time flatness of RMSE is large when the time change is large. In addition, it can be seen that the optimum support length is different when the moving object has a steep temporal change such as a typhoon and when it moves slowly like a tide. In other words, the temporal change of RMSE in the case of the tide is concentrated within 2 and the time average of RMSE is also within 9.

A larger time change than that is the steep part of the time change associated with the movement of clouds in frame numbers 1-5. On the other hand, in the case of a typhoon, the time variation of RMSE is concentrated within about 4 and the time average of RMSE reaches 18. Furthermore, the support length of the part of the RMSE that has a large change over time is often 16 and is followed by 8, 4, and 2.

Therefore, it can be seen that the DWT / IDWT analysis based on the basis function with a long support length is suitable for detecting and tracking moving objects with large temporal changes. In summary, it can be construed that when detecting and tracking a phenomenon that changes rapidly, it is necessary to extend the support length and view the observation data over a wide time range.

V. CONCLUSION

In this paper, the author first show that spatiotemporal feature extraction based on Daubechies basis functions is effective for detecting moving objects such as typhoons, clouds, and tidal waves in Earth observation satellite image data, and analyzing movement vectors.

The author proposed a method for optimizing the support length of the basis function used in that case. In other words, the three-dimensional discrete wavelet transform based on the basis function of the support length in the assumed range (1, 2, 4, 8, 16, 32, etc.) is applied to the time-series earth observation satellite image data in one stage, and low frequency components (LLL Component (reverse discrete wavelet transform) is used to evaluate the squared deviation between the original data and the reconstructed data, and this is large (edge component due to movement of object: high-frequency component is extracted greatly). The optimum support length is determined with such support length.

The author evaluated the effectiveness of the proposed method by applying it to time-series data of geostationary meteorological satellite GOES images, and found that the Daubechies basis function, which has a long support length, can be used to analyze the motion vector of an object that moves rapidly like a typhoon. It was confirmed that the spatiotemporal feature extraction based on the spatiotemporal feature extraction based on the Daubechies basis function, which has a short support length, is effective for the motion vector analysis of the phenomenon such as the tidal movement.

VI. FUTURE RESEARCH WORKS

In the future, the author will continue to validate of the proposed method with a variety of time series of data.

ACKNOWLEDGMENT

This is a report of the research results (or part of it) by the Grant-in-Aid for Scientific Research (Special Researcher Incentive). The author would like to express my gratitude to the Grant provider. The author would like to thank Dr. Kaname Seto of former student of Saga University and Dr. Leland M. Jameson of Naval Research Laboratory for their contribution of this study. The author, also, would like to thank Professor Dr. Hiroshi Okumura and Professor Dr. Osamu Fukuda for their valuable discussions.

REFERENCES

- [1] Kohei Arai, Basic Theory of Wavelet Analysis, Morikita Publishing (November 2000).
- [2] Kohei Arai, Leland Jameson, How to use earth observation satellite data by wavelet analysis, Morikita Publishing (July 2001).
- [3] Kohei Arai, Self-study wavelet analysis, published by Modern Science Co., Ltd. (June 2006).
- [4] Saito, N., Simultaneous noise suppression and signal compression using a library of orthogonal bases and the minimum description length criterion, In Wavelets in Geophysics, Foufoula-Georgiou, E. and Kumar, P. (ed), Academic Press, pp.299-324, 1994.
- [5] Yuki Matsushima, Shingo Shirahata, Wataru Sakamoto, Selection of Wavelet Basis Functions by AIC, Applied Statistics, 33, 2, 201-219, 2004.
- [6] Corbett J., Leduc J.-P., and Kong M., Analysis of de-formational transformations with spatiotemporal continuous wavelet transforms, in Proceedings of IEEE-ICASSP, 1999.
- [7] Duval-Destin M. and Murenzi R., Spatio-temporal wavelets : Applications to the analysis of moving patterns, Progress in Wavelet Analysis and Applications, Editions Frontières, Gif-sur-Yvette, France, 1993.
- [8] J.-P.Leduc, F.M ujica, R.M urenzi, and M .J.T. Smith, Spatio temporal Wavelets : A Group-Theoretic Construction for Motion Estimation and Tracking. SIAM J. Appl. Math. Society for Industrial and Applied Mathematics 61, 2, 596-632, 2000.
- [9] F.M ujica, J.-P.Leduc, R.M urenzi and M.J.T. Smith, A New Motion Parameter Estimation Algorithm Based on the Continuous Wavelet Transform IEEE Transactions on Image Processing, 9, 5, 873-888, 2000.
- [10] Bing Wang, Rong Chun Zhao, Motion tracking using the continuous wavelet transform and EKF models, Proceedings of the International Computer Congress 2004, 155-161, World Scientific Publishing Co. Pte. Ltd., 2004.
- [11] Braut P., A New Scheme For Object-Oriented Video Compression And Scene Analysis, Based On Motion Tuned Spatio-Temporal Wavelet Family and Trajectory Identification, IEEE-ISSPIT03, 2003.
- [12] Kohei Arai, Method for support length determination of base function of wavelet for edge and line detection as well as moving object and change detections, International Journal of Research and Reviews on Computer Science, 2, 4, 1133-1139, 2011.
- [13] Kohei Arai, Tomoko Nishikawa, Method for car in dangerous action detection by means of wavelet Multi-Resolution Analysis based on appropriate support length of base function, International Journal of Advanced Research in Artificial Intelligence, 2, 4, 13-17, 2013.
- [14] Kohei Arai, Prediction method of El Nino Southern Oscillation event by means of wavelet based data compression with appropriate support length of base function, International Journal of Advanced Research in Artificial Intelligence, 2, 8, 16-20, 2013.
- [15] Kohei Arai and M.Saka, Bright band height assignment with precipitation radar data based on Multi-Resolution Analysis: MRA of wavelet analysis, Advances in Space Research, 37, 12, 2197-2201, 2006.
- [16] Kohei Arai and Yuji Yamada, Improvement of secret image invisibility in circulation image with Dyadic wavelet based data hiding with run-length coding, International Journal of Advanced Computer Science and Applications, 2, 7, 33-40, 2011.
- [17] Kohei Arai, Yuichiro Eguchi and Yoichiro Kitajima, Extraction of line features from multifidus muscle of CT scanned images with morphological filter together with wavelet multi resolution analysis, International Journal of Advanced Computer Science and Applications, 2, 8, 60-66, 2011.
- [18] Kohei Arai, CO₂ concentration change detection in time and space domains by means of wavelet analysis of MRA: Multi Resolution Analysis, International Journal of Advanced Computer Science and Applications, 2, 8, 82-86, 2011.
- [19] Kohei Arai, Polarimetric SAR image classification with high frequency component derived from wavelet multi resolution analysis: MRA, International Journal of Advanced Computer Science and Applications, 2, 9, 37-42, 2011.
- [20] Kohei Arai, Embedded object detection with radar echo data by means of wavelet analysis of MRA: Multi Resolution Analysis, International Journal of Advanced Computer Science and Applications, 2, 9, 27-32, 2011.
- [21] Kohei Arai, Maximum entropy method based blind separation by means of neural network with feature enhancements by wavelet analysis, International Journal of Research and Reviews on Computer Science, 2, 4, 1116-1122, 2011.
- [22] Kohei Arai, Method for support length determination of base function of wavelet for edge and line detection as well as moving object and change detections, International Journal of Research and Reviews on Computer Science, 2, 4, 1133-1139, 2011.
- [23] Kohei Arai, Rosa Andrie, Human gait gender classification using 2D discrete wavelet transforms energy, International Journal of Computer Science and Network Security, 11, 12, 62-68, 2011.
- [24] Kohei Arai, Visualization of 3D object shape complexity with wavelet descriptor and its application to image retrievals, Journal of Visualization, DOI:10.1007/s, 12650-011-0118-6, 2011.
- [25] Kohei Arai, Visualization of 3D object shape complexity with wavelet descriptor and its application to image retrievals, Journal of Visualization, 15, 2, 155-166, 2012.
- [26] Kohei Arai, C.Rahmad, Wavelet based image retrieval method, International Journal of Advanced Computer Science and Applications, 3, 4, 6-11, 2012.
- [27] Kohei Arai, DP matching based image retrieval method with wavelet Multi Resolution Analysis: MRA which is robust against magnification of image size, International Journal of Research and Review on Computer Science, 3, 4, 1738-1743, 2012.
- [28] Kohei Arai, Wavelet based change detection for four dimensional assimilation data in space and time domains, International Journal of Advanced Computer Science and Applications, 3, 11, 71-75, 2012.
- [29] Kohei Arai, Method for image source separation by means of independent component analysis: ICA, Maximum entropy method, and wavelet based method, International Journal of Advanced Computer Science and Applications, 3, 11, 76-81, 2012.
- [30] Kohei Arai, I.N.Abdulah, H.Okumura, Identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation, International Journal of Computer Applications, 61, 1, 34-38, 2013.
- [31] Kohei Arai Method for object motion characteristics estimation based on wavelet Multi-Resolution Analysis: MRA, International Journal of Advanced Research in Artificial Intelligence, 2, 1, 25-32, (2013).
- [32] Kohei Arai, Indra Nugraha Abdulah, Hiroshi Okumura, Identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation, International Journal of Computer applications, 61, 1, 1-5, (2013).
- [33] Kohei Arai, Comparative study on discrimination methods for identifying dangerous red tide species based on wavelet utilized classification methods, International Journal of Advanced Computer Science and Applications, 4, 1, 95-102, 2013.
- [34] Kohei Arai, Indra Nugraha Abdullah, Hiroshi Okumura, Identification of ornamental plant functioned as medicinal plant based on redundant discrete wavelet transformation, International Journal of Advanced Research in Artificial Intelligence, 2, 3, 60-64, 2013.
- [35] Kohei Arai, Tomoko Nishikawa, Method for car in dangerous action detection by means of wavelet Multi-Resolution Analysis based on

- appropriate support length of base function, International Journal of Advanced Research in Artificial Intelligence, 2, 4, 13-17, 2013.
- [36] Kohei Arai, Indra Nugraha Abdullah, Hiroshi Okumura, Rie Kawakami, Improvement of automated detection method for clustered micro calcification base on wavelet transformation and support vector machine, International Journal of Advanced Research in Artificial Intelligence, 2, 4, 23-28, 2013.
- [37] Kohei Arai, Prediction method of El Nino Southern Oscillation event by means of wavelet based data compression with appropriate support length of base function, International Journal of Advanced Research in Artificial Intelligence, 2, 8, 16-20, 2013.
- [38] Kohei Arai, Method for data hiding based on Legall 5/2 (Cohen-Daubechies-Feauveau: CDF 5/3) wavelet with data compression and random scanning of secret imagery data, International Journal of Wavelets Multi Solution and Information Processing, 11, 4, 1-18, B60006 World Scientific Publishing Company, DOI: 10.1142/SO219691313600060, 1360006-1, 2013.
- [39] Kohei Arai, Image retrieval and classification method based on Euclidian distance between normalized features including wavelet descriptor, International Journal of Advanced Research in Artificial Intelligence, 2, 10, 19-25, 2013.
- [40] Kohei Arai, Rosa Andrie Asmara, Gender classification method based on gait energy motion derived from silhouettes through wavelet analysis of human gait moving pictures, International Journal of Information Technology and Computer Science, 6, 3, 1-11, 2014.
- [41] Kohei Arai, Multifidus muscle volume estimation based on three dimensional wavelet Multi Resolution Analysis: MRA with buttocks Computer Tomography: CT images, International Journal of Advanced Research in Artificial Intelligence, 2, 12, 9-15, 2013.
- [42] Kohei Arai, Method for object motion characteristics estimation based on wavelet Multi resolution Analysis: MRA, International Journal of information Technology and Computer Science, 6, 1, 41-49, DOI: 10.5815/ijitcs, 2014.01.05, 2014.
- [43] Kohei Arai, Rosa Andrie Asmara, Gender classification method based on gait energy motion derived from silhouette through wavelet analysis of human gait moving pictures, International Journal of Information technology and Computer Science, 5, 5, 12-17, 2013.
- [44] Kohei Arai, Method for object motion characteristics estimation based on wavelet Multi resolution Analysis: MRA, International Journal of Information Technology and Computer Science, 1, 41-49, DOI: 10.5815/ijitcs, 2014.01.05, 2014.
- [45] Kohei Arai, Indra Nugraha Abdullar, H.Okumura, Comparative study of feature extraction components for several wavelet transformations for ornamental plants, International Journal of Advanced Research in Artificial Intelligence, 3, 2, 5-11, 2014.
- [46] Kohei Arai, Rosa Andrie Asmara, Human gait gender classification using 3D discrete wavelet transformation feature extraction, International Journal of Advanced Research in Artificial Intelligence, 3, 2, 12-17, 2014.
- [47] Kohei Arai, Cahya Rahmad, Image Retrieval Method Utilizing Texture Information Derived from Discrete Wavelet Transformation Together with Color Information, International Journal of Advanced Research on Artificial Intelligence, 5, 10, 1-6, 2016.
- [48] Cahya Rahmed Kohei Arai, Arief Prasetyo, Noriza Arigki, Noble Method for Data Hiding using Steganography Discrete Wavelet Transformation and Cryptography Triple Data Encryption Standard: DES, International Journal of Advanced Computer Science and Applications IJACSA, 9, 11, 261-266, 2018.
- [49] Kohei Arai, Phytoplankton discrimination method with wavelet descriptor based shape feature extraction from microscopic images, Journal of RIMS Signal and time frequency analysis, Kokyuroku edited by Ryuichi Ashino, Kyoto University, 50-86, 2019.
- [50] Kohei Arai, Change Detection Method with Multi-temporal Satellite Images based on Wavelet Decomposition and Tiling, Journal of Advanced Computer Science and Applications, Vol. 12, No. 3, 56-61, 2021.
- [51] Kohei Arai, Circle Feature Extraction from Remote Sensing Satellite Images Based on Polar Coordinate Representation of Wavelets, FICC 2021: Advances in Information and Communication, Springer, vol.1, pp 558-566, 2021.

AUTHORS' PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is a Science Council of Japan Special Member since 2012. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 55 books and published 620 journal papers as well as 450 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. <http://teagis.ip.is.saga-u.ac.jp/index.html>.

DeepfakeNet, an Efficient Deepfake Detection Method

Dafeng Gong¹

Wenzhou Polytechnic, Wenzhou, China
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia

Yogan Jaya Kumar², Ong Sing Goh³

Zi Ye⁴, Wanle Chi⁵
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia

Abstract—Different CNNs models do not perform well in deepfake detection in cross datasets. This paper proposes a deepfake detection model called DeepfakeNet, which consists of 20 network layers. It refers to the stacking idea of ResNet and the split-transform-merge idea of Inception to design the network block structure. That is, the block structure of ResNeXt. The study uses some data of FaceForensics++, Kaggle and TIMIT datasets, and data enhancement technology is used to expand the datasets for training and testing models. The experimental results show that, compared with the current mainstream models including VGG19, ResNet101, ResNeXt50, XceptionNet and GoogleNet, in the same dataset and preset parameters, the proposed detection model not only has higher accuracy and lower error rate in cross dataset detection, but also has a significant improvement in performance.

Keywords—DeepfakeNet; deepfake detection; data enhancement; CNNs; cross dataset

I. INTRODUCTION

Face is one of the most representative features in human beings' biometrics, with high recognition. At the same time, with the rapid development of face synthesis technology, the security threat brought by face tampering is becoming more and more serious. Especially in the era of mobile phones highly popular and social networks increasingly mature, the deepfake video using deep network model to replace face spreads rapidly in social media and the Internet, such as deepfacelab, faceswap [1].

At the end of 2018, the Dutch Deeptech laboratory released a Deepfake development report[2] showing: Deepfake's global search volume has stabilized to as much as 1 million by December 2018, with at least 14678 fake videos, including 96% of pornography, and risk content such as violence, political sensitivity, advertising, contraband, etc., disguised in the video. The appearance of AI face changing is undoubtedly a great impact on its objective authenticity.

The threat of fake face video to society is increasing, and it has attracted widespread attention from academia and industry. There have been some related studies, and there are also international competitions for face-changing video detection[3]. According to the features used, the existing fake face video detection technologies are roughly divided into three categories[4]: based on traditional manual features, based on biological features, and based on neural network extraction

of features. The first type of method mainly draws on the idea of image forensics and analyzes a single frame of images. Typical methods include the use of image quality measurement and principal component analysis[5] and the use of local binary pattern (LBP) features[6]. The second type of method mainly uses the unique biological information of the face. X. Yang et al. [7] proposed a model that can divide the facial landmarks into two groups according to the degree of influence during the tampering process, and use different landmarks to estimate the head posture direction and compare the differences. As a basis for discrimination, F. Matern et al. [8] found that the diffuse reflection information presented by the pupils of the two eyes in the fake face is inconsistent; The study by P. Korshunov and S. Marcel [9] uses both the video image and audio information to compare the lips in the true and false video. The difference between action and voice matching distinguishes whether there is tampering; S. Agarwal et al. [10] pointed out that every person has a unique movement pattern, and changing faces leads to a mismatch between the target object and the source object's movement patterns, which can be measured from the forehead, cheeks, nose, etc. The features are extracted from the movement changes of the region for classification decision. The third type of method mainly learns human faces by constructing a convolutional neural network, and extracts higher-dimensional semantic features for classification. Some researchers regard it as a conventional classification problem. The study by A. Khodabakhsh et al. [6] uses classic classification models such as AlexNet[11], VGG-19[12], ResNet[13], Inception and Xception[14] for image recognition to detect. D. Afchar et al. [15] Built Meso-4 and MesoInception-4, and S. Tariq et al. [16] built ShallowNet to detect single-frame images; B. Bayar and M. C. Stamm [17] pointed out that in the problem of tamper detection, tampering traces are more important than image content information. The MISLnet of the constrained convolutional layer suppresses the image content when extracting the shallow features; D. Guera et al. [18] considers the time domain information in the video and combines the convolutional neural network with the sequential neural network to find the continuous frame features in the fake face video inconsistency; S.-Y. Wang et al. [19] uses the ResNet-50[13] network model to detect different GAN composite images and Deepfake face images.

From the experimental results given in the above research, the high performance of the algorithm based on neural network extraction features often depends on the dataset used. In the

cross-dataset, due to the large domain offset, that is, the size of the dataset target is different, most of the algorithms and models performs poorly, such as VGG19, GoogleNet, XceptionNet, ResNet50. Background complexity, resolution, and the quality of the synthetic fake face are different, which makes the data distribution vary greatly, which leads to the model's inability to make correct judgments and poor detection results, resulting in a significant drop in performance during cross-dataset [6].

This paper proposes a solution to the generalization performance of face-swapping video detection. Different from the above-mentioned method based on feature detection, and starts from the image and believes that fake face tampering is a special splicing tampering problem. According to the fact that face changing mainly operates on part of the face area without modifying the content of other images, DeepfakeNet is proposed, which is a detection method based on image segmentation and deep residual network. The key contributions in this study are: (1) Extracting video data from multiple data sources, and creating a unified experimental data set through data enhancement methods; (2) Proposing an improved structure based on ResNeXt[20], as whether face change occurs judgment basis for tampering; (3) A model parameter is trained to obtain better detection results.

This paper first introduces the generation principle of deepfake and the current detection technology of deepfake; Then, the DeepfakeNet detection algorithm proposed in this paper is described, and the network structure of the model is explained; Then set up the experimental environment, expand the data set using data enhancement technology, and use relevant standards to compare the mainstream models to draw conclusions; Finally, the possible direction of further efforts in this field is put forward.

In this study, Prof. & Dr. Goh put forward the overall idea; Dr. Yogan gives guidance and improvement; Dafeng Gong designs algorithm, implementation, and analysis conclusion; Zi Ye preprocesses the dataset; Wanle Chi verifies the model.

II. RELATED TECHNOLOGY

A. The Basic Architecture of Deepfake

Deepfake is a neural network trained with an unsupervised learning method. It regenerates the original input after encoding the distorted face image, and expects this network to have the ability to restore any face. The overall process of implementing Deepfake to change the face is shown in Fig. 1, that is, the final realization of FaceA to replace FaceB in a video or image. First, obtain images with A/B faces from the video or image collection, and perform face detection and face alignment. Specifically, on the basis of face detection, the location of key facial features is performed, and the detected faces are normalized and aligned through affine transformation, which can intercept half-face or full-face facial images; then The intercepted A/B face image is sent to the neural network for training and conversion to obtain A face with B expression, action, environment and other conditions, and then the output face image is overlaid on the original image (Fig. 1), and the edges are smoothed then re-synthesize the video.

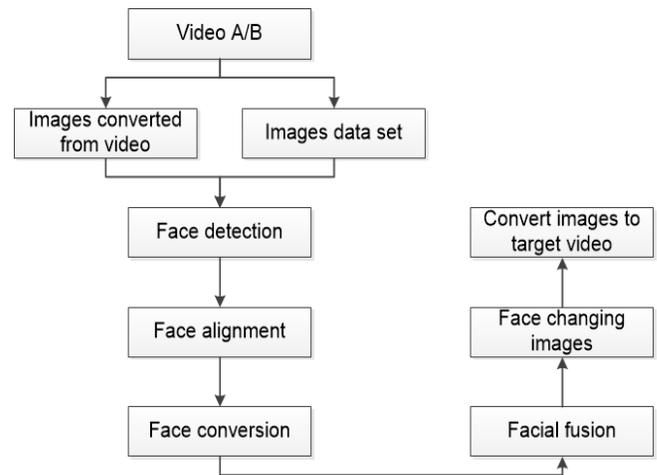


Fig. 1. Face-Swapping Process through Deepfake.

The above process is based on GAN network, and the generation process of data can be represented by the following two public announcements. Goodfellow, I. J et al. [21] proposed data is generated in the following way:

$$\lim_{\sigma \rightarrow 0} \nabla_x E_{\epsilon \sim N(0, \sigma^2)} f(x + \epsilon) = \nabla_x f(x) \quad (1)$$

The function of optimization of multilayer perceptrons is:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (2)$$

In the above equation, $D(x)$ represents the probability of D judging x from real data, $D(G(z))$ represents the probability of D judging $G(z)$ as real data, Z represents random noise, and $G(z)$ is the probability of generating data, that is, the probability of D judging as fake, and $1-D(G(z))$ is the probability of D judging as true; for D , the probability of other judging as true is maximized, while for G , the objective function is minimized.

B. Deepfake Detection

The development of Deepfake technology also gave birth to the corresponding detection technology. Yuezun Li et al. [22] used a deep neural network model to combine the convolutional neural network (CNN) and the recurrent neural network (RNN) to form a long and short-term memory network (LRCN). It is believed that fake videos generally do not show the characteristics of blinking, breathing, and eye movement, so that it can detect whether the face in the image or video is real or generated by AI; Matern et al. [8] proposed using image artifacts to detect facial forgery; Yuezun Li et al. [23] trained CNN to detect this facial artifact in face-changing videos; Haodong Li et al. [24] used the difference in color components between real images and forged facial images to distinguish between authenticity and fake images; Yang et al. [7] proposed an SVM support vector machine classification method based on the mismatch between the head pose and the position of the important facial features; Koopman et al. [25] predicted that the tampering of the facial region would affect the local illumination response inconsistency noise in the video frame, and proved that the noise can be used distinguish between original video and face-changing video, but a larger

data set must be studied to determine the correlation between the two; because FakeApp generates Deepfake fake video, it will cause inconsistencies within and between frames, Guera et al. [18] proposed the use of recurrent neural networks (RNN) to detect this anomaly in videos; Afchar et al. [15] proposed two low-level neural networks for detection; Korshunov et al. [5] evaluated the reliability of several detection methods and proposed: Advanced VGG and Facenet-based neural network face recognition algorithms cannot distinguish face-changing videos from original videos. Detection algorithms based on lip shape and voice inconsistency cannot distinguish Deepfake fake videos, while image quality detection based on support vector machine classification. The technology can detect high-quality (128×128) face-changing images with low error.

However, algorithms based on neural network extraction of features can often achieve higher accuracy in cross-dataset detection, and the main drawback is that the performance of cross-dataset detection drops sharply, and there is a problem of insufficient generalization performance[6].

III. DEEPPFAKE DETECTION ALGORITHM

This study refers to the ResNeXt network [20], using ResNet's stacking ideas [13] and Inception's split-transform-merge ideas [26]. The calculation process can be expressed as follows:

$$y = x + \sum_{i=1}^C T_i(x) \tag{3}$$

Where x is a short-cut, each feature undergoes a linear transformation, C is the cardinality of simple Inception, and T_i is any transformation, such as a series of convolution operations.

Its basic block structure is the same as ResNeXt, as shown in Fig. 2. Each box represents a layer, and the meanings of the three data representations are: input data channel, filter size, output data channel. The advantage is to improve the accuracy through a wider or deeper network under the premise of ensuring the amount of FLOPs and parameters. For each block structure, the convolution kernel is grouped by channel to reduce the dimensionality of the data to form 32 parallel branches, and then respectively perform convolution transformation and feature change on 32 low-latitude data, and then aggregate them back to the original by addition dimension. The final network structure is shown in Table I, consisting of a 20-layer network, here called DeepfakeNet. This network architecture is shown in Fig. 3.

There are five groups of convolutions in this network. The input image size of 1st group is 224x224, and the size of output data of 5th group is 7x7, which is reduced by 32 (2^5) times. Each time, the stride is set to 2 on the first layer of each group of convolutions, and each time is reduced by a factor of 2, a total of 5 times is reduced.

The overall architecture and process of the detection model is mainly composed of 3 main parts: preprocessing data module, extracting feature module and deep learning model module. In the data preprocessing module, the video data set is first processed. For frame images, then enhance the data set. Since only the face of the person in the video frame is the

detection target, the extracted video frame is intercepted, and the features of face images are extracted by CNN. At the same time, the powerful extracting feature ability of CNN can be used to more accurately judge images whether it has been modified. After sufficient training and verification, the DeepfakeNet model is continuously improved to obtain better results.

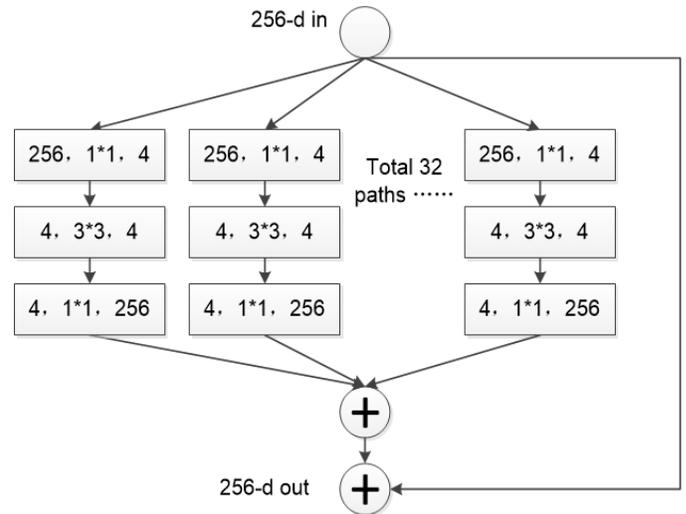


Fig. 2. A Block of ResNeXt with Cardinality = 32, with Roughly the Same Complexity.

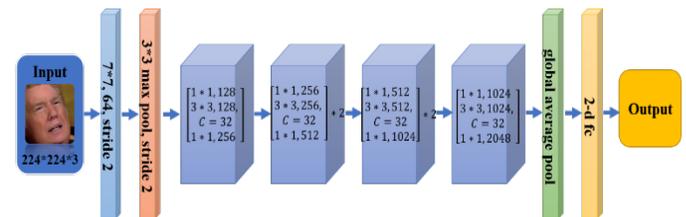


Fig. 3. DeepfakeNet Architecture.

TABLE I. DEEPPFAKENEt WITH A 32*4D TEMPLATE.

state	output	DeepfakeNet (32*4d)
Conv1	112*112	7*7, 64, stride 2
Conv2	56*56	3*3 max pool, stride 2
		$\begin{bmatrix} 1 * 1, 128 \\ 3 * 3, 128, C = 32 \\ 1 * 1, 256 \end{bmatrix} * 1$
Conv3	28*28	$\begin{bmatrix} 1 * 1, 256 \\ 3 * 3, 256, C = 32 \\ 1 * 1, 512 \end{bmatrix} * 2$
		$\begin{bmatrix} 1 * 1, 512 \\ 3 * 3, 512, C = 32 \\ 1 * 1, 1024 \end{bmatrix} * 2$
Conv4	14*14	$\begin{bmatrix} 1 * 1, 1024 \\ 3 * 3, 1024, C = 32 \\ 1 * 1, 2048 \end{bmatrix} * 1$
		$\begin{bmatrix} 1 * 1, 1024 \\ 3 * 3, 1024, C = 32 \\ 1 * 1, 2048 \end{bmatrix} * 1$
	1*1	global average pool 2-d fc, softmax
params		10.87 * 106
FLOPs		2.05 * 109

IV. EXPERIMENTAL SETUP

A. Lab Environment

In order to verify the performance and effectiveness of the detection model, this study selects videos from the deepfake open source datasets such as FaceForensics++[27], TIMIT[5] and Kaggle competitions[3] for experiments, as shown in Table II. Since the deep learning model in this article reads the training set according to a sequence of consecutive frames, the video must be converted into a sequence of frame images.

In order to obtain an input image of a uniform size, the videos in each dataset are divided into frames, and the face is located by the convolutional neural network detector in the Dlib library frame by frame, and k ($k > 1$) times the face is taken as the center of the face frame. The size of the image area, sampled to the size of 224×224 , as the input image. In order to effectively train and test the network, this paper expands the training samples to enhance the diversity of the data, so that the model can adapt to a wide range of application environments, and has a wide range of applications in target recognition and target detection.

Common methods of enhancing dataset include stretching, rotating, flipping, etc., as shown in Table III. Perform the following operations on each image: compress or stretch to between 0.75 and 1.25 times; each image is generated from 30 degrees left to 30 degrees right, every 2 degrees; brightness changes from the original brightness 0.75 to 1.25 times of, each 0.1 times difference generates one; at the same time, each image is flipped horizontally and vertically. These operations are performed in order to obtain a sufficiently large set of enhanced data[15][28].

This experiment was done on an ubuntu server with 2 Intel Xeon Silver 4214 CPUs, 192GB RAM, and NVIDIA Tesla V100 32GB PCIe GPU. The deep learning model is implemented using Python language and Pytorch framework.

In order for the model to fully learn the feature information of the data set, this study sets the number of iterations of the model epoch to 300, and sets the loss function of the deep learning model to the mean-square error (MSE) loss function. The calculation method is as follows:

$$L_{loss} = \frac{1}{2m} \sum_{i=1}^m (\hat{y}_i - y_i)^2 \quad (4)$$

In the formula, m is the number of samples, y_i is the label of the sample, and \hat{y}_i is the predicted value of the model. The loss function L_{loss} can well show the degree of fit between the true label and the predicted results. The smaller the value, the better the degree of fit. In deep learning, the optimization algorithm is used to optimize the model, so that the value of L_{loss} tends to the minimum, and the fitting ability of the model is increased. The dropout used in the training of this paper is set to 0.8, the learning rate is set to 1×10^{-5} , and the optimization algorithm is Adam [20]. The experimental data set is divided into a training dataset and a verification dataset. Each experiment was randomly assigned to generate them, corresponding to 80% and 20% respectively. The model is trained on the training dataset and verified on the verification dataset.

TABLE II. COMPOSITION OF DATA SOURCES

Data Source	Total Quantity	Number of Videos Selected
FaceForensics++	4000	3200
Kaggle	117812	21767
TIMIT	1199	960

TABLE III. APPROACHES OF ENHANCING DATA

Types	stretching	rotating	Brightness change	Flip horizontally/vertically
Params	[0.75, 1.25]	[-30, 30]	[0.75, 1.25]	Yes/Yes

B. Evaluation Measures

In order to evaluate the performance based on the temporal and spatial feature consistency detection model, this paper selects a variety of metrics to evaluate the model. First of all, for classification models, accuracy is often used to evaluate the global accuracy of a model. The higher the accuracy, the better the accuracy of the model. The calculation formula of the accuracy rate is as follows:

$$A_{accuracy} = \frac{T_{TP} + T_{TN}}{T_{TP} + T_{TN} + F_{FP} + F_{FN}} \quad (5)$$

In this formula, T_{TP} is true positive (TP), which refers to the number of fake images that are correctly classified; T_{TN} is true negative (TN), which refers to the number of real images that are correctly classified; F_{FP} is False positive (FP), which refers to the number of fake images that have been misclassified; F_{FN} is false negative (FN), which refers to the number of real images that have been misclassified.

In order to evaluate the model more comprehensively, in addition to selecting the accuracy rate, this article also selects the area under roc curve (AUC) as the evaluation index in addition to the receiver operating characteristic curve (ROC). The ROC curve is based on the predicted results of the model. Sort the samples by size, use the predicted probability of each sample as the threshold value one by one in this order, calculate the false positive rate (false positive rate, F_{PR}) and the true case rate, that is, the recall rate (true positive rate, T_{PR}), and Take F_{PR} as the horizontal axis and T_{PR} as the vertical axis. Among them, the calculation formulas of F_{PR} and T_{PR} are as follows:

$$F_{PR} = \frac{F_{TP}}{T_{TN} + F_{FP}} \quad (6)$$

$$T_{PR} = \frac{T_{TP}}{F_{FN} + T_{TP}} \quad (7)$$

ROC curve can well represent the generalization performance of a model. AUC is the area under the ROC curve. The larger the AUC value, the better the performance. The calculation formula of AUC is as follows:

$$A_{AUC} = \frac{1}{2} \sum_{i=1}^m (F_{PR}^{(i+1)} - F_{PR}^{(i)}) \times (T_{PR}^{(i)} - T_{PR}^{(i+1)}) \quad (8)$$

In this formula, m is the number of samples.

V. EXPERIMENTAL RESULT

A. Analysis of Algorithm Effectiveness

After a lot of experiments with the same parameters, such as Table IV (all models uses the same preset parameters), we get the data curves as shown in Fig. 4 and Fig. 5. It can be seen from Fig. 4 that as the number of training increases, the Loss function value of the model on the validation dataset and the training dataset gradually decreases, indicating that the model's fitting ability is getting stronger and stronger, which fully illustrates the effectiveness of the experimental model in this paper.

Fig. 5 shows that as the number of model iterations increases, the accuracy of the model's classification prediction on the training set and validation set also gradually increases, indicating that the model's effect is getting better and better. It can also be seen from Fig. 5 that the accuracy of the model tends to be stable around the 160th epoch. The training accuracy of the final model can reach about 97.13%, and the accuracy of the verification data set is about 96.69%, indicating that the model is good; classification and detection results. With the increase in the number of model iterations, the model's fitting ability has been slightly improved, but due to the strong fitting ability of the deep learning model, it is prone to overfitting the training dataset. In order to avoid over-fitting, this paper adopts an early stopping strategy in the experiment. The early stopping strategy is often used in deep learning model training, that is, when the loss function value of the model does not improve for a period of time, the training of the model is terminated.

As shown in Fig. 6, it is part of our experimental results. The number on the top of each small image represents the probability that the operation result is true or false, the final prediction result and the labelled value. For example, the number of the small image in the upper right corner is (0.02, 0.98 | 1 | 1), which indicates the result of operation with DeepfakeNet model. The probability of 0.02 is true, and the probability of 0.98 is false, so the prediction result is false (1), which is consistent with the actual labelled data (also false (1)).

In order to verify the effectiveness of the model, this paper uses common models to carry out comparative experiments based on the same dataset and the same preset parameters, and the comparative results are shown in Table V. It can be seen from Table V that the model proposed in this study has better accuracy than VGG, GoogleNet, XceptionNet, ResNet, ResNeXt and so on.

TABLE IV. PRESET PARAMETERS

Params	Value
batch_size	128
epochs	300
dropout	0.8
max_lr	0.00001
sample_ratio	2.0

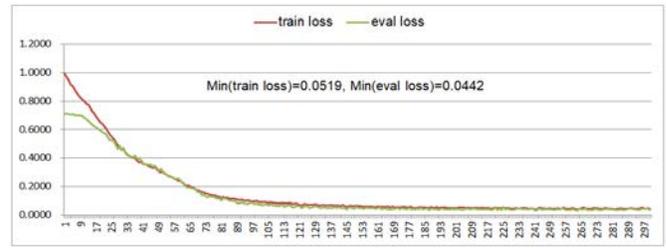


Fig. 4. Curve of Loss Value Changing with Training Times.

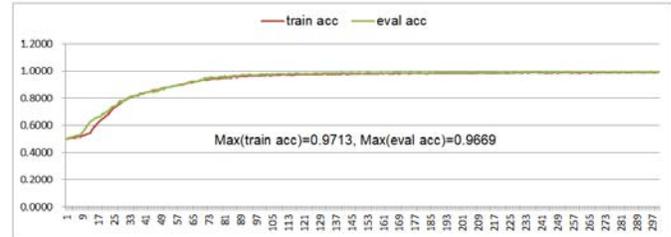


Fig. 5. The Curve of Accuracy with Training Times.



Fig. 6. Samples of Experimental Results.

TABLE V. ACCURACY COMPARISON OF EXPERIMENTAL RESULTS

Models	Accuracy (%)
VGG19	80.22
GoogleNet	88.94
XceptionNet	92.03
ResNet101	93.78
ResNeXt50	94.36
Model of this paper (DeepfakeNet)	96.69

In order to compare the performance of the models in many aspects, AUC was selected as the comparison index. The comparison results are shown in Table VI. The AUC value can not only reflect the detection effect of models, but also represent the generalization ability of the model. From Table VI, it can be seen that the model in this paper has advantages over other models, indicating that the detection effect and generalization performance of the model in this paper are better than other models.

TABLE VI. AUC VALUE COMPARISON OF EXPERIMENTAL RESULTS

Models	AUC
VGG19	0.83
GoogleNet	0.92
XceptionNet	0.92
ResNet101	0.93
ResNeXt50	0.94
Model of this paper (DeepfakeNet)	0.96

B. Analysis of Algorithm Performance

This section discusses the computational complexity of each algorithm, and compares it according to the number of floating-point operations (FLOPs) and the number of parameters for each model. Generally speaking, if the same effect is achieved, the smaller the number of FLOPs and parameters, the better. Table VII shows FLOPs and the number of parameters of some current mainstream network architectures. Although Table VII shows that GoogleNet is superior in the number of FLOPs and parameters, combined with Table V and Table VI, its accuracy (88.94) and AUC (0.85) are compared with the corresponding data of DeepfakeNet (98.69 and 0.98, respectively). Compared with ResNeXt50, the FLOPs of this model (2.05) is about 48% of it (4.27), and the parameters number of this model (10.87) is about 43% of it (25.08). Overall, performance of the network structure proposed in this study is better.

TABLE VII. COMPARISON OF NUMBER OF FLOPS AND PARAMS

Models	FLOPs	Params
VGG19	19.67 * 10 ⁹	145.77 * 10 ⁶
ResNet101	7.85 * 10 ⁹	44.6 * 10 ⁶
ResNeXt50	4.27 * 10 ⁹	25.08 * 10 ⁶
XceptionNet	3.81 * 10 ⁹	22.8 * 10 ⁶
GoogleNet	1.51 * 10 ⁹	6.13 * 10 ⁶
Model of this paper (DeepfakeNet)	2.05 * 10 ⁹	10.87 * 10 ⁶

VI. CONCLUDING REMARKS

At present, most of the popular fake face video detection algorithms use deep network to extract features. The main reason for the poor cross-dataset performance is that the deep network is easy to learn too many features in the dataset, resulting in poor generalization performance. This paper treats fake face video detection as a special image mosaic tampering detection problem, and uses image segmentation and deep residual network to predict the tampered area, reduces the impact of different training datasets, and improves the generalization performance of the detection algorithm. The experimental results on multiple popular face-swapping video dataset show that compared with other similar algorithms, the method in this paper greatly reduces the average error rate of cross-dataset detection while maintaining high accuracy in the dataset detection. The algorithm has good generality. The method in this paper can obtain good fake face video detection performance in different data sources, which shows that the

idea of improving generalization performance in this paper is general. Future improvements include expanding the scope of the training set, solving the precise detection of faces with different video quality, optimizing the network model, and developing a more complex and effective face tampering video detection network to improve usage.

ACKNOWLEDGMENT

This work was supported by Wenzhou basic scientific research project of Wenzhou Science and Technology Bureau in 2020 (No. G2020033). We would also like to thank Universiti Teknikal Malaysia Melaka for the collaboration.

REFERENCES

- [1] Z. Zhang and Q. Liu, "Detect Video Forgery by Performing Transfer Learning on Deep Neural Network," Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov., pp. 415–422, 2020, doi: 10.1007/978-3-030-32591-6_44.
- [2] D. B. V., "The state of Deepfakes: reality under attack," 2019.
- [3] "DeepFake Detection Challenge." [Online]. Available: <https://www.kaggle.com/c/deepfake-detection-challenge>.
- [4] H. Yongjian, "Deepfake Videos Detection Based on Image Segmentation with Deep Neural Networks," J. Electron. Inf. Technol., vol. 43, no. 1, pp. 162–170, Jan. 2021, doi: 10.11999/JEIT200077.
- [5] P. Korshunov and S. Marcel, "DeepFakes: a New Threat to Face Recognition? Assessment and Detection," arXiv Prepr., pp. 1–6, 2018.
- [6] A. Khodabakhsh, R. Ramachandra, K. Raja, P. Wasnik, and C. Busch, "Fake Face Detection Methods: Can They Be Generalized?," in 2018 International Conference of the Biometrics Special Interest Group, BIOSIG 2018, 2018, pp. 1–6, doi: 10.23919/BIOSIG.2018.8553251.
- [7] X. Yang, Y. Li, and S. Lyu, "Exposing Deep Fakes Using Inconsistent Head Poses," in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2019, doi: 10.1109/ICASSP.2019.8683164.
- [8] F. Matern, C. Riess, and M. Stamminger, "Exploiting visual artifacts to expose deepfakes and face manipulations," Proc. - 2019 IEEE Winter Conf. Appl. Comput. Vis. Work. WACVW 2019, pp. 83–92, 2019, doi: 10.1109/WACVW.2019.00020.
- [9] P. Korshunov and S. Marcel, "Speaker inconsistency detection in tampered video," in European Signal Processing Conference, 2018, pp. 2375–2379, doi: 10.23919/EUSIPCO.2018.8553270.
- [10] S. Agarwal, H. Farid, Y. Gu, M. He, K. Nagano, and H. Li, "Protecting world leaders against deep fakes," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Work., pp. 38–45, 2019.
- [11] G. E. KRIZHEVSKY, Alex, SUTSKEVER, Ilya, HINTON, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process. Syst., vol. 25, pp. 1097–1105, 2012.
- [12] Z. A. Simonyan K, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv Prepr., 2014.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, doi: 10.1109/CVPR.2016.90.
- [14] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017, pp. 246–253, doi: 10.1109/CVPR.2017.195.
- [15] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "MesoNet: A compact facial video forgery detection network," 2018 IEEE Int. Work. Inf. Forensics Secur., pp. 1–7, 2018, doi: 10.1109/WIFS.2018.8630761.
- [16] S. Tariq, S. Lee, H. Kim, Y. Shin, and S. S. Woo, "Detecting both machine and human created fake face images in the wild," in Proceedings of the ACM Conference on Computer and Communications Security, 2018, p. 81–87, doi: 10.1145/3267357.3267367.
- [17] B. Bayar and M. C. Stamm, "Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection," IEEE Trans. Inf. Forensics Secur., vol. 13, no. 11, pp. 2691–2706, 2018, doi: 10.1109/TIFS.2018.2825953.

- [17] D. Guera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," in Proceedings of AVSS 2018 - 2018 15th IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2019, pp. 1–6, doi: 10.1109/AVSS.2018.8639163.
- [18] S.-Y. Wang, O. Wang, R. Zhang, A. Owens, and A. A. Efros, "CNN-generated images are surprisingly easy to spot... for now," arXiv Prepr., 2019.
- [19] K. Xie, S. Girshick, R. Dollár, P. Tu, Z., & He, "Aggregated Residual Transformations for Deep Neural Networks," IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 1492–1500, 2017.
- [20] I. J. Goodfellow et al., "Generative Adversarial Networks," in Advances in Neural Information Processing Systems, 2014, pp. 2672–2680, doi: 10.3156/jsoft.29.5_177_2.
- [21] Y. Li, M.-C. Chang, H. Farid, and S. Lyu, "In actu oculi: Exposing ai generated fake face videos by detecting eye blinking," 2018 IEEE Int. Work. Inf. Forensics Secur., pp. 1–7, 2018.
- [22] Y. Li and S. Lyu, "Exposing DeepFake Videos By Detecting Face Warping Artifacts," arXiv Prepr., 2019.
- [23] J. Li, H., Li, B., Tan, S., Huang, "Detection of deep network generated images using disparities in color components," arXiv Prepr., 2018.
- [24] M. Koopman, A. M. Rodriguez, and Z. Geradts, "Detection of Deepfake Video Manipulation," 20th Irish Mach. Vis. Image Process. Conf., pp. 133–136, 2018.
- [25] and A. R. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, "Going deeper with convolutions," IEEE Comput. Soc., 2014.
- [26] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics: A large-scale video dataset for forgery detection in human faces," arXiv. 2018.
- [27] Y. Gao, Y. Hu, Z. Yu, Y. Lin, and B. Liu, "Evaluation and Comparison of Five Popular Fake Face Detection Networks," Yingyong Kexue Xuebao/Journal Appl. Sci., vol. 37, no. 5, pp. 590–608, 2019, doi: 10.3969/j.issn.0255-8297.2019.05.002.

A New Key Generation Technique based on Neural Networks for Lightweight Block Ciphers

Sohel Rana^{1*}, M. Rubaiyat Hossain Mondal²

Institute of Information and Communication Technology
Bangladesh University of Engineering and Technology
Dhaka 1205, Bangladesh

A. H. M. Shahariar Parvez³

Department of Computer Science and Engineering
Dhaka University of Engineering and Technology
Dhaka 1700, Bangladesh

Abstract—In recent years, small computing devices used in wireless sensors, radio frequency identification (RFID) tags, Internet of Things (IoT) are increasing rapidly. However, the resources and capabilities of these devices are limited. Conventional encryption ciphers are computationally expensive and not suitable for lightweight devices. Hence, research in lightweight ciphers is important. In this paper, a new key scheduling technique based on neural network (NN) is introduced for lightweight block ciphers. The proposed NN approach is based on a multilayer feedforward neural network with a single hidden layer with the concept of nonlinear activation function to satisfy the Shannon confusion properties. It is shown here that NN consisting of 4 input, 4 hidden, and 4 output neurons is the best in key scheduling process. With this architecture, 5 unique keys are generated from 64 bit input data. Nonlinear bit shuffling is applied to create enough diffusion. The 4-4-4 NN approach generates the secure keys with an avalanche effect of more than 50 percent and consumes less power and memory, thus ensuring better performance than that of the existing algorithms. In our experiments, the memory usage and execution cycle of the NN key scheduling technique are evaluated on the fair evaluation of lightweight cryptographic systems (FELICS) tool that runs on the Linux operating system. The proposed NN approach is also implemented using MATLAB 2021a to test the key sensitivity by the histogram and correlation graphs of several encrypted and decrypted images. Results also show that compared to the existing algorithms, the proposed NN-cipher algorithm has lower number of execution cycles and hence less power consumption.

Keywords—Lightweight cryptography; IoT; resource limited devices; neural network; avalanche effect; FELICS; MATLAB

I. INTRODUCTION

Lightweight cryptography [1] is that the scaled-down version of traditional cryptography which target is to provide security for devices which resource capacity is restricted. While thinking in computer communication all people want to prevent his/her message from the malicious person as well as the valid one must receive/decrypt original message easily. However, there is a clear trade-off between lightweight and security of ciphers: Hence, a decent level of security is often achieved in such sort of resource-constrained devices. In recent years, the research community has been focusing on designing cryptographic primitives which are suited to those resource-constrained devices [2]. Conventional cryptographic algorithms like RSA [3] mostly perform well in powerful devices; therefore, lightweight algorithms do not seem to be

necessary for them. Resources like read-only memory (ROM), random access memory (RAM), processing speed, and battery power are limited for resource-constrained devices like Embedded systems, radio frequency identification tags (RFID), and sensor networks. Hence, the lightweight block ciphers become essential to ensure the security of these devices. Different types of cryptographic algorithms like Advanced Encryption Standard (AES) [1], Data encryption standard (DES) [1], PRESENT [4], etc. are used for resource-constrained devices. In most resource-limited devices, lightweight block ciphers use such design architectures which will ensure enough security while keeping execution cycles as less as possible. Most of the ciphers follow the Feistel Architecture like Secure Internet of Things (SIT) [5], SIMON [6], Speck [6], etc. or by Substitution-Permutation Network like PRESENT [4], AES [7], etc. or by using both Architectures like DES, SIMON [6] to supply enough Shannon's confusion and diffusion properties in cipher text. Key scheduling in the block ciphers should perform in a secure way because the security of the ciphers depends on the secret keys which are used in every round of a block cipher. To make round keys strong, different complex number theories like modular arithmetic, prime factorization, Euclidian algorithms, etc. are applied in key generation techniques that end in hamper the performance of resource-limited devices. Good key scheduling must have two properties; randomness to generate unique keys and a high avalanche effect to ensure high key sensitivity. A single bit change in the key should change at least 50% bit in the cipher-text so that an attacker cannot easily predict a plain-text or keys through a statistical attack of encrypted message. That type of effect in cipher text is regarded as an avalanche effect. To implement a strong cipher, the avalanche effect should be considered as one of the primary design objectives.

A. Motivation

To research on cryptography is a challenging and interesting topic. Cryptography is the heart of secure data transmission. It includes complex mathematics, advanced programming, advanced number theories, etc. With the increasing usage of resource-constrained devices, lightweight block ciphers will be essential to provide security for those devices in near future. HP investigate that above seventy percent of re-source-limited devices are vulnerable to attacks [8]. There is a trade-off between the safety and performance of a low-powered small computing device. Since most of the cipher proposed to date is based on [1] complex number theory

*Corresponding Author

i.e., Modular arithmetic, Prime factorization, GCD testing algorithms, etc. As conventional ciphers use complex number theory to meet (Avalanche Effect) [1] Shannon confusion and diffusion properties so these ciphers generally become computationally expensive that hinders the performance of resource-limited devices. For that reason, it becomes challenging to implement these heavy algorithms in small computing devices for ensuring security. Hence, we focus on developing a simple and less power-consuming key generation technique using feedforward neural networks [9] (NN).

B. Contribution

This paper presents a NN-based key generation technique for lightweight block ciphers to provide the lightweight devices' security and meet the challenges of limited resource utilization. This paper puts NNs into the key scheduling process to generate strong round keys. This is done to achieve a more avalanche effect: more than 50% of the output bit is changed for a single bit change in input. This property is also called Shannon confusion. In the used NN, every bit in the input neuron is connected to all the output neuron bits. So, a change of a single bit can affect all bits in the output. It has addressed nonlinearity also since the sigmoid activation function is applied in hidden layer of proposed NN architecture. The proposed technique is computationally lightweight because NN uses simple mathematical operations like addition and multiplication to generate output. It also ensures nonlinearity as NN uses nonlinear activation functions like [10] sigmoid and Step Function. Normally this nonlinear transformation is done by S-BOX [11] like in AES [7], DES [7]. We faced few challenges to put the NN in the key generation technique. These are as follows.

- The choice of a NN to be used.
- The size of the input, hidden, and output neurons.

We evaluated 6 different feedforward NN architectures: NN4-4-4, NN4-3-4, NN4-2-4, NN4-1-4, NN4-5-4, and NN4-6-4. Among these, NN4-4-4 and NN4-3-4 perform better others. The performance of NN4-5-4 is also good; however it has higher computational complexity. We used 4-4-4 architecture as its average avalanche effect is higher than NN4-3-4, although the computational power of 4-4-4 is little more expensive than NN4-3-4. We used MATLAB 2021a to test the randomness by entropy, correlation, and histogram and key sensitivity i.e., Avalanche effect by encrypting images. Thus, the proposed algorithm ensures enough security and consumes less power.

II. BACKGROUND AND LITERATURE REVIEW

Lightweight devices have security vulnerabilities. Particularly, the devices used in IoT have security threats since these devices may remain without supervision for long hours [12]. Conventional ciphers require higher computations [13]. Most of the modern cryptographic algorithms proposed are based on complex number theory like [3] RSA, [1] ElGamal, and SPN [11] network like AES and [11] Feistel architecture like [7] DES. The primary focus of using these primitives is to create keys and ciphertext more secure by ensuring more avalanche effect i.e., more confusion and diffusion in the ciphertext. The main problem with these primitives is

computation-ally expensive, which hampers the performance of resource-limited devices if implemented.

A. NN in Cryptography

The NN is a core concept of computer science that can be implanted into cipher to achieve a higher avalanche effect in cryptography. As every neuron of input is connected to each neuron of output so a change of a single bit in input can affect all bits (neurons) of output. The NN also provides nonlinear transformation like confusion by using nonlinear activation functions like sigmoid. We can fix the output of a neuron by training the NN for different weights. Backpropagation is a good choice to train the NN. After that, a trained network with fixed weights can be implanted to cipher to make enough confusion in the ciphertext. Since NN is also a one-way transformation, reverse engineering attack is not effective.

B. Other Related Works

The authors of article [9], proposed a block cipher based on NN for cryptography. They designed an encryption algorithm based on NN to provide security in resource-constrained devices. Authors applied a [17] backpropagation network as supervised learning to train the fully connected feedforward NN. They claimed that the NN performs consistently and unconditionally throughout the encryption as well as decryption process. However, their algorithm was computationally expensive due to the extra burden of the training process of NN.

In [14], the authors evaluate the performance and security of modern [18] lightweight ciphers like TEA, HIGHT, KATAN, and KLEIN which are instigated especially in resource-constrained devices. To evaluate the performance metrics like memory and power consumption, the authors used the AtTiny45 microcontroller as a resource-limited device. Besides, they assessed the level of Shannon confusion and diffusion to testing the avalanche effect.

In [2], authors proposed a cipher for lightweight devices consisting of two core concepts of genetic algorithm, namely, two-point crossover and coin flip mutation. They also tested their proposed cipher on Fair Evaluation of Lightweight Cryptographic Systems (FELICS) to compare execution cycle and memory usages.

The author in [5] presented a symmetric cipher that combined together Feistel architecture and Substitution Permutation Network (SPN) to avail the linear and non-linear transformation in cipher text. They proposed a cipher that includes: key generation and the encryption process. The key scheduling section generates 5 unique keys by taking 64 bits master key as input from the user. After initial permutation, 64-bit input is grouped into 4 blocks each of which is 16-bit data in size. Every 4 blocks are fitted as input for f-function as shown in Fig. 1. A 4x4 matrix is used to transform the output of the f-function. Here the only source of nonlinearity is the usage of the matrix. The F-function consists of two P-Boxes; used in key scheduling is just the linear transformation. We introduced NN-function in replace of F-function to provide both nonlinearity and high avalanche affect which results in high key sensitivity.

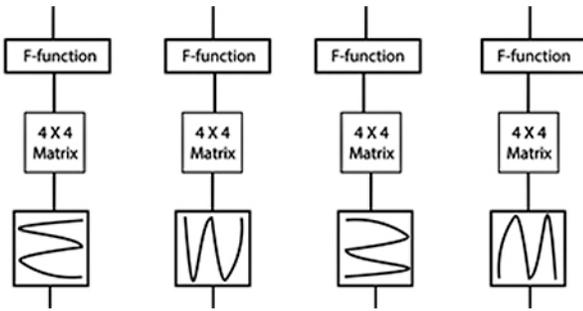


Fig. 1. 4x4 Matrix Format of Key Generation.

III. PROPOSED NN APPROACH

This section describes our proposed NN cipher approach. The proposed technique is a symmetric key block cipher based on a feedforward NN with a nonlinear activation function. The NN approach is suitable for low-powered and resource-limited devices. Using NN is to generate the keys strong enough by ensuring the avalanche effect more than 50%. We put the NN with nonlinear activation function into the key scheduling process to generate round keys used in different rounds of the encryption process. As every neuron of input is connected to each neuron of output so a change of a single bit in input can affect all bits (neurons) of output. These properties of NN meet the standard of avalanche effect that is more than 50%. We analyze the different sizes of NN to test the avalanche effect. Our experiment shows that NNs having too many or too few neurons in hidden layers than that of the input layer have less avalanche effect. We use NN consisting of 4 input neurons, 4 hidden neurons, and 4 output neurons in the key scheduling process. The size of the keys and plaintext is 64 bits.

A. Key Expansion

In a block cipher, the most fundamental component is the keys that are used to perform the encryption as well as decryption. If the key that was used to generate cipher text is compromised, the security is totally broken. Therefore, the illumination of the key should be as difficult as possible. To prevent data from different statistical attacks like chosen cipher text, chosen plain text, differential attack, etc. the sensitivity of the keys must be too high. Even if the attacker assumes the key that differs only a single bit from the original key, the result of decryption with that assumed key should be like cipher text. To ensure higher key sensitivity, the algorithm uses NN, with a nonlinear activation function for generating a key with an avalanche effect of more than 50%. Fig. 2 illustrates the process of the proposed key scheduling details.

The proposed algorithm needs 64-bit data as input to generate five unique keys for five rounds of the encryption process. These 64 bits of data are divided by 4 bits which generate 16 networks. Each network consists of 4-bit data. The proposed technique uses NN in 4 networks to create nonlinear transformation, and others are P and Q transformation to make enough diffusion in generated keys. A feedforward NN of NN 4-4-4 (4 input neurons, 4 output neurons, and 4 neurons of a hidden layer) is put on 1st, 6th, 11th, 15th networks, receipts 4 bits as input, and generates the 4-bit output with sigmoid and step activation function which is a nonlinear activation function. By using the nonlinear activation function in the hidden layer, the manufactured key is getting stronger with more confusion. Nonlinear bit shuffling is applied to create enough diffusion (linear transformation) in generated keys.

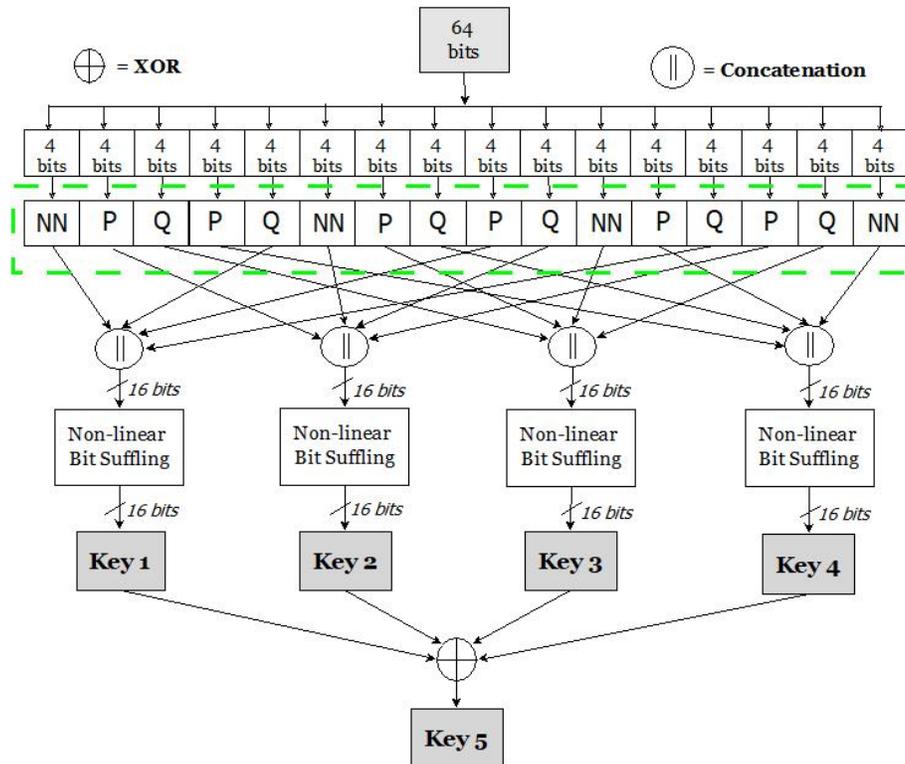


Fig. 2. NN based Key Scheduling Process.

After that the four 16-bit key (K1, K2, K3, K4) are generated, which consists of the output of four-bit networks. The next step is to apply XOR operation on every key K1, K2, K3, and K4 to generate the fifth key K5.

B. NN as a Function

In this paper, we use NN instead of the f-function [5], which is a combination of some linear transformation, i.e., permutation Box. The one-way properties of NN are more powerful to generate strong keys, and the avalanche effect of the proposed NN is more than 50 percent. This property of NN ensures the higher key sensitivity to protect different statistical cryptanalysis. Moreover, NN has less computational complexities because simple mathematical operations like addition and multiplication are used to produce output. Without NN, if anyone wants to generate the key with a higher avalanche effect, they must have to use complex number theory like prime factorization, Random numbers, and Modular arithmetic etc. For example, like RSA, large Prime numbers and modular exponential arithmetic are used to generate keys. Hence, there will be large computation power which is not acceptable for lightweight devices. In this proposed algorithm, a feedforward NN as shown in Fig. 3 is applied. The weights are trained with machine learning for a better output result. In a feedforward NN, input data travels in one direction only, passing through artificial neural nodes and exiting through output nodes. It has unidirectional forward propagation [9] but no backward propagation. Weights are static here. An activation function is fed by inputs which are multiplied by weights. The network has no cycles or loops.

Here, $X_i = [X_1, X_2, X_3, X_4]$ are four neurons of input layer that join with their corresponding weights to 4 neurons of hidden layers $H_j = [H_1, H_2, H_3, H_4]$ to generate 4 neurons of output layer $O_k = [O_1, O_2, O_3, O_4]$. The equation to calculate the each neuron of hidden layer is

$$h_j = \sum_{i=1}^n W_{ij}X_i + b_i \tag{1}$$

where h_j = hidden layer output, W_{ij} = weights, X_i = input bit and b_i =bias value.

The equation to calculate the each neuron of hidden layer is

$$O_k = \sum_{j=1}^n h_jW_{jk} + b_j \tag{2}$$

where O_k = desired output, h_j =hidden layer output, W_{ij} = weights and b_j = bias value of hidden layer.

C. Activation Function

Activation function produces nonlinearity into the output of a neuron. We used sigmoid function [10] is as activation function in the hidden layer.

$$\phi(H_j) = \frac{1}{1+e^{-H_j}} \text{ for } j = 1, 2, 3, 4 \tag{3}$$

We used step function [10] as activation function in output layer because we need either 0 or 1 as output in output layer. The output is a certain value, A_1 for the case where the input summation has a values beyond a particular threshold, and A_0 when it is less than the threshold. The perceptron values were $A_1 = 1$ and $A_0 = 0$.

$$F(O_k) = \begin{cases} 1, & O_k > 0; \\ 0, & O_k \leq 0; \end{cases} \text{ for } k = 1, 2, 3, 4 \tag{4}$$

From the truth table shown in Table I, it can be seen that the rate of bit changing is average 50% upper. This type of bit changing can make Shannon's confusion and diffusion easily. The security of the proposed key generation is standard for its image analysis because its correlation comes out in uniform patterns.

D. The Selection of Architecture NN 4-4-4

In cryptography, the keys must be generated by secure enough key scheduling algorithms. Generally, Different popular hashing algorithms like Secure Hash Algorithm (SHA), (Message Digest) MD5 etc., are used to generate rounds key to encrypt the plaintext. The most important issue in generated keys is the sensitivity of the keys. Any change in the key or the plaintext will lead to a significant change in the ciphertext so that an attacker cannot easily predict a plaintext or keys through a statistical analysis of ciphertext. This property is regarded as avalanche effect. To implement a strong cipher, avalanche effect should be considered as one of the primary design objective. The avalanche effect was identified by [1] "Shannon's property of confusion", however, the term was first mentioned by Horst Feistel. A strong encryption algorithm should always satisfy the [1] Avalanche effect > 50% criteria.

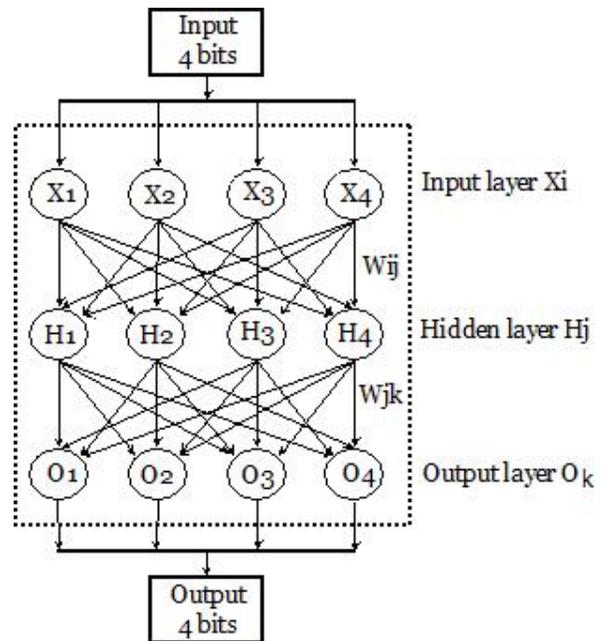


Fig. 3. Feed Forward NN (NN 4-4-4).

TABLE I. SAMPLE OUTPUT OF NN 4-4-4

Input	Output
1	1
1	0
1	0
0	1

An encryption algorithm that does not satisfies this property can favor an easy statistical attack. Considering the higher avalanche effect and lower computational power, we proposed the freed forward NN of 4 input neurons, 4 hidden neurons, and 4 output neurons for key scheduling process. As shown in Table II, we evaluate avalanche effect for different combinations of hidden neurons. We collect the outputs in python language. We do not consider all combinations of a single bit to test the avalanche effect but run the program for different combinations and put the average of collected outputs. According to the used weights, threshold values and activation functions, we got 53.125 % of avalanche effect (1-bit change) for architecture NN 4-4-4, NN 4-3-4, and NN 4-5-4 others are zeros. However, for 4-bit change, architecture NN 4-4-4 has the highest avalanche effect.

E. Computational Complexity of NN

In order to calculate the actual computational complexity of NN, it is necessary to know both the complexity for each operation and the number of operations. Assume that Computational Cost $C.C = M * 3 + A$ where M denotes the number of Multiplications and A denotes the number of addition in a NN. Here in the above equation, M is multiplied by 3 because in integer arithmetic multiplication is usually 3 times appreciably slower than addition. We do not consider the complexity of the activation function for simplicity of calculation.

TABLE II. EFFECT IN ROUND KEYS ON A SMALL CHANGE IN MASTER KEY

NN I/P - H - O/P	Affect in output for 1 bit changed in input (%)	Affect in output for 4 bit changed in input (%)	Computational cost
NN 4-4-4	53.125	59.375	32 Multiplications + 32 Additions
NN 4-3-4	53.125	56.25	24 Multiplications + 24 Additions
NN 4-2-4	00.00	00.00	16 Multiplications + 16 Additions
NN 4-1-4	00.00	00.00	8 Multiplications + 8 Additions
NN 4-5-4	53.125	56.25	40 Multiplications + 40 Additions
NN 4-6-4	00.00	00.00	48 Multiplications + 48 Additions

TABLE III. WEIGHED COMPUTATIONAL COST

NN I/P - Hidden - O/P	Computational Cost (Cycle)
NN 4-6-4	192
NN 4-5-4	160
NN 4-4-4	128
NN 4-3-4	96
NN 4-2-4	64
NN 4-1-4	32

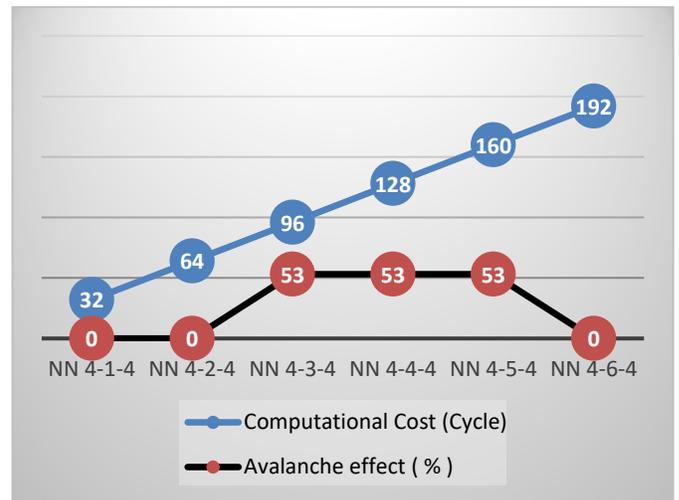


Fig. 4. Impact on Computational Power and Avalanche effect by each Increasing Hidden Layer.

Here $M = 4$ input weights * 4 inputs + 4 hidden weights * 4 output neurons = 32, similar for $A = 32$. So, $C.C = 32 * 3 + 32 = 128$. Accordingly, we calculated for all the NNs shown in Table III. The impact on computational power and avalanche effect by each increasing hidden layer are illustrated in Fig. 4. Horizontal axis represents the number of hidden layers for 4 input and output neurons.

The computational power cost increases linearly as we increase the hidden layer one by one. On the other hand, a number of architecture, namely NN 4-4-4, NN 4-3-4 and NN 4-5-4 have the avalanche effect of 53.125% (> 50%) which satisfy the standard of Shannon's property of confusion. We observed that as he hidden layers exceed the input or output neuron, avalanche effect remains the same and sometimes starts to fall. The performance of architecture with 5 hidden neurons is not improved. The optimal number of hidden [8] units should be smaller than the number of inputs. So, we selected architecture NN 4-4-4 to generate keys with high sensitivity.

F. Key Sensitivity for Ciphertext

To keep data secret from different statistical analyses like chosen ciphertext, chosen plain text, differential attack, encrypted ciphertext must change drastically (at least 50%) for the change of single bit of keys or plain text. We tabulated keys with single bit change and corresponding ciphertext for all mentioned architectures. Proposed architecture showed the avalanche effect more than 50%. All most entire avalanche effect depends on the NN in key scheduling. Table IV shows 6 pairs of cipher text for six different pair of main keys while a pair of keys differs only in single bit to each other.

G. Permutation (P and Q Table)

We used the P and Q tables as reported in the literature [5]. These tables perform linear transformations resulting in diffusion. However, we used P and Q table as per our new design of the key scheduling. The transformations made by P are shown in Table V.

TABLE IV. AVALANCHE EFFECT IN CIPHER TEXT FOR SINGLE BIT CHANGED IN KEY

NN I/P - H - O/P	Input Keys (64 bits) in hex	Cipher text(64 bits) in hex	Avalanche Effect (%)
NN 4-4-4	12 34 56 78 9a bc de fa	04 36 12 d2 97 8f 1b af	53.125
	12 34 56 78 9a bc de fb	48 d b8 55 2d 30 5b 80	
NN 4-3-4	12 34 56 78 9a bc de fa	04 36 12 d2 97 8f 1b af	53.125
	12 34 56 78 9a bc de fb	48 d b8 55 2d 30 5b 80	
NN 4-2-4	12 34 56 78 9a bc de fa	d3 29 51 da ec 8f c8 b4	00.00
	12 34 56 78 9a bc de fb	d3 29 51 da ec 8f c8 b4	
NN 4-1-4	12 34 56 78 9a bc de fa	d3 29 51 da ec 8f c8 b4	00.00
	12 34 56 78 9a bc de fb	d3 29 51 da ec 8f c8 b4	
NN 4-5-4	12 34 56 78 9a bc de fa	04 36 12 d2 97 8f 1b af	53.125
	12 34 56 78 9a bc de fb	48 d b8 55 2d 30 5b 80	
NN 4-6-4	12 34 56 78 9a bc de fa	d3 29 51 da ec 8f c8 b4	00.00
	12 34 56 78 9a bc de fb	d3 29 51 da ec 8f c8 b4	

Here, Plain text used is 0xab cd 12 34 87 65 01 35

TABLE V. P TABLE

Input key(k_i)	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Generated key(k_i)	3	F	E	0	5	4	B	C	D	A	9	6	7	8	2	1

The transformations made by Q are shown in the Table VI.

TABLE VI. Q TABLE

Input key(k_i)	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Generated key(k_i)	9	E	5	6	A	2	3	C	F	0	4	D	7	B	1	8

H. Non-Linear Bit Shuffling

We used the nonlinear bit-shuffling as reported in the literature [2]. In the nonlinear bit-shuffling, a concatenated 16-bit is transferred to each block of non-linear bit shuffling. After that, a random number is computed from the input of 16-bit data which is again logically combined with 16-bit input by performing a bitwise XOR operation. The output of the XOR operation is transferred to bit shuffling as well as to the perfect shuffling block sequentially to create enough diffusion in generated keys.

I. Encryption and Decryption Process

We used the encryption algorithm which was proposed by [2]. The encryption process consists of Feistel architecture with G-function [2] which is based on the concept of two operators of genetic algorithms: mutation and crossover. Fig. 5 illustrates the flow of operation for a single among five rounds. The encryption process takes a block of 64-bit plain messages as input at a time. The first round uses the first key K1 generated by NN based key generation technique and then K2, K3, K4, K5 for 2nd, 3rd,4th, and 5th rounds sequentially.

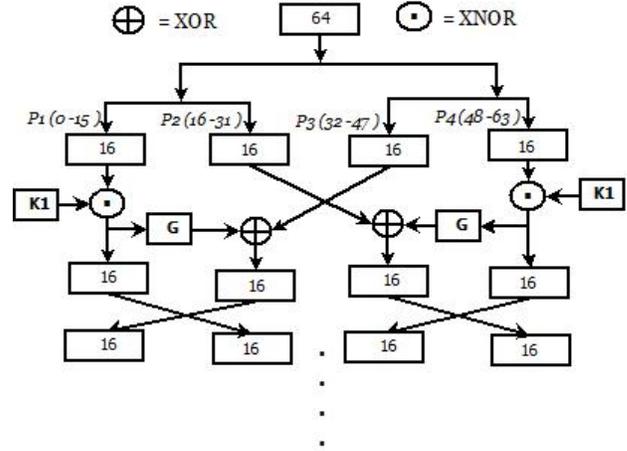


Fig. 5. A Single Round of Encryption Process.

$$RO_{i,j} = \begin{cases} Px_{i,j} \odot K_i ; j = 1 \text{ and } 4 \\ Px_{i,j+1} \oplus EG_{li} ; j = 2 \\ Px_{i,j-1} \oplus EG_{ri} ; j = 3 \end{cases} \quad (5)$$

The 64-bit message is equally split into four 16 bit segments. According to the Feistel structure, swapping, XOR, XNOR operations are performed among the split blocks to increase the avalanche effect in cipher text. An XNOR operation is performed between round key and left as well as rightmost blocks separately. Then the output of the XNOR operation is feed to G-function as input. The 4th block and output of the left G-function are again XORed and the 2nd block and output of the right G-function are XORed as well. After that, a swapping operation is executed among the 4 blocks except for the last round. The equation (5) represents the process of how a single rounds encrypt plaintext into cipher text. Finally, every four blocks are combined together to generate a block of 64-bit cipher text. The decryption process is the opposite of the encryption process. This time last key is used first.

J. G-Function

We replicated the G-function from the earlier research work. The G-function [2] is based on the concept of two operators of genetic algorithms: mutation and crossover. This function takes 16 bits as input and first, split equally into two eight-bit segments. Both two 8 bit data are transformed by using a substitution box that is called S-Box. After performing a two-point crossover to both outputs of S-Box, a coin flip mutation operation occurs. Finally, the 16-bit output is generated.

IV. EXPERIMENTAL SETUP

The proposed approach is initially written in a popular structure language C. We used CodeBlocks as IDE. The coded cipher is not dependent on any machine to be executed. We also used a benchmark tool ‘Fair Evaluation of Lightweight Cryptography Systems (FELICS) to measure memory usage and execution cycles that run on Linux Ubuntu. The FELICS tool is open access and free to install. We also implemented our proposed NN approach in MATLAB 2021a to evaluate the security strength of keys. The proposed NN approach is evaluated based on key sensitivity, execution cycles, bridge histogram, and correlation histograms. We also assessed the NN cipher in terms of memory usage and clock cycles to generate keys, cipher text, and regenerate plaintext.

A. Evaluation Parameters

The security strength of the proposed algorithm is tested based on key sensitivity, execution cycles, bridge histogram and correlation histograms. We also assessed the memory utilization and execution cycles for key generation, encryption and decryption of this algorithm. The FELICS [15] provides a command-line interface like GCC (GNU Compiler Collection) to test and build any lightweight cryptographic code. They provide documentation to facilitate the implementation as shown in Fig. 6. We can compile our implementation and test whether ours is runnable in FELICS or not. It provides three scenarios against which we can test our code.

We used FELICS to collect clock cycles required for the key generation, encryption, and decryption process of different reported ciphers along with the proposed cipher. We measure the program memory, RAM, and actual code size as well. Table VII shows the comparative results of different ciphers for AVR architecture. It can be seen that among the methods considered, the proposed NN method requires the lowest total execution cycles. Though other ciphers like HIGHT [16] have fewer cycles for key generation, the overall execution cycle of NN is fewer than that of others. The NN cipher needs the lowest RAM to execute and hence, it is memory efficient also.

TABLE VII. EXECUTION CYCLES OF CIPHERS ON AVR ARCHITECTURE

CIPHER	Device	Block size	Key size	CODE SIZE	RAM	Cycles (Key generation)	Cycles (encryption)	Cycles (decryption)
AES[7]	AVR	128	128	23090	720	3274	5423	5388
HIGHT[16]	AVR	64	128	13476	288	1412	3376	3401
LEA[14]	AVR	128	128	3700	432	4290	3723	3784
PRESENT[4]	AVR	64	80	1738	274	2570	7447	7422
RC5	AVR	64	128	20044	360	26793	4616	4652
Simon[6]	AVR	64	96	1370	188	2991	1980	1925
Speck[6]	AVR	64	96	2552	124	1509	1179	1411
SIT[5]	AVR	64	64	826	22	2130	876	851
G-cipher[2]	AVR	64	64	1228	34	1630	792	789
Proposed NN-Cipher	AVR	64	64	1228	34	1483	792	789

Fig. 7 presents bar chart comparisons among various reported ciphers with NN approach. For each ciphers, bar chart shows the status of required clock cycles to generate keys, cipher text from plaintext, plaintext from cipher text, as well as overall execution cycles. The chart shown in Fig. 7 clearly demonstrates that the NN cipher executes in less clock cycles, improving over the other reported ciphers especially than SIT [5] and G-cipher [2] which we followed most. So, the proposed NN cipher is more power efficient than that of other ciphers.

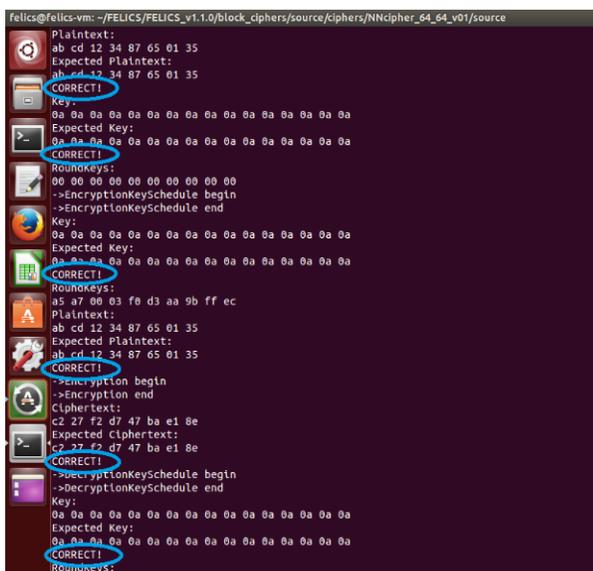


Fig. 6. Testing the Proposed Approach in FELICS.

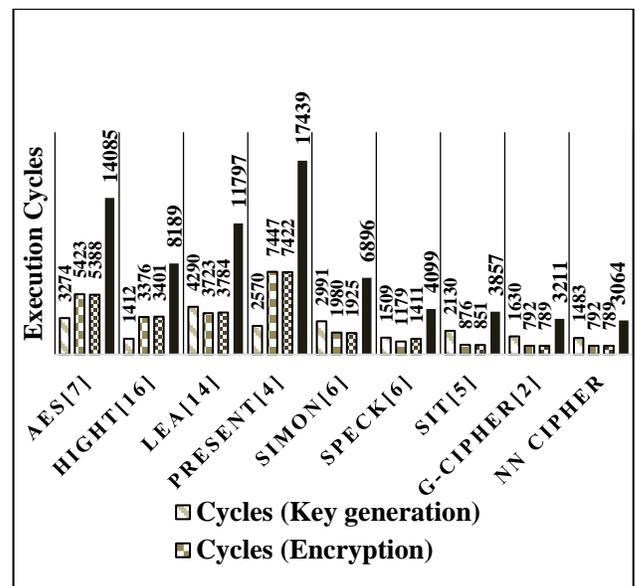


Fig. 7. Execution Cycle of Ciphers for Hardware Implementation.

B. Analysis of Key Sensitivity

The NN approach is also demonstrated in MATLAB® which encrypts an image and then decrypts the image with the correct key for a visual observation key sensitivity. After that the images are decrypted by using a wrong key with a single bit alteration from the original key. This is also a test for the avalanche effect of the keys. This is also a test for the avalanche effect of the keys. The decryption is non-recognizable if even one bit changes in the original keys. Fig. 8 shows that for NN-cipher, the encrypted images can only be decrypted with the correct key.

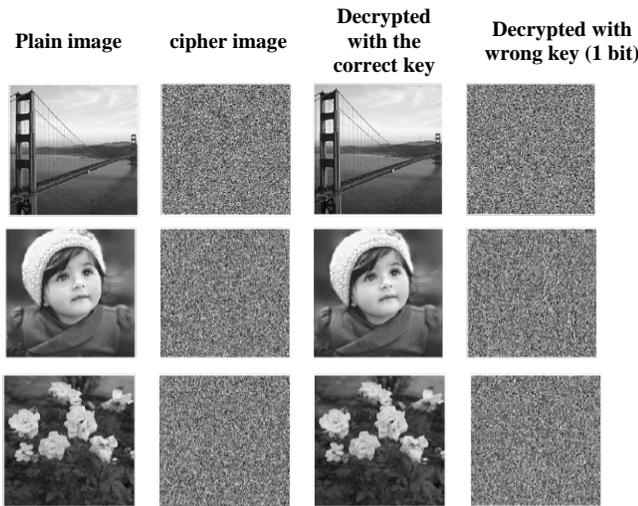


Fig. 8. Analysis of Key Sensitivity.

C. Histogram and Correlation Comparisons

Fig. 9 presents the bridge histograms of original and encrypted images. The vertical line indicates the pixel available in the image, and the horizontal axis refers to image intensity. The histogram of the image shows the uniform relationship that ensures the strength of the cipher image. So any statistical attacks will not be effective to predict the plain image from the cipher image without using the correct keys.

Fig. 10 shows the correlation graph of the considered original images and the encrypted images. The correlation graph of plain image shows linear relationship that is higher positive correlated value. However, the correlation graph of cipher image shows highly randomness that is the negative values. Hence, the negative correlation values of encrypted images indicate the strong security strength of proposed cipher.

D. Power Consumption

For calculating the total power consumed by an algorithm on a particular device, first we need the execution cycle of that algorithm. By using the following equation (6), we can compute the power consumption of an algorithm on a particular device:

$$E = I * V_{cc} * N * \tau \tag{6}$$

Where, V_{cc} denotes the operating voltage and I indicates operating current used up for T seconds (unit in Ampere). τ refers to the clock period as well as N indicates the required number of execution cycle. If f be the operating frequency of

the particular device in Hertz then we can calculate the time period of the particular device that is $\tau = 1/f \text{ sec/cycle}$.

According to the absolute maximum rating (AMR) of dataset, usually maximum operating voltage of [19] Atmel Atmega128 is 5V, Maximum current is 40mA and operates at 16 MHz. Fig. 11 demonstrates the of energy consumption of different reported ciphers along with the proposed NN approach for a single block of data. The bat chart shows that NN approach consume less power than that of others.

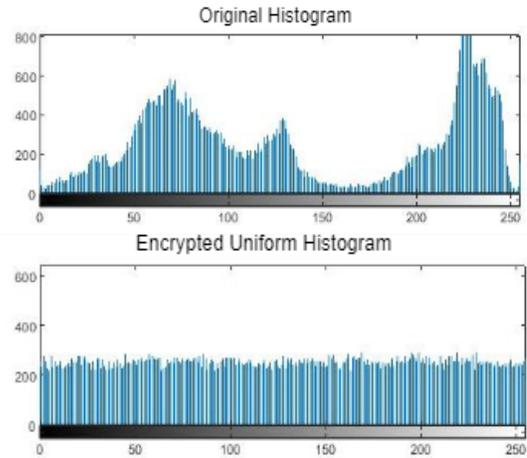


Fig. 9. Bridge Histogram.

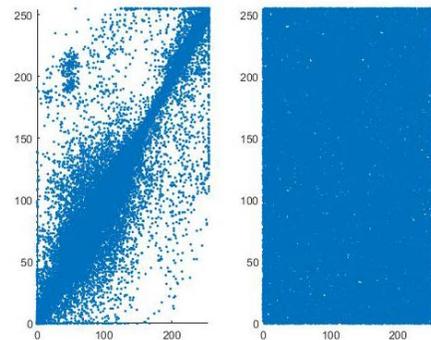


Fig. 10. Bridge Correlations for Encrypted and Decrypted Image.

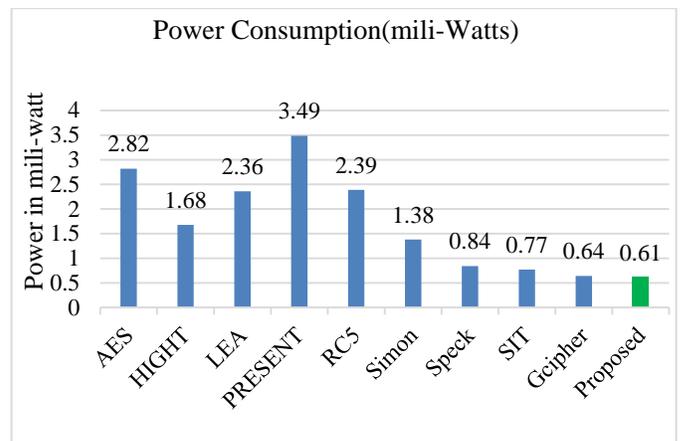


Fig. 11. Energy Consumption Comparison of Ciphers.

V. CONCLUSION AND FUTURE WORK

The security, as well as performance of resource-limited devices, is an important issue. For this purpose, a lightweight cryptographic algorithm using NN is proposed in this paper. This NN-cipher algorithm has lower key generation cycles and less power consumption than the existing ciphers. The bridge histogram and the bridge correlation plot indicate that the NN-cipher can reliably encrypt images. Moreover, the key sensitivity results indicate that for NN-ciphers, encrypted images can only be successfully decrypted using the actual key. Hence, the proposed cipher will be an excellent solution of security for those devices that are resource-limited. Besides, We intend to perform more mathematical analysis as well as hardware implementation on our proposed algorithm to investigate its security strength as future work. More evaluation metrics like the randomness of generated keys using the Chi-square test can be considered for further evaluation.

REFERENCES

- [1] William Stallings "Cryptography and Network Security Principles and Practices", Fourth Edition, Publisher: Prentice Hall, November 16, 2005.
- [2] Sohel Rana, Saddam Hossain, Hasan Imam Shoun and Dr. Mohammad Abul Kashem, "An Effective Lightweight Cryptographic Algorithm to Secure Resource-Constrained Devices" International Journal of Advanced Computer Science and Applications (IJACSA), 9(11), 2018. DOI: 10.14569/IJACSA.2018.091137.
- [3] Ekhlas Abbas Albahrani, Tayseer Karam Alshekly . "Image Cipher System Based On RSA And Chaotic Maps, Hindawi, Security and Communication Networks." Security and Communication Networks, Article ID:5586959, pp. 18 Pages, <https://doi.org/10.1155/2021/5586959> (2021).
- [4] Kostas Papapagiannopoulos. "High throughput in slices: the case of PRESENT, PRINCE and KATAN64 ciphers". Radboud University Nijmegen, Department of Digital Security. 2016.
- [5] Muhammad Usman, Irfan Ahmed, M. Imran Aslam, Shujaat Khan and Usman Ali Shah. "SIT: A Lightweight Encryption Algorithm for Secure Internet of Things" Iqra University, Defence View and Department of Electronic Engineering (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 1, 2017.
- [6] Tomer Ashur, Atul Luykx. "An Account of the ISO/IEC Standardization of the Simon and Speck Block Cipher Families." Security of Ubiquitous Computing Systems. Edited by © 2020 Springer Nature Switzerland AG. Part of Springer Nature, Vol License CC BY 4.0, pp. 63-78 Switzerland, DOI: 10.1007/978-3-030-10591-4_4 2021.
- [7] Biao Xing¹, DanDan Wang¹, Yongquan Yang¹, Zhiqiang Wei², Jiajing Wu¹, Cuihua He . "Accelerating DES and AES Algorithms for a Heterogeneous Many-core Processor." International Journal of Parallel Programming (2021) , Int., pp. 49:463–486 , <https://doi.org/10.1007/s10766-021-00692-4>. (2021).
- [8] S. A. Kumar, T. Vealey, and H. Srivastava, "Security in Internet of Things: Challenges, solutions and future directions", in 2016 49th Hawaii International Conference on System Sciences (HICSS), IEEE, 2016, pp.5772-5781.
- [9] Eva Volna , Martin Kotyba, Vaclav Kocian, Michal Janosek, CRYPTOGRAPHY BASED ON NEURAL NETWORK, Proceedings 26th European Conference on Model-ing and Simulation ©CMS Klaus G. Troitzsch, Michael Möhring.
- [10] Chigozie Enyinna Nwankpa, Winifred Ijomah, Anthony Gachagan, And Stephen Mar-shall, "Performance Analysis Of Various Activation Functions In Artificial Neural Net-works", Journal Of Physics Conference Series 1237:022030, June 2019; DOI: 10.1088/1742-6596/1237/2/022030.
- [11] Sohel Rana, Wadud, Ali Azgar, Dr. M Abul Kashem, "A Survey Paper of Lightweight Block Ciphers Based on Their Different Design Architectures and Performance Metrics" International Journal of Computer Engineering and Information Technology, June 2019, Volume 11, Issue 6.
- [12] P. Wang, Professor S. Chaudhry, S. Li, T. Tryfonas and H. Li, "The internet of things: a security point of view", Internet Research, vol. 26, no. 2, pp. 337-359, 2016.
- [13] S. Wang, Z. Zhang, Z. Ye, X. Wang, X. Lin, and S. Chen, "Application of environmental internet of things on water quality management of urban scenic river", International Journal of Sustainable Development & World Ecology, vol. 20, no3, pp. 216-222, 2013.
- [14] Vikash Kumar Jha, "Cryptanalysis of Lightweight Block Ciphers" Aalto University School of Science Degree Programme of Computer Science and Engineering, Master's Thesis, November 18, 2011.
- [15] D. Dinu, A. Biryukov, J. Großschädl, D. Khovratovich, Y. L. Corre, L. Perrin, "FELICS – Fair Evaluation of Lightweight Cryptographic Systems", University of Luxembourg, July 2015.
- [16] Bohun Kim, Junghoon Cho, Byungjun Choi, Jongsun Park, and Hwajeong Seo Hindawi. "Compact Implementations of HIGHT Block Cipher on IoT Platforms. Hindawi, Security and Communication Networks, Volume 2019, Article ID 5323578.
- [17] Tope Komal, Rane Ashutosh, Rahate Roshan, S.M. Nalawade, Encryption and Decryption using Artificial Neural Network, International Advanced Research Journal in Science, Engineering and Technology, Vol. 2, Issue 4, April 2015.
- [18] Kerry A. McKay, Larry Bassham, Meltem Sönmez Turan, Nicky Mouha. Report on Lightweight Cryptography, National Institute of Standards and Technology, USA, March 2017.
- [19] Utsav Banerjee, Lisa Ho, and Skanda Koppula. "Power-Based Side-Channel Attack for AES Key Extraction on the ATmega328 Microcontroller". Massachusetts Institute of Technology. ResearchGate. December 201.

Fully Automated Ontology Increment's User Guide Generation

Google AliceBot and AIML based Algorithmic Approach

Kaneeka Vidanage¹, Rosmayati Mohamad², Noor Maizura Mohamad Noor³, Zuriana Abu Bakar⁴
Faculty of Ocean Engineering Technology & Informatics, University Malaysia Terengganu (UMT), Malaysia^{1,2,3,4}
Faculty of Computing, Department of Data Science, NSBM Green University, Sri Lanka¹

Abstract—This research focuses on a domain and schema independent user-guide generation for ontology increments. Having a user guide or a catalogue/manual is vital for quick and effective knowledge dissemination. If a user guide can be generated for an ontology as well, there could be ample advantages. Stakeholders can scan across the user guide of the ontology and verify the eligibility of it, against the intended purposes. Additionally, this could be useful in ontology's version management requisites and knowledge verification requirements as well. Even though, ontology construction being iterative and incremental operational, there will be several intermediate versions before it reaches to the fine-tuned final version. Therefore, manual user guide creation will be a tedious and impossible operation. Consequently, this research focuses on a novel algorithmic approach to domain and schema independent ontology verbalization. A special algorithm is created to alter the functionality of Google's AliceBot to work as a verbalizer, instead of a chatterbot. Artificial Intelligent Modelling Language (AIML) technology is utilized to create the templates for the ontology specific knowledge embeddings. This entire process is fully automated via the proposed novel algorithm, which is a key contribution of this research. Eventually, the generated user guide generation tool is evaluated against three different domains with the involvement of fifteen stakeholders and 82% of averaged acceptance has been yielded.

Keywords—AliceBot; artificial intelligent modelling language; ontologist; verbalizing

I. INTRODUCTION

Ontologies are recognized as domain rich conceptualizations [1] which are both machine and human-readable [2-3]. Further, as of its enriched ability on conceptualization, it's ideal to encode specialized human knowledge [4]. Subsequently, encoded knowledge will be machine-readable accomplishing endless domain-specific reasoning and knowledge representation necessities. Because of those unique features, the popularity of ontology-based applications escalated drastically. Therefore, presently, there're thousands of applied ontologies developed in numerous domains such as biology, agriculture, bioinformatics, law, management, etc. [5-6]. Applied ontologies are used to overcome issues coming from non-computing domains and by using the aforementioned benefits of the ontologies, most of those issues can be effectively resolved [30]. Construction of an accurately defined applied ontology is a complex process, which requires both ontologists and domain specialists to work hand-in-hand with mutual understanding throughout the entire

process. As a methodical workflow to fulfil the effective bridging of domain specialists and ontologists, "collaborative ontology engineering" has been emerged out as a separate niche under the umbrella of ontology engineering [7-9].

One of the crucial necessities to be fulfilled in collaborative ontology engineering is the proper glueing in-between domain specialists and ontologists [10]. Unless effective participation and collective contribution reaching towards an error-free applied ontology will not be realistic. Researchers have mentioned, ontologists need to have reasonable insight on the domain to be modelled and vice-versa domain specialists should have a sound understanding of the essential basics associated with semantic web and knowledge modelling. Once this state is achieved only, collective, and effective participation of both parties can be expected, leading towards the construction of an error-free applied ontology schemata [11-12].

Even though, there is a critical bottleneck caused due to shortage or illiteracy of comprehension, on semantic concepts, experienced by non-computing domain specialists such as lawyers, medical doctors, agricultural specialists, etc. [13-15].

Ontology construction, being a complex, iterative and incremental task, it's expected at the end of each iteration, domain specialists should cross-reference and verify, that the knowledge provided by them, are accurately and consistently modelled, and embedded to the ontology by the ontologists [16-17]. Then the errors located can be corrected then and there, without waiting until it reaches to complex conceptual flaws resulting in an erroneous schema. In accomplishing this requirement, lack of or no literacy in semantic concepts is a strong hindrance. Because to properly understand an ontological taxonomy defined, the user should have reasonable knowledge associated with basic object-oriented concepts such as inheritance and semantic concepts such as triple concept, data properties, object properties, disjoint classes, symmetric classes and, the concept of individuals, etc. Though the person understands those aspects, the next step is writing an appropriate SPARQL or SQWRL query to verify the accuracy of the knowledge embeddings. Properly understanding the schemata of the ontology and along with the use of accurate syntaxes, forming up an accurate SPARQL query is a challenging task, even for a computer scientist at once. Therefore, obviously, it's an unrealistic goal to be expected from a non-computing domain specialist like a medical doctor, business manager or lawyer. Even if the SPARQL query is

written and executed, results will be returned as a triple with URLs and pre-processing of those are required to derive the answers in plain English, where this is also could not be a feasible task based on the competency level of a non-computing domain specialist [13-15].

All these obstacles are constraining the domain specialists' involvement, in accomplishing knowledge verification necessities, which is a crucial step in collaborative ontology engineering [9,18]. On the other hand, the creation of a catalogue (i.e., user guide) for the ontology's structure is very important for its latter maintenance and knowledge diffusion requirements [21-22]. However, manually fulfilling this task could escalate the workload of the ontologists. Further, lately, if any alteration occurred in the schemata, the entire catalogue needs to be re-written or updated accordingly. This is going to be a highly effort-consuming and tedious task on the shoulders of the ontologists [21-22].

This research is focusing on a fully automated mechanism to verbalize (i.e., output the knowledge encoded in the ontology, in its natural language form) the entire ontology, despite its domain or the schemata [19-20]. This will resolve the technological challenges non-computing domain specialists need to face and it will revoke the burden of manual user guide creation efforts from the ontologists' workload as well. Therefore, both domain specialists and ontologists are benefited from this novel contribution. The key contributions of the proposed technique against the existing mechanisms are:

- The proposed technique can work with any domain → Domain independent.
- The proposed technique can work with any schema → Schema independent.
- No external / manual configuration requirements.
- Converts semantic contents in the ontology increments into layman understandable English.

II. LITERATURE REVIEW

In the process of collaborative ontology engineering, the collective opinion of both the ontologists and domain specialists are very vital. There needs to be a proper workflow to accomplish the collective opinion derivation requirements [23-24]. Otherwise, in a collaborative environment, multiple people will raise multiple viewpoints and try to stick to their perspectives. This will ultimately lead to the issue of "Tragedies of Commons" [25-26]. Hence, in Shneiderman's "Information Seeking Mantra" concept, Shneiderman has pointed a proper workflow to methodically integrate the dispersed viewpoints of the stakeholders to reach towards an overall collective insight, at the end [27]. Unfortunately, the idea of Shneiderman's "Information Seeking Mantra" requires a specialist tool support to fulfil its workflow steps. Information Seeking Mantra concept first requests the stakeholders to get an overall idea about the problem of concern. In accomplishing this request, Shneiderman suggests the use of both visualization and verbalization tool support. Next, zoom towards the required information only. Both verbalization and specially defined visualization techniques can fulfil the second step's requirements as well. The third and

fourth steps are focusing on, filtering unnecessary information and look for information on demand, respectively [27-29]. As the outcome of this research, though a special prototype is proposed to address all requirements of Shneiderman's "Information Seeking Mantra", this paper's scope is constrained to discuss its verbalization feature only, to manage the scope of this paper.

Technically, verbalization is defined as the process of translating axioms defined in ontology to natural language [19-20]. Most of the existing verbalization systems rely on the complex Natural Language Generation (NLG) pipeline to convert axioms into natural language [31]. This is a complex, technological pipeline where all the phases need to be accurately fulfilled, to get an understandable natural language output. Namely, those steps are defined as content selection step, discourse planning step, lexicalization step, aggregation step, generation of referring expressions and finally linguistic realization step [32].

Among all those steps, the discourse planning step is very vital to achieve coherent verbalization output. The discourse planning step utilizes the 'Rhetoric Structure Theory (RST)' for the coherent organization of the text [65]. RST is based on two main conceptions as nucleus and satellite. The nucleus represents the significant axioms associated with the considered domain and the satellite represents the associative properties linked with the nucleus which are required to elaborate the nucleus [65]. Therefore, if the identification of the nucleus and satellite didn't occur in a domain-specific manner, it will adversely affect the clarity of the verbalized contents [31,33]. For that reason, there is a manual phase with the domain specialist and the ontology engineer to properly assign weights to the axioms defined in the domain considered. Afterwards, with the help of the pre-defined rule sets, it will automate the RST, assuring appropriate discourse planning, leading towards accurate and coherent verbalization.

The problem that arises here is the inability of using the same verbalizer for any other domain. Complex prior configurations, which is referred to as portal configuration is a must. This makes a verbalization ready framework to become domain-dependent always [32]. This is a key limitation associated with the existing verbalizing techniques.

The next restriction is the necessity of annotations to enrich the semantic realization of the concepts in the ontology. Again, this request additional effort from the ontology engineers and in most cases eligible foundational de-facto standard meta-models are (i.e. Dublin Core, FOAF) needed to be incorporated into the ontology. Because the majority of the existing verbalization frameworks are configured to link with the pre-defined annotated endpoints of those de-facto standard meta-models only. This poses an additional overload to the ontologists, and it acts as a modelling restrictor also [34-35]. Therefore, the free will of the ontologists and domain specialists are restrained, as they need to plan everything in a manner to suit up with the de-facto standard meta-models.

As the final disadvantage, it can be pointed, that most of the existing verbalizers produce Control Natural Language (CNL) which is resembling to assembly language and it's not colloquial English that can be understood by laymen.

Therefore, another Natural Language Processing (NLP) layer must be introduced to overcome the barrier of converting technical English constructs to its colloquial format, which will be another processing overhead. One of the main causes for this is, the existing evolution of verbalizers have evolved up to the level of Attempt to Control English (ACE), which is a form of Control Natural Language (CNL). In CNL, verbalizers attempt to extract the triple formulations in the ontology and to exactly covert them into the English language, where the contextualized connectivity and colloquialism will be lost [36-38].

TABLE I. EXISTING SYSTEM ANALYSIS

Tool	Restrictions / Limitations
NaturalOWL [31, 52]	Extensive domain-specific configuration efforts.
LODE [53]	Ontology increment must be published on the web and the accessibility url must be according to the cool_URI format. XSLT script of the LODE is configured to work with only standardized metamodels like FOAF, Dublin-Core, etc. This acts as a modelling restrictor.
SWAT Tools [54]	Extensive redundancies in the HTML link sequences generated for the ontology. Split information problem with lots of dispersed information here and there. Make the role of the domain specialist very difficult.
MIKAT [55]	Verbalizer is statically attached to the breast cancer domain. It cannot work with any other domain. Fully domain-dependent.

TABLE II. EXISTING ALGORITHM ANALYSIS ACCORDING

Algorithm	Deficiencies
CNL - ER Algorithm [56]	Works only for entity relationship diagram's verbalizations. Does not work for ontologies.
Contextual Verbalization Algorithm [36]	On-demand verbalization only. Generated results cannot be saved for later use. Additionally, the verbalized outputs are primitive and restricted to the level of ACE.
Semantic Expressions Algorithm [57]	The output is not in the natural language. It's in System Verilog format. This doesn't cater to the requirement to be fulfilled.
FST Verbalizer Algorithm [58]	Incapable to verbalize an ontology
Variable Verbalization Algorithm [59]	Based on domain-specific training of a machine learning model. Doesn't address the domain independence requirement. It's not possible to organize a training dataset, without fixing the domain.
Semantic Refinement Algorithm [60]	Extensive domain-specific configuration effort.

That's the main reason for the verbalized output to look very primitive and the flow seems inconvenient to interpret by the end-users. Most of the deficiencies associated with verbalizers are schema and domain-specific [39-41]. Among those, some of the general deficiencies are reviewed above. Hence, it can be easily concluded, through the existing verbalization mechanisms, the afore-mentioned research gaps of domain-dependence, excessive human involvement associated with configurations and CNL based less colloquialism are not properly resolved. Hence, the emphasis of this research is to propose a novel approach to overcome the aforesaid shortages.

Table I contains a comparison of famous existing verbalization tools, their limitations and why they cannot resolve the issue of domain and schema independent verbalization by producing a colloquial user guide.

Further, several existing verbalization algorithms are also reviewed to recognize their deficiencies. Table II contain the details of verbalizer algorithmic analysis.

Therefore, according to the discussion conducted in the literature review section, it's apparent there is a research gap to be resolved. The following sections of the paper discuss the steps followed to fulfil the recognized research gap.

III. METHODOLOGY

After completion of an intense and extensive systematic literature assessment [42], it's concluded that the aforesaid research gaps are still not been resolved properly.

Subsequently, the blend of the think-a-loud protocol [43] and systems thinking [44] notions are used to collectively brainstorm on the problem of concern and ultimately AIML (Artificial Intelligence-based modelling/mark-up language) is selected to implement the proposed solution, as AIML is recognized as an ideal technology for creating natural language software agents with an XML dialectic [44-48].

Table III contains the analysis results of the technology review conducted.

As per the collective brainstorming results logged in Table III, it was determined AIML is the ideal technology platform to be utilized to resolve the research problem. Because it has broad external integration support and no domain-related training datasets are required to train a domain-specific model. Hence, AIML ideally matches up with the domain and schema independence requirement.

Design science research methodology is selected as of the investigative nature of this research [47]. Implemented version of the prototype is quantitatively and qualitatively assessed on its functionality Fig. 1 exposes the application workflow of design science research methodology's operation for this research.

The importance of the research problem being investigated and its timely relevance was already justified via the literature review results studied. Additionally, a comprehensive literature analysis was conducted again on existing verbalization tools and algorithms to explicitly justify the deficiencies unresolved. Consequently, it was recognized:

- a) Domain and schema independent ontology verbalization tool with.
- b) Zero configuration effort.
- c) For the fully automated user guide generation of the iterative and incremental ontology increments.

as the consolidated research, objective to be accomplished from this research.

Henceforth, via an adequate amount of collective brainstorming, the following algorithm was designed. The proposed algorithm comprises three main operational phases as depicted in Fig. 2.

TABLE III. TECHNOLOGY SELECTION

Technology	Remark
DialogFlow [61]	<ul style="list-style-type: none"> -The training requirement is a dire constraint that needs human involvement for different contexts (done via manually typing potential phrases for the training) -The basis is on Machine Learning & domain-specific training → Makes it a domain-dependent solution, as a dataset will be domain-specific -A regular expression for pattern matching support is not available -Advanced features are not freely available -Can integrate with only one webhook – interaction where external knowledge bases are restricted.
RAZA[63]	<ul style="list-style-type: none"> -Write domain-specific stories – Manual training task. → Makes it a domain-dependent solution, as a dataset will be domain-specific -Training the dialogue model – Expose the domain-specific user stories for training purposes -Complex and a challenging learning curve before the usage -Memory hungry technology – which will slow the entire system performance - External integration assistance is very minimal. -Overfitting and underfitting issues.
IBM Watson [62]	<ul style="list-style-type: none"> -Integration with third party resources is difficult -Steep learning curve -Costly
AIML / ALICE [45-48]	<ul style="list-style-type: none"> -No datasets required for training purposes -Intelligence is extracted from the knowledge scripts – can be auto-generated via axioms extraction from the ontology -Manual integrations and expansions are also supported -ALICE contains a robust collection of AIML scripts to make the bot more intelligent -Freely available -Easy to use -A lot of potential for external integrations -Stimulus-response model can be used to organize knowledge -A regular expression for pattern matching support is also available

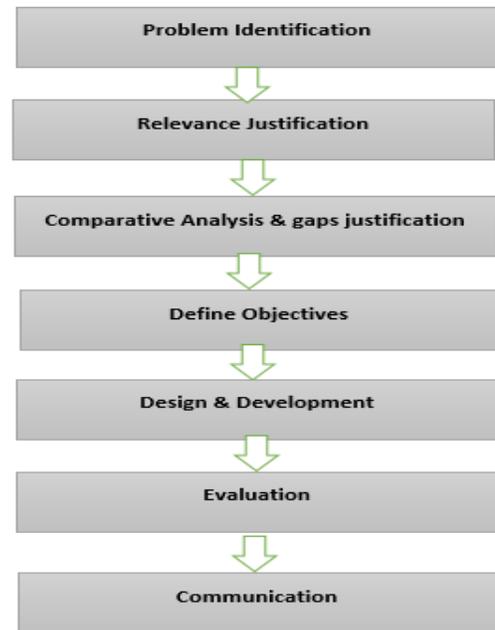


Fig. 1. Research Process.

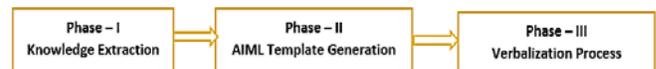


Fig. 2. Algorithm Phases.

Phase-I - [Knowledge Extraction]

Start
Upload RDF / OWL version of the ontology increment to be verbalized.

Check for the format as RDF or OWL.
Trigger format-specific knowledge extraction logic.

While [Until EOF == TRUE]

Extract class information

Extract data properties

Extract object properties

Extract class-specific individuals (if existing)

Stow them appropriately in different relations of the

RDBMS.

End While

Ontology increment to be verbalized can be directly uploaded to phase-I of the algorithm. Phase-I of the algorithm contains code snippets for knowledge extraction from both RDF(Resource Description Framework) and OWL (Web Ontology Language) formats. Henceforth, the extracted knowledge elements will be separately stowed in the database relations as classes, data properties, object properties etc. Implementation of phase-I of the algorithm can be depicted as in Fig. 3 code snippet.

```
if (fileType.equals("rdf"))
{
String namespace=modelLoader(path);// Derive main namespace of the Knowledge Model
Document doc=XMLDocumentLoader(path);// Initialize XML dom parser
individualExtractorRDF(doc,path,namespace);// Extract Individual By Individual
}
else if (fileType.equals("owl")) // If the file extension is OWL
{
    OWLOntologyManager manager = OWLManager.createOWLOntologyManager();
    File file = new File(path);
    OWLOntology ontology = manager.loadOntologyFromOntologyDocument(file);
    individualExtractorOWL(ontology,path);
}
```

Fig. 3. Knowledge Extraction Code Snippet.

Phase-II - [Auto-generation of the ontology specific AIML template]

The main purpose of this phase is ontology increment specific AIML template generation. This can be identified as a critical contribution to this research. Reasons for the choice of AIML technology is already elaborated in Table III. The proposed pseudo-code for the phase-II operation can be listed as mentioned below:

Start

Sequentially extract stored semantic elements from the database relations.

Stow them in individually created ArrayLists

Initiate a blank baseline AIML template (Fig. 4)

```
<xml>
<aiml>
  <category>
    Extract elements from one specific ArrayList (i.e., classes).
    <pattern>
      Derive the name of the semantic element type stowed in the selected ArrayList
    <template>
      I=0
      While (I >= ArrayList. Size ())
      Extract specified semantic element contents from the selected ArrayList
      Continue the appending process until reaching the end of the ArrayList size
      I++
      End While
    </template>
  </category>
  Repeat the same process for all semantic elements residing inside all the ArrayLists.
</aiml>
</xml>
End
```

```
<?xml version="1.0" encoding="UTF-8"?>
<aiml>
  <category>
    <pattern>GENERAL</pattern>
    <template>What is an ontology ? :-
      It's defined as a domain rich conceptualisation, which is a knowledge model, comprises of variety of domain related classes
      What is a class ? :-
      A class is a structural building block, representing an important perspective associated with the domain considered.
      There are numerous of relationship types available in linking one class with another.</template>
  </category>
  <category>
    <pattern>CLASS</pattern>
    <template>placeholder_1</template>
  </category>
  <category>
    <pattern>INHERITENCE_GENERAL</pattern>
    <template>Inheritance is much like parent-child feature resemblance.
      There exists parent classes and child classes. Child class inherits, attributes and behaviours from it's parent class.</template>
  </category>
```

Fig. 4. Snapshot of the Baseline AIML Template.

As depicted in Fig. 4, the generalistic baseline AIML template structure's placeholder contents will be filled by the information extracted from phase-II of the algorithm. Subsequently, the generalized baseline AIML template will be ontology increment specific. This will be done in a fully automated manner by phase-II of the algorithm.

Phase-III - [Verbalization Process]

The main purpose of phase -III is for the generation of the verbalized user guide. The pseudocode operation of the phase-III is as visible below.

Start

Load AliceBot Engine (Fig. 5)

Submit a class-specific semantic structure name (i.e. Data Properties of the Student Class) as a request object to the AliceBot Engine.

AliceBot Engine will traverse across the autogenerated AIML template and seek the most matching category

Locate for class-specific template contents

Derive the class specific<template> contents residing inside the matching category.

Cross-references the extracted contents against the values of the class-specific attributes via executing a parameterized database query.

Stow the extracted AIML pattern sequences and class-specific verbalized contents in an ArrayList.

Repeat the same process and stow all statements to be verbalized inside verbalizedrepositArr ArrayList.

K=0

While [K >= verbalizedrepositArr.size()]

Iterate across the ArrayList contents.

iText PDF Report Plugin (

verbalizedrepositArr.get(K).toString())

K++

End While

Generate PDF version of the verbalized user guide for the specific ontology increment

End

AliceBot is a Google chatterbot, developed on the foundations of AIML technology. With the help of the aforementioned algorithm designed, the operation of the chatterbot is altered to an automated verbalizer, which is a significant research contribution.

As a means of enhancing the validity of the experiment conducted, the same verbalized user guide generation process is conducted for three different domains. Those are on COVID-19 ontology increment, aquaculture ontology increment and the criminal law ontology increment. Snapshots of the taxonomical structures of those ontology increments are included in the appendix of this paper.

The proposed evaluation workflow is depicted in Fig. 8. The proposed evaluation framework utilized in this research can be pointed out as another contribution to this research. Both quantitative and qualitative emphasis were blended in this proposed framework.

First, the operationalization step was carried out. In there several open-ended questions were compiled against the research's objectives. This mapping of the questions in the questionnaire with the research's objective is the process of operationalization [64]. This will make sure the responses collected via the questions in the questionnaire are very much relevant and matching across with the requirements of the research. The list of open-ended questions defined for this evaluation is as mentioned below:

- 1) Have you been notified about the existing verbalization mechanisms?
- 2) In contrast to those, have you identified any positive capabilities of the proposed mechanism?
- 3) Do you think it will facilitate the role of ontology increment verification?
- 4) Can you elaborate on how it will facilitate the role of the inspectors?
- 5) What are the deficiencies you identified in the proposed verbalization mechanism?

In the pre-warm-up setup, all stakeholders were exposed to a specially created synoptic video clip about the research conducted. This phase acts as a retrospect and summarizes the significant aspects of the research conducted to all the stakeholders involved in the evaluation. This was done, before the official commencement of the evaluation process as it will resolve all unclear areas associated with this experimental setup.

In the controlled interview sessions, a face-to-face interview series were conducted with nine domain specialists belonging to the criminal law, COVID-19 and aquaculture domains and six ontologists. The above listed five questions were the main basis for the controlling of the interview sessions with the fifteen stakeholders. All controlled interview sessions were video recorded to facilitate later interpreting requirements. The recording was done, via getting prior approval and consent from all the involved participants and it was utilized only for research purposes and not for any other personnel gains.

10	20	30	40	50	60	70	80	90	100
Very Poor		Fairly OK – Major Revisions			Good & Acceptable – Minor Revisions			Exceptional	

Fig. 9. Close-Ended Response Grid.

During the thematic extraction phase, all recorded interviews were transcribed into a textual format. Henceforth, those were iteratively examined for several turns by the involved research staff. All information gathered through the repetitive analysis were segregated into a few generalistic themes. At the point of the analysis, initially at a drastic rate new themes started to emerge out, by the time of reaching the ninth transcription, a saturation of the themes were noted, where the same themes commenced repeating again and again. This characteristic was recognized as reaching the saturation state of the interview findings.

The theme extraction allowed to recognize of mostly insightful areas of the research. It was not feasible to collect all important opinions only via quantitative terms. Therefore, the qualitative phase enforced through controlled interview sessions created the opportunity to recognize significant user insights.

Henceforth, a close-ended question series was compiled to gather more insights on the located themes. This allows us to derive attention to details on the specific aspects with a numerical emphasis. A special rating grid was utilized to extract stakeholder opinions as depicted in Fig. 9.

Following five questions were provided in a close-ended format and requested to rate the opinions:

- 1) Natural English verbalization of the technical semantics of the ontology increment is accurate.
- 2) Acts as a manual / user guide for the ontology increment, enforcing offline usage as well.
- 3) Boosts comprehension when concepts are verbalized in layman terms.
- 4) Domain and schema independent, configuration free operation.
- 5) How would you rate the verbalization assistance provided by the tool support?

Following Table IV summarizes the averaged response scores derived via all 15 stakeholders involved in this experiment.

TABLE IV. RESPONSE SCORES

Verbalized User Guide [VU]	Law	85%
	COVID-19	80%
	Aquaculture	83%
	Averaged	82%

A summarized collection of the qualitative interpretations and insights gathered through the controlled interview session were depicted in Table V.

TABLE V. QUALITATIVE RESPONSE OVERVIEW

Verbalized User Guide [VU]	1.	Surface level enlightenment for the domain specialists about the conceptualizations of the ontology increments
	2.	Boosts blended comprehension when visualization and verbalizations are integrated during analysis.
	3.	Offline information reviewing facility is important, as doesn't need to stick to a computer screen all the time.
	4.	Acts as a user guide/manual for each version of the ontology increment.
	5.	Fully automated, configuration free, domain and schema independent operation.
	6.	Natural language (i.e., English) representation of the semantic concepts in a domain specialist friendly manner.

As the final phase of the evaluation process, the iterative framework was applied to reflective asses on the research objective accomplishment. Iterative framework [49-51] is an established framework to evaluate the efficacy of research objective accomplishment logically. The operation of the iterative framework is governed via three different but interconnected questions. Reflective evidence must be provided for each section in place. The discussion associated with the iterative framework steps were logged in Table VI.

TABLE VI. ITERATIVE FRAMEWORK

Steps in Iterative Framework	Reflective Evidence
01 → What are the data telling me?	<p>Quantitative Metrics:- As depicted in Table IV, multiple domain-specific qualitative opinion scores were utilized to validate the effectiveness of the constructed verbalization prototype and its operational effectiveness.</p> <p>Qualitative Assessment:- With the involvement of the stakeholders contributed for the ontology increment constructions, empirical assessment of the verbalization prototype was done, in terms of the results returned, accuracy, usability, technical aid provided & etc. Further, as visible in Table V, the stakeholders' reflective opinion themes were logged</p> <p>As a collective reconciliation, it can be concluded both the quantitative and qualitative experimental phases have yielded satisfactory results.</p>
02 → What do I want to know?	The overall operational efficacy of the verbalization algorithm/prototype developed in terms of domain/schema independence knowledge verbalization.

03 → Is there a dialectical relationship between step 01 & 02?	<p>In the quantitative phase of evaluation, the verbalization prototype is exposed to multiple ontology increments in three different domains. In all those, experiments quantitative matrices are calculated to determine the overall efficacy of the verbalization prototype and it's apparent the overall operation has yielded successful results.</p> <p>In the qualitative assessment phase, stakeholders' opinions were thematically assessed, and refined outcomes were tabulated in Table V.</p> <p>Both, quantitative and qualitative evaluation phases conducted on the criteria of the domain/schema independence verbalization, have collectively yielded successful outcomes.</p> <p>Therefore, as the overall final reflection, as per the iterative framework rationale, it can be concluded as, there is a positive and satisfactory link between step-01 and step-02, which reflects the overall efficacy of the verbalization prototype/algorithm resulted from this research.</p>
---	---

V. CONCLUSION

Effective synchronization between ontologists and domain specialists are a must to accomplish the effective operation of the collaborative ontology engineering goals. To fulfil this purpose, domain specialists should closely involve in the verification process of knowledge embeddings present in each ontology increment. Because error-free, domain-oriented applied ontology construction is an iterative and incremental operation. Hence, it's an extremely effective practice to expose the ontology increment for the cross-validations of the knowledge embeddings, at the end of each iteration.

Verbalization is recognized as a very effective procedure for fulfilling this necessity, as it's a natural language representation mechanism of the internal knowledge embeddings of the ontology of concern.

Usually, lack of technical literacy on semantic concepts, querying skill sets will act as a barrier for the non-computing domain specialists (i.e. lawyers, medical doctors, business consultants, bankers, etc. ...), to complete their role of cross-validation in collaborative ontology engineering workflow.

But as conversed in this paper, almost, all of the existing verbalizers have lots of limitations, confiding their operation to a statically attached one domain or complex configuration phases or technically complex verbalized results that are not easy to interpret. Therefore, this research focused on addressing those shortcomings of the existing verbalizers and uplifting the operational efficacy of the verbalizers to the next level, by making them operate in a domain and schema independent manner.

In this research, AIML based Alice bot's operational flow is transferred into a form of a verbalizer, by the introduction of

a fully automated newly defined algorithm, which is a significant technical contribution resulting out from this research. Henceforth, its operational accuracy is quantitatively and qualitatively evaluated, where both mechanisms have yielded an overall of 82% acceptance.

Domain and schema independent ontology verbalization with no manual configurations and fully automated user guide construction for the ontology of concern are two critical application-level contributions yielding out from this research.

But as one limitation, it can be concluded that this verbalizer will only work for ontologies with a lexicon-based schematic structure, as the backbone of this prototype is developed on top of a chatter bot's architecture.

REFERENCES

- [1] Spasic, I., Ananiadou, S., McNaught, J., & Kumar, A. (2005). Text mining and ontologies in biomedicine: Making sense of raw text. *Briefings in Bioinformatics*, 6(3), 239-251.
- [2] Kashyap V (2008) Ontologies and Schemas. *The Semantic Web*, 79-135. doi:10.1007/978-3-540 764526_5
- [3] Abeyesiriwardana, P. C., & Kodituwakku, S. R. (2012). Ontology-Based Information Extraction for Disease Intelligence. *International Journal of Research in Computer Science*, 2(6), 7-19. doi:10.7815/ijorcs.26.2012.051
- [4] Katifori, A., Vassilakis, C., & Dix, A. (2010). Ontologies and the brain: Using spreading activation through ontologies to support personal interaction. *Cognitive Systems Research*, 11(1), 25-41. doi:10.1016/j.cogsys.2009.02.001
- [5] Hendler, J., & Berners-Lee, T. (2010). From the Semantic Web to social machines: A research challenge for AI on the World Wide Web. *Artificial Intelligence*, 174(2), 156-161. doi:10.1016/j.artint.2009.11.010
- [6] Ramanathan V. Guha (2013). "Light at the End of the Tunnel". *International Semantic Web Conference 2013 Keynote*.
- [7] De Nicola, A., & Missikoff, M. (2016). A lightweight methodology for rapid ontology engineering. *Communications of the ACM*, 59(3), 79-86. doi:10.1145/2818359 doi:10.1093/bib/6.3.239
- [8] Ingram, J., & Gaskell, P. (2019). Searching for meaning: Co-constructing ontologies with stakeholders for smarter search engines in agriculture. *NJAS - Wageningen Journal of Life Sciences*, 100300. doi:10.1016/j.njas.2019.04.006
- [9] Strohmaier, M., Walk, S., PPSchko, J., Lamprecht, D., Tudorache, T., Nyulas, C., ...Noy, N. F. (2013). How Ontologies are Made: Studying the Hidden Social Dynamics Behind Collaborative Ontology Engineering Projects. *SSRN Electronic Journal*. doi:10.2139/ssrn.3199036
- [10] Narock, T., & Fox, P. (2015). *The Semantic Web in Earth and Space Science. Current Status and Future Directions*. Amsterdam, Netherlands: IOS Press.
- [11] McDaniel, M., & Storey, V. C. (2019). Evaluating Domain Ontologies. *ACM Computing Surveys*, 52(4), 1-44. doi:10.1145/3329124
- [12] Façanha, R. L., Cavalcanti, M. C., & Campos, M. L. (2019). A Systematic Approach to Review Legacy Schemas Based on Ontological Analysis. *Metadata and Semantic Research*, 63-75. doi:10.1007/978-3-030-14401-2_6
- [13] Trokanas, N., & Cecelja, F. (2016). Ontology evaluation for reuse in the domain of Process Systems Engineering. *Computers & Chemical Engineering*, 85, 177-187. doi:10.1016/j.compchemeng.2015.12.003
- [14] Munir, K., & SherazAnjum, M. (2018). The use of ontologies for effective knowledge modelling and information retrieval. *Applied Computing and Informatics*, 14(2), 116-126. doi:10.1016/j.aci.2017.07.003
- [15] Elve, A. T., & Preisig, H. A. (2018). From ontology to executable program code. *Computers & Chemical Engineering*. doi:10.1016/j.compchemeng.2018.09.004
- [16] Saha, S., Usman, Z., Li, W., Jones, S., & Shah, N. (2019). Core domain ontology for joining processes to consolidate welding standards. *Robotics and Computer-Integrated Manufacturing*, 59, 417-430. doi:10.1016/j.rcim.2019.05.010
- [17] Elgammal, A., & Turetken, O. (2015). Lifecycle Business Process Compliance Management: A Semantically-Enabled Framework. *2015 International Conference on Cloud Computing (ICCC)*. doi:10.1109/cloudcomp.2015.7149646
- [18] Simperl, E., & Luczak-Rösch, M. (2013). Collaborative ontology engineering: a survey. *The Knowledge Engineering Review*, 29(1), 101-131. doi:10.1017/s0269888913000192
- [19] Cheng, Y., He, F., Lu, X., & Cai, W. (2019). On the role of generating textual description for design intent communication in feature-based 3D collaborative design. *Advanced Engineering Informatics*, 39. Retrieved from <https://doi.org/10.1016/j.aei.2019.02.003>
- [20] Venugopal, & Kumar. (2015). *Ontology Verbalization using Semantic-Refinement*. IOS Press. doi:0000-0000/15/\$00.00
- [21] MacArthur, J.A., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., Pendlington, Z.M., Welter, D., Burdett, T., Hindorf, L., Flicek, P., Cunningham, F., & Parkinson, H.E. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research*.
- [22] Poveda-Villalón, M., García-Castro, R., & Gómez-Pérez, A. (2014). Building an ontology catalogue for smart cities.
- [23] Bonacin, R., Nabuco, O., & Pierozzi, I. (2016). Ontology models of the impacts of agriculture and climate changes on water resources: Scenarios on interoperability and information recovery. *Future Generation Comp. Syst.*, 54, 423-434.
- [24] Strohmaier, M., Walk, S., PPSchko, J., Lamprecht, D., Tudorache, T., Nyulas, C., ...Noy, N. F. (2013). How Ontologies are Made: Studying the Hidden Social Dynamics Behind Collaborative Ontology Engineering Projects. *SSRN Electronic Journal*. doi:10.2139/ssrn.3199036
- [25] Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243-1248. Retrieved January 29, 2020, from www.jstor.org/stable/1724745
- [26] Nagle, Frank, The Digital Commons: Tragedy or Opportunity? A Reflection on the 50th Anniversary of Hardin's Tragedy of the Commons (December 13, 2018). Harvard Business School Strategy Unit Working Paper No. 19-060. Available at SSRN: <https://ssrn.com/abstract=3301005> or <http://dx.doi.org/10.2139/ssrn.3301005>
- [27] May, P. Hanrahan, D. A. Keim, B. Shneiderman, and S. Card, "The state of visual analytics: Views on what visual analytics is and where it is going," in 2010 IEEE Symposium on Visual Analytics Science and Technology, 2010, pp. 257-259.
- [28] Caldarella, E.G., Rinaldi, A.M.: An approach to ontology integration for ontology reuse. In: 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI) (2016). <https://doi.org/10.1109/iri.2016.58>
- [29] Bonacin, R., Nabuco, O., & Pierozzi, I. (2016). Ontology models of the impacts of agriculture and climate changes on water resources: Scenarios on interoperability and information recovery. *Future Generation Comp. Syst.*, 54, 423-434.
- [30] Strohmaier, M., Walk, S., PPSchko, J., Lamprecht, D., Tudorache, T., Nyulas, C., ...Noy, N. F. (2013). How Ontologies are Made: Studying the Hidden Social Dynamics Behind Collaborative Ontology Engineering Projects. *SSRN Electronic Journal*. doi:10.2139/ssrn.3199036
- [31] Androutopoulos, I., Lampouras, G., & Galanis, D. (2013). Generating Natural Language Descriptions from OWL Ontologies: the NaturalOWL System. *Journal of Artificial Intelligence Research*, 48, 671-715. doi:10.1613/jair.4017
- [32] Bouayad-Agha, N., Casamayor, G., & Wanner, L. (2014). Natural Language Generation in the context of the Semantic Web. *Semantic Web*, 5, 493-513.
- [33] Sportelli, F., & Franconi, E. (2016). Formalisation of ORM Derivation Rules and Their Mapping into OWL. On the Move to Meaningful

- Internet Systems: OTM 2016 Conferences, 827-843. doi:10.1007/978-3-319-48472-3_52
- [34] Dai, Y., Zhang, S., Chen, J., Chen, T., and Zhang, W. Semantic Network Language Generation Based on a Semantic Networks Serialization Grammar. *World Wide Web* 13, 3(2010), 307–341.
- [35] Demir, S., Carberry, S., and McCoy, K. A Discourseware Graph-based Content Selection Framework. In *Proceedings of the 6th International Natural Language Generation Conference (INLG) (2010)*, pp. 17–27.
- [36] Bojars, U., Liepins, R., Gruzitis, N., Cerans, K., & Celms, E. (2016). Extending OWL Ontology Visualizations with Interactive Contextual Verbalization. *VOILA@ISWC*.
- [37] Keet, C. M., & Khumalo, L. (2016). On the verbalization patterns of part-whole relations in isiZulu. *Proceedings of the 9th International Natural Language Generation conference*. doi:10.18653/v1/w16-6629
- [38] Kaarel Kaljurand and Norbert E. Fuchs. 2007. Verbalizing owl in attempt to controlled English. In *Proceedings of Third International Workshop on OWL: Experiences and Directions*, Innsbruck, Austria (6th–7th June 2007), volume 258
- [39] Guimarães, M. A., Zisman, R. P., & Renzi, A. B. (2019). Pharmaceutical Online Store Project: Usability, Affordances and Expectations. *Advances in Intelligent Systems and Computing*, 523-534. doi:10.1007/978-3-030-20040-4_47
- [40] Sun, Mi, Olsson, Paulsson, & Harrie. (2019). Utilizing BIM and GIS for Representation and Visualization of 3D Cadastre. *ISPRS International Journal of Geo-Information*, 8(11), 503. doi:10.3390/ijgi8110503
- [41] Dorothy, B., & S. Ramesh Kumar, B. (2018). DORBRI: An Architecture for the DoD Security Breaches Through Quantum IoT. *International Conference on Computer Networks and Communication Technologies*, 491-496. doi:10.1007/978-981-10-8681-6_44
- [42] Whittemore R. 2005. Combining evidence in nursing research: methods and implications. *Nursing Research* 54(1):56-62
- [43] Renzi, A.B., Chammas, A., Agner, L., Greenspan, J.: Startup Rio: user experience and startups. In: *Design, User Experience, and Usability: Design Discourse*, Los Angeles, CA, pp. 339–347. Spring (2015)
- [44] Arnold, R.D., & Wade, J. (2015). *A Definition of Systems Thinking: A Systems Approach*
- [45] Wallace, *The Elements of AIML Style*, ALICE A. I. Foundation, Inc. March 28, 2003
- [46] Wallace, R. S. (2014). AIML 2.0 Working Draft, <https://docs.google.com/document/d/1wNT25hJRyupcG51aO89UcQEiG-HkXRXusukADpFnDs4/pub>
- [47] McCarthy, J. (1980). "Circumscription—A Form of Non-Monotonic Reasoning." *Artificial Intelligence* 13(1–2): 27–39.
- [48] Madhumitha, Keerthana, & Hemalatha. (2019). Interactive Chatbot Using AIML. *Advanced Networking & Applications*, 217-223.
- [49] Stockdale, R., & Standing, C. (2006). An interpretive approach to evaluating information systems: A content, context, process framework. *European Journal of Operational Research*, 173, 1090-1102.
- [50] Boyatzis, R. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*.
- [51] Srivastava, P., & Hopwood, N. (2009). *A Practical Iterative Framework for Qualitative Data Analysis*.
- [52] Androutsopoulos, I., Lampouras, G., & Galanis, D. (2013). Generating Natural Language Descriptions from OWL Ontologies: the NaturalOWL System. *J. Artif. Intell. Res.*, 48, 671-715.
- [53] Gonano, C.M., Tomasi, F., Mambelli, F., Vitali, F., & Peroni, S. (2014). Zeri e LODE. Extracting the Zeri photo archive to linked open data: formalizing the conceptual model. *IEEE/ACM Joint Conference on Digital Libraries*, 289-298.
- [54] Third, A., Williams, S., & Power, R. (2011). *OWL to English: a tool for generating organised easily-navigated hypertexts from ontologies*.
- [55] Bontcheva, K., & Wilks, Y. (2004). Automatic Report Generation from Ontologies: The MIAKT Approach. *Natural Language Processing and Information Systems*, 324-335. doi:10.1007/978-3540-27779-8_28
- [56] Hossain, B., Rajan, G., & Schwitter, R. (2019). CNL-ER: A Controlled Natural Language for Specifying and Verbalising Entity Relationship Models. *ALTA*.
- [57] Krishnamurthy, R., & Hsiao, M. (2020). Transforming Natural Language Specifications to Logical Forms for Hardware Verification. *2020 IEEE 38th International Conference on Computer Design (ICCD)*, 393-396.
- [58] Sak, H., Beaufays, F., Nakajima, K., & Allauzen, C. (2013). Language model verbalization for automatic speech recognition. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 8262-8266.
- [59] Perera, V., Selvaraj, S.P., Rosenthal, S., & Veloso, M. (2016). Dynamic generation and refinement of robot verbalization. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 212-218.
- [60] Venugopal, V.E. (2018). *Ontology Verbalization using Semantic-Refinement*.
- [61] Janarthanam, S. (2017). *Hands-On Chatbots and Conversational UI Development: Build chatbots and voice user interfaces with Chatfuel, Dialogflow, Microsoft Bot Framework, Twilio, and Alexa Skills*.
- [62] Ralston, K., Chen, Y., Isah, H., & Zulkermine, F. (2019). A Voice Interactive Multilingual Student Support System using IBM Watson. *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, 1924-1929.
- [63] Bocklisch, T., Faulkner, J., Pawlowski, N., & Nichol, A. (2017). *Rasa: Open Source Language Understanding and Dialogue Management*. ArXiv, abs/1712.05181.
- [64] Jespersen, L.N., Michelsen, S.I., Holstein, B., Tjørnhøj-Thomsen, T., & Due, P. (2018). Conceptualization, operationalization, and content validity of the EQOL-questionnaire measuring quality of life and participation for persons with disabilities. *Health and Quality of Life Outcomes*, 16.
- [65] Fonseca, c. (2016). *A new theory for the study of texts: theory of rhetoric structure - Rethorical structure theory*.

APPENDIX

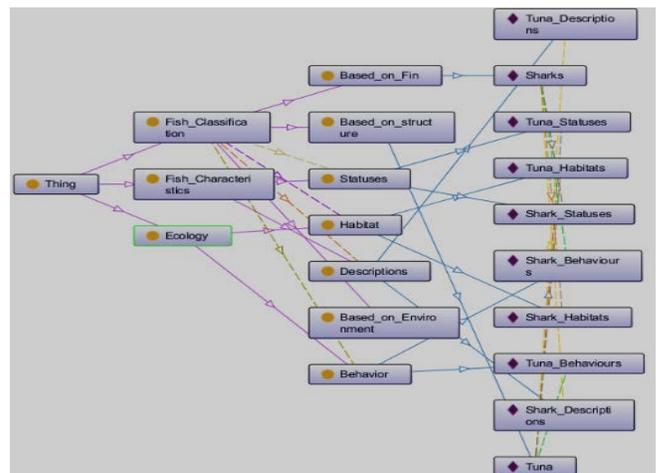


Fig. 10. Ontology Increment for Aquaculture Domain.

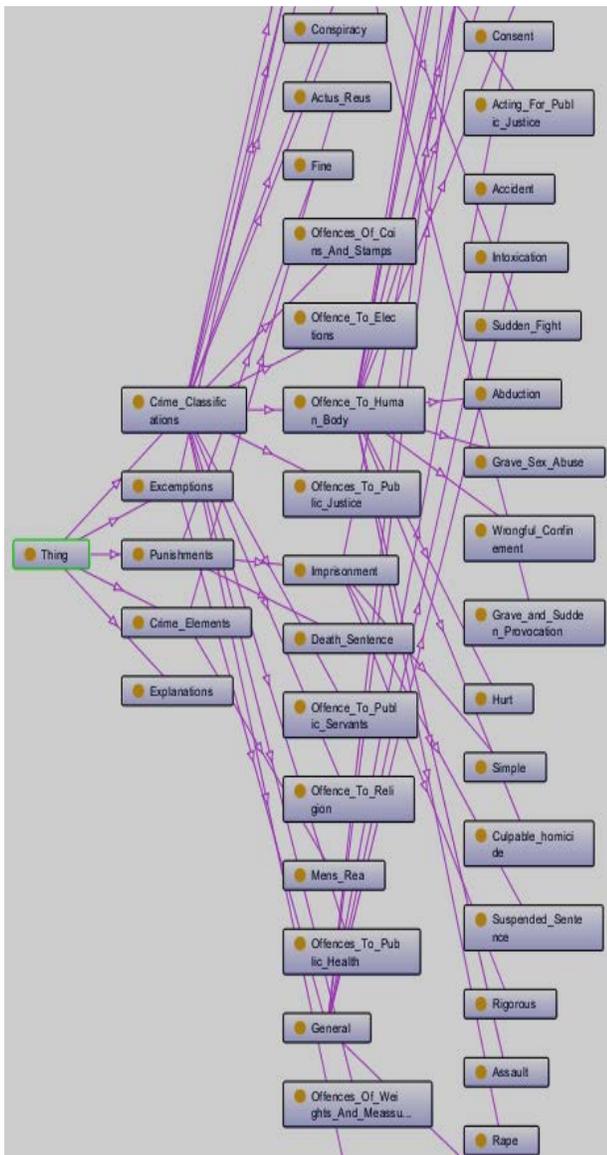


Fig. 11. Ontology Increment for Criminal Law.

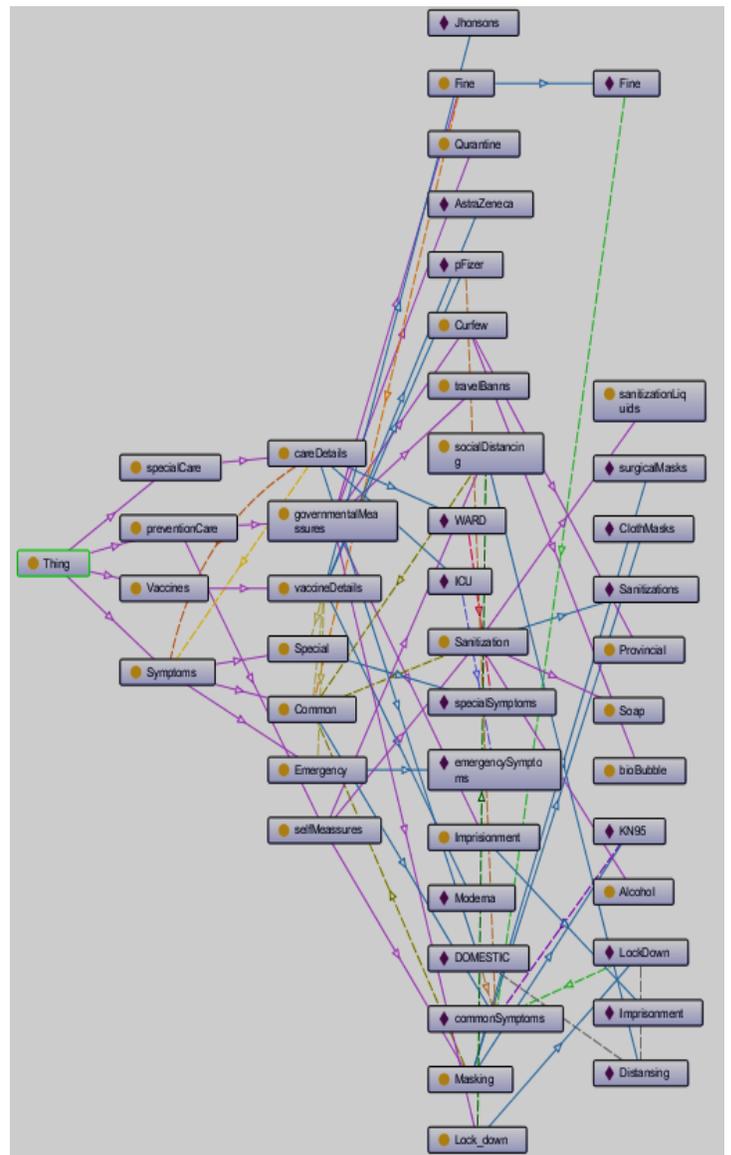


Fig. 12. Ontology Increment for COVID-19.

Design of an Efficient RPL Objective Function for Internet of Things Applications

Sonia Kuwelkar¹, H.G. Virani²

Electronics and Telecommunications Department
Goa College of Engineering, Goa, India

Abstract—Over the decade, a rapid growth in use of smart devices connected and communicating over the Internet is seen in various domains. The IPv6 Routing protocol for Low power and Lossy Networks (RPL) is the routing backbone of such IOT networks. RPL is a proactive, distance vector protocol which constructs the routes based on an Objective Function. The performance of RPL protocol largely depends on the design of Objective Function. Depending on application requirements, the RPL standard offers flexibility in design of the objective function and scope for improving the routing process. In this paper, an efficient Objective Function, RPL-FZ, is proposed. Speedy communication across nodes, low energy consumption and reliable data delivery is key to achieve quality of Service. Considering this, RPL-FZ uses relevant metrics like Residual Energy of Node, Delay and ETX (Expected Transmission count) to make the routing decisions. The metrics are combined using fuzzy logic technique to obtain a single metric Quality for each neighbor node. The neighbor with highest value of Quality is chosen as best parent to forward sensed data toward the collection unit. The proposed objective function RPL-FZ is integrated in the Contiki OS and network simulations are performed using the COOJA simulator. The performance evaluation reveals that RPL-FZ achieves 7% higher Packet Delivery rate, 8% lower energy consumption and 8% lesser latency as compared to single metric based standard objective functions OF0 and MRHOF.

Keywords—Internet of things; low power Lossy Networks; IPv6 routing protocol for LLN; objective function; fuzzy logic

I. INTRODUCTION

The Internet of Things finds application in Automatic Meter Reading, Industrial Monitoring, Healthcare, Home automation, Surveillance and Weather monitoring. The sensor nodes employed in these networks have limited processing capability, constrained memory and are battery powered [1]. The links which connect the nodes support low data rates and have comparatively lower packet delivery ratios. These networks support multipoint-to-point, point-to-multipoint and point-to-point data traffic. When deployed in practical application field, these networks can scale to density of over thousand nodes. Such unique requirements classify an IOT network as a Low power and Lossy Network. The design of Routing protocol for these networks is a challenging task. The conventional routing protocols designed for Wireless Sensor Networks, ADHOC networks and MANETs are not suitable to adapt to the lossy and dynamic nature of these networks. To cater to the specific requirements of these networks, the IPv6-Routing Protocol for Low power and Lossy Network (RPL) [2] were proposed by the IETF ROLL working group.

RPL is a distance vector routing protocol which creates Destination Oriented Directed Acyclic graphs (DODAG) identical to tree topology. These DAGs have a Root node, which is generally the border router or gateway, and all other nodes in the network direct their data towards the root node. The Root node initiates the DODAG construction by broadcasting a DODAG Information Object (DIO) message to all neighboring nodes. The DIO message holds network sensitive information like DODAG ID, RPL Instance ID, Objective Function, metrics, constraints and Rank. When a particular node receives a DIO message from a neighbor, it adds the sender nodes routing related info to its parents list as a candidate parent node (CPN). A best parent will be chosen from the candidate parents list to forward the data to the DODAG root. The selection of preferred parent node is governed by the metric specified in Objective function. RPL standard lists two objective functions, OF0 [3] and MRHOF [4]. The objective function MRHOF (Minimum Rank with Hysteresis Objective Function) utilizes the link quality metric-ETX (Expected transmission count) to decide the preferred parent and best path to root. Whereas, OF0 (Objective Function Zero) uses hop count as a decision criterion. Both the Objective Functions (OF) rely on a single metric to make the best parent choice. Once the best parent is selected, the node will compute its own Rank with respect to its parent Rank. The DODAG root has the lowest Rank and every other node computes a rank which increases proportionately with its distance from the root node. Every node transmits the updated DIO message to its neighboring nodes. A Destination Advertisement Object (DAO) message is sent by nodes, except the root, to advertise their addresses and prefixes to their parents and to populate the routing tables with prefixes of their children. A node which has newly joined the network can solicit a DIO message from its neighbour using a DODAG Information Solicitation (DIS) message. The DIO message transmission is triggered by the expiration of a Trickle Timer. To reduce the control overhead the trickle timer interval is doubled after a DIO is transmitted. Fig. 1 shows a DODAG and the routing messages in the network.

The routing mechanism and data forwarding techniques in RPL is simple and flexible to suit various applications. RPL utilizes minimum memory for routing table information and minimizes routing signalling. It is included in powerline communications like G3-PLC and in Zigbee IP specifications [8].

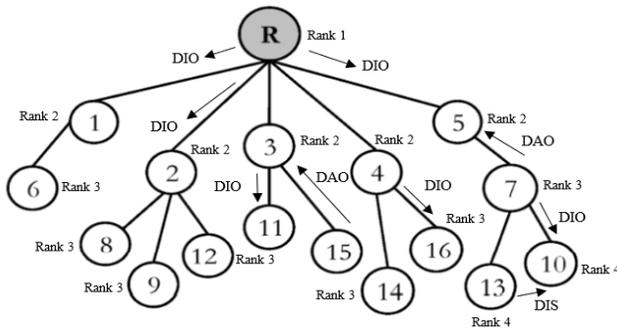


Fig. 1. DODAG.

RPL is also incorporated in IP stack of open-source operating systems like Contiki OS, Tiny OS, LiteOS, T-Kernel, EyeOS and RIOT meant for low power and constrained nodes.

To improve the performance of a RPL based network, the following contributions are presented in this paper:

1) An efficient Objective Function RPL-FZ, which makes routing decisions and best parent selection based on three metrics namely Residual Energy (RE), Delay and ETX (Expected Transmission Count) is proposed. Fuzzy logic is a very popular technique for Multi criteria decision making. These three metrics are combined using Fuzzy logic process to compute a single metric Quality. A node will compute the Quality values of every neighbor depending on Residual Energy, Delay and ETX values. The neighbor with maximum Quality is chosen as the preferred parent node for forwarding data to the root. The fuzzy Inference system is triggered at every node upon receiving a DIO message.

2) The designed Objective function RPL-FZ is extensively tested in COOJA simulator of Contiki OS by varying network density from 70 nodes to 100 nodes and increasing the data transmission rate from 2 packets/min to 10 packets/min. Improvement in Power consumption, Latency and Packet Delivery Ratio is observed in comparison to the RPL standard Objective Functions OF0 and MRHOF.

The paper is organized as follows: In section II, similar work done by other researchers is reviewed. In section III, detailed design of the proposed Objective function RPL-FZ using the fuzzy logic process is presented. In section IV, the effectiveness of the Objective function RPL-FZ is assessed through simulations. In section V, the conclusions and future work are presented.

II. RELATED WORK

RPL being the standard protocol for routing in Low power Lossy Networks, has attracted attention of researchers across the world. In this section, previous work related to optimization of RPL protocol is discussed. Some researchers have evaluated the performance of the protocol under different network conditions [5-6] and shown practical implementation [7]. Some have proposed novel methods of link estimation, neighbor table management [8], interference reduction and

load balancing [9]. Similar to our study, techniques to optimize the objective function using fuzzy logic are presented [10] – [13].

Abuein et al [5] in their study have evaluated the performance of OF0 and MRHOF objective functions of RPL under the random and grid topologies. The network density is medium and varies from 50 to 85 nodes. They concluded that the network performs best when the density is between 50-65 nodes and the RX is 60%. In an earlier study [6] the authors have analyzed the performance of OF0 and MRHOF in two network scenarios. In scenario 1, simulations were carried out by varying network densities from 30 to 70 in area of 300x300 m² and using a single sink node. In scenario 2, node density was varied from 70 to 110. Power consumption, Packet Deliver Ratios and Network convergence time was computed. It was observed that MRHOF had advantage over OF0 when node density was low, but when the network density increased beyond 90, OF0 performed better than MRHOF.

Kitagawa et al [7] have illustrated the practical implementation of the RPL protocol. Pulse sensors send the patients pulse data continuously using Zigbee technology to a remote Raspberry Pi based internet gateway. The sensors route the data using the RPL protocol. Received signal strength, Path loss and Average power consumption are plotted. It is noted that position of sensors affects the power consumption.

Ancillotti et al [8] have addressed the issue of passive link monitoring in RPL. A unique cross layer design for link estimation and novel technique for effective neighbor table management is suggested. A hybrid link estimation approach is used which improves the packet delivery ratios. S Kharche et al [9] have proposed a routing algorithm DN-RPL based on deep neural network. This algorithm attempts to resolve the issue of interference in 6LoWPANs. The deepnets based routing DN-RPL is compared with machine learning based RPL (ML-RPL).

Adeeb Saaidah et al in their work [10] have suggested an improved Objective function OFRRT-FUZZY for RPL. It combines three metrics Received Signal Strength Indicator, Throughput (TH) and Remaining Energy (RE). They have used the popular fuzzy logic process to merge these metrics and make routing decision. Simulations performed using COOJA simulator show that OFRRT-FUZZY performs better than OF0 and MRHOF. Mah Zaib Jamil et al [11] have proposed an Objective function based on a new metric ELT (Expected Lifetime metric) of node. ELT is computed by measuring ratio of residual energy to energy used. Comparison analysis of RPL implementation with ELT, HOP and ETX metrics is done. Hanane et al [12] have made an attempt to improve RPL by designing an objective function OF-EC for RPL which depends on three metrics namely ETX, hop count and Energy Consumption. They have used the Fuzzy Logic process to design this Objective Function.

In [13] the authors have proposed a cross layer Objective function ELITE. A metric SPR is proposed and used for route formation. SPR (Strobe per Packet Ratio) measures the required strobes per packet at the MAC layer. ELITE was

successful in reducing the average number of strobes needed per packet and average amount of energy consumed by sensor node.

III. DESIGN OF RPL-FZ USING FUZZY LOGIC

In this section, detailed design of proposed Objective Function RPL-FZ is explained. The RPL standard allows the designer flexibility in choice of metrics for Objective Function design [14]. In our work we have chosen three metrics namely, Residual Energy, Delay and Expected transmission count in the Objective function design. The three metrics are chosen considering the real time, reliability and low power consumption requirements of Quality of service. These three metrics have to be combined to obtain a single decision metric for routing. There are several multi criteria decision making techniques which can be used. However, the sensor nodes have limited processing and memory capability. Considering this Fuzzy logic is the ideal option [15]. Fuzzy logic is an accurate method widely used to combine several values to obtain a single precise output [16]. It is ideal in LLN as the implementation is supported by the sensor devices and can enhance network performance. In this work, the Mamdani Fuzzy Inference model [17] is used to obtain accurate output.

The selection of the best parent by a node depends on the routing metric contained in the Objective Function. Hence the routing metric plays an important role in enhancing the overall Network performance. The standard RPL utilizes either Link Quality or Hop count as the routing metric. This may lead to a node with less residual energy being continuously chosen as the Parent node, thus further draining its energy and eventually affecting the Network lifetime [18]. Hence it is necessary to consider the Node Residual Energy as well as Latency in routing process. The metrics considered in this work are as follows:

ETX (Expected transmission count): It represents the number of times transmission is required to successfully deliver a packet to its destination. A good link should be able to deliver the packet in first attempt. Ideally, a low ETX value signifies a good and reliable link.

Delay: It is the time taken for a packet to reach its destination. For real time application delay should be minimum.

Residual Energy: When a node is deployed in the network it has a level of Energy E_0 . Its energy reduces during its network operations. The difference between initial energy E_0 and energy consumed is called the Residual Energy. The intent here is to avoid selecting a node which is running low on energy as the best parent. This eventually helps in prolonging the network lifetime.

A. Combining Routing Metrics using Fuzzy Inference System

The Fuzzy Inference System is shown in Fig. 2. It contains a Fuzzifier unit, Fuzzy inference engine with the Rule base and the Defuzzifier unit.

The fuzzification process consists of following steps:

Fuzzification: In this step, the crisp inputs are specified in terms of linguistic variables. The membership functions for

the linguistic variables are defined. In our work the crisp input variables are ETX, Delay and Residual Energy.

Fuzzy inference System: The FIS combines fuzzified inputs based on set of AND-OR rules and calculates the output.

Aggregation: If the output depends on more than one rule than it is unified using aggregation method.

Defuzzification: This is process of converting the fuzzified output into a crisp value. In our work this is the metric Quality.

The first step is to specify the linguistic variables for ETX, RE and Delay. The linguistic variables for the ETX input variable are short, average and long. A link with less than 3 signifies as good link and more than 12 is considered undesirable. The Delay term is represented with linguistic variables small, average and high. A value of 3000 units signifies high delay for packet delivery. The Residual Energy is denoted using terms low, average and full. At start of deployment the nodes have full energy of value 255. The membership functions for RE, ETX and Delay are shown in Fig. 3. Rectilinear (Triangular or trapezoidal) membership functions can be chosen to quantify the linguistic variables. In this work, the fuzzy sets are specified using Trapezoidal membership functions to achieve better results.

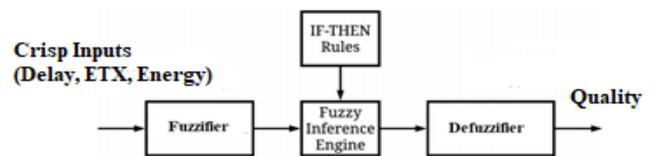


Fig. 2. Block Diagram of Fuzzy Inference System.

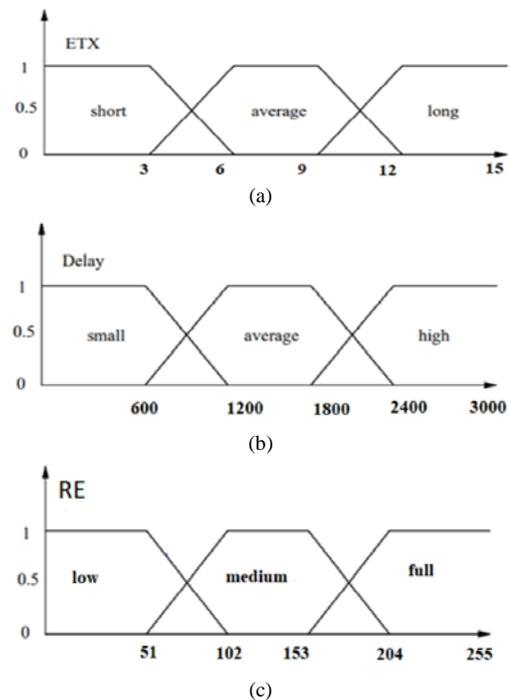


Fig. 3. Membership Functions (a) ETX (b) Delay and (c) Residual Energy.

The formula defining the membership in the short fuzzy set of ETX is given by

$$\mu(\text{short})(ETX) = \begin{cases} 0, ETX > 6 \\ \frac{6 - ETX}{6 - 3}, 3 \leq ETX \leq 6 \\ 1, ETX \leq 3 \end{cases}$$

The formula for membership in the Average fuzzy set of ETX is given by

$$\mu(\text{avg})(ETX) = \begin{cases} 0, ETX < 3 \text{ or } ETX > 12 \\ \frac{ETX - 3}{6 - 3}, 3 \leq ETX \leq 6 \\ 1, 6 \leq ETX \leq 9 \\ \frac{12 - ETX}{12 - 9}, 9 \leq ETX \leq 12 \end{cases}$$

The formula defining the membership in long fuzzy set of ETX is given by

$$\mu(\text{long})(ETX) = \begin{cases} 0, ETX < 9 \\ \frac{ETX - 9}{12 - 9}, 9 \leq ETX \leq 12 \\ 1, ETX > 12 \end{cases}$$

Similarly, the membership levels can be defined for the fuzzy sets of Residual Energy and delay inputs.

A set of IF-THEN fuzzy rules are stored in the database. These are applied to the fuzzy inputs in the fuzzy inference engine. Different fuzzy inputs are connected by AND, OR and NOT fuzzy operators. The three inputs ETX, RE and Delay have associated with them three fuzzy sets each. The rule base will therefore consist of $3^3 = 27$ rules. The Quality output has associated with it seven fuzzy sets, namely very_bad, bad, degraded, average, satisfactory, good and excellent. The membership function of Quality is shown in Fig. 4. Some of the rules used to compute Quality are shown in Table I.

The rules are stored in database as follows:

- 1) If ETX is short and Delay is small and RE is full then Quality is excellent.
- 2) (If ETX is short and Delay is small and RE is medium then Quality is good) OR (If ETX is average and DELAY is small and RE is full then Quality is good).
- 3) (If ETX is short and DELAY is high and RE is full then Quality is satisfactory) OR (If ETX is average and DELAY is average and RE is full then Quality is satisfactory).
- 4) (If ETX is average and DELAY is average and RE is medium then Quality is Average) OR (If ETX is long and DELAY is small and RE is medium then Quality is Average).
- 5) (If ETX is short and DELAY is average and RE is low then Quality is degraded) OR (If ETX is average and DELAY is small and RE is low then Quality is degraded).
- 6) (If ETX is long and DELAY is average and RE is medium then Quality is bad) OR (If DELAY is high and ETX is long and RE is medium then Quality is bad).
- 7) If ETX is long and DELAY is high and RE is low then Quality is very_bad.

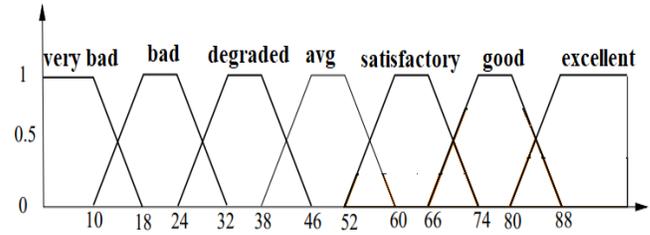


Fig. 4. Quality Membership Function.

TABLE I. QUALITY VALUES

ETX	DELAY	RE	QUALITY
short	small	full	excellent
short	small	medium	good
average	small	full	good
short	high	full	satisfactory
average	average	full	satisfactory
short	high	medium	average
average	average	medium	average
short	average	low	degraded
average	small	low	degraded
long	average	medium	bad
long	high	medium	bad
long	high	low	very_bad

The Quality value varies from 0 to 100. A Quality score is assigned to each candidate parent node. For a particular candidate node when the combination of metrics yields a good Quality score it can be selected as preferred parent.

During defuzzification, the value of Quality is computed using the center of gravity (COG) defuzzification method:

$$M = \frac{\sum_{i=1}^N Vi \times \mu A (Vi)}{\sum_{i=1}^N \mu A (Vi)}$$

where M represents the defuzzified value, N corresponds to number of activated rules, V_i is the domain value of triggered rule i, and μA is the membership value.

B. Procedure of Parent Selection using RPL-FZ

In RPL the DIO messages are sent after regular intervals determined by the trickle timer. If a node does not receive a DIO, then it will transmit a DIS to solicit a DIO from a neighbor. When a node receives the first DIO message it will add the sender node as a candidate parent node (CPN) in its routing table. The metrics ETX, Delay and Residual Energy for the neighbor are obtained. If no other candidate parent node is present in the list, then this node itself is chosen as Preferred Parent. If more than one candidate is there in list then the Fuzzy Inference System is triggered and routing metrics-ETX, RE and Delay are sent as input. After defuzzification, the Quality scores of the candidate parents are compared with the Quality score of existing preferred parent node (PPN). A Candidate node which has the maximum Quality value is chosen as the new Preferred Parent Node. The

rank is calculated and DAO message is sent to the newly selected best parent node. The old PPN is removed and the new updated DIO message is multicast to neighboring nodes.

This process is indicated in flowchart of Fig. 5.

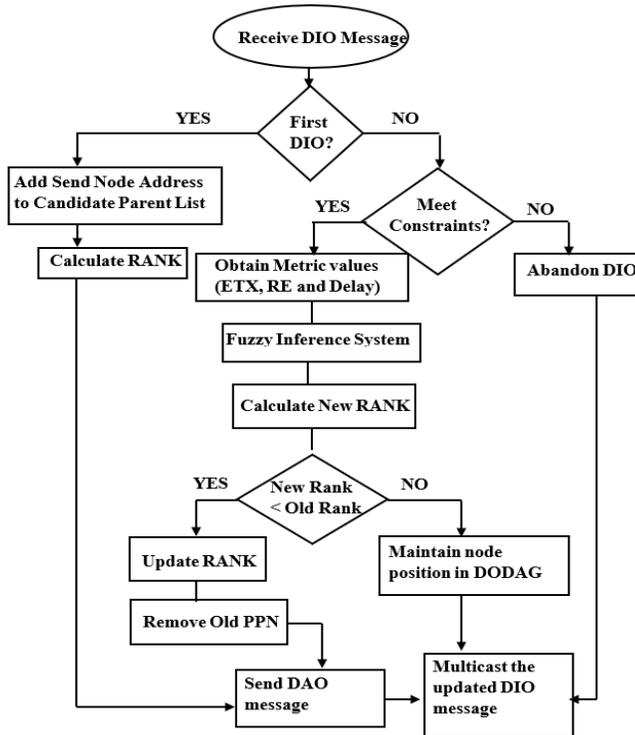


Fig. 5. Flowchart for Parent Selection.

IV. RESULTS AND DISCUSSION

The designed Objective function RPL-FZ is incorporated into the Contiki OS [19]. Simulations are performed using the COOJA network simulator available in Contiki OS [20]. It is a widely used open-source tool. The simulations are performed for evaluating the objective function under varying network density and varying data transmission rates.

The simulation parameters are shown below in Table II.

TABLE II. SIMULATION SETTINGS

Simulator Used	(Contiki OS 2.7) COOJA
Emulated Nodes	T-mote Sky
Network Density	70, 90, 100 nodes
Deployment Type	Random Topology
MAC Layer	Contiki MAC
Transmission Ratio	100%
Network Area	400m * 400m
Transmission Range	30m
Number of Sink Nodes	1
Data Transmission Rate	2 pkts/min, 5 pkts/min and 10 pkts/min
Simulation Time	3 hours

A performance comparison is made between the standard RPL Objective Functions OF0, MRHOF and the Fuzzy based RPL-FZ by varying the density of nodes from 70 to 100, and increasing data transmission rates from 2 p/min to 10 p/min. A single sink was used to collect the data form the devices.

The following performance parameters are computed from the simulation test logs using Perl scripts.

A. Simulation Results

1) *Average power consumption:* It is seen that RPL with MRHOF consumes highest power. This is because MRHOF considers only the transmission count in routing decision and does not consider energy of node. An RPL implementation with the MRHOF objective function can select paths with unreliable links since the link estimation is done by μ IPv6 using passive link monitoring techniques and RPL has no control over it [21]. Thus, it is likely that the routing decisions with MRHOF may be based on outdated link statistics. It is seen form the experiments that MRHOF performs badly as the network density and data rate increases.

At 70 nodes and data rate of 2 p/min, the power consumption with MRHOF is 5% higher than with RPL-FZ. As the node density increases, more traffic in the network causes the power consumption to rise. The sensors have to route more packets, causing them to be in ON mode for longer time duration. At 100 nodes, RPL-FZ consumption is 6% lower than OF0 and 9% lower compared to MRHOF. As the packet rate increases to 10 p/min, power consumption increases further. Higher traffic leads to congestion, collisions and packet loss due to packets being dropped. This causes retransmission of packets. Under these strained conditions of high density and high data rate, RPL-FZ consumes 8% lower power than MRHOF. Although the sensor nodes require extra computational power for the fuzzy process and calculation of rank based on the Quality score, the final power consumption of the Network with RPL-FZ is much lesser as compared to OF0 and MRHOF. RPL-FZ considers the remaining energy of a node during parent selection, thus ensuring that a node low on energy does not end up repeatedly being selected as parent and depleting its energy further. This shows that a network with RPL-FZ will have better lifetime as compared to OF0 or MRHOF. Fig. 6 shows the comparative power consumption of RPL-FZ, MRHOF and OF0.

2) *Packet delivery ratio:* The sensor nodes continuously transmit data to the Root node. Most of the packets are successfully received, but some are lost due to collisions and congestion. The ratio of total packets successfully received at the Sink to the total packets transmitted in the network is termed as Packet Delivery Ratio [22]. Fig. 7 shows the graphs of Packet Delivery Ratios for OF0, RPL-FZ and MRHOF for node density varying from 70 to 100 and packet rates from 2p/min to 10p/min. It can be seen from the graphs that PDR is better at lower packet rates and low node density. At low packet rate 100% packet delivery is obtained with RPL-FZ. RPL with OF0 does not consider the link quality, and when selecting a parent with minimum rank it may actually end up

selecting a congested parent and thus end up dropping packets. RPL-FZ considers ETX as a decision metric and thus ensures better packet delivery. As the sensor data transmission rate increases from 2 ppm to 10 ppm and node density from 70 to 100 it is seen that more packets are lost. This happens due to packet collisions and congestion in the network when packet rate increases. RPL is generally suitable for networks with low data rates. The PDR is also depended on the topology of the network. Our study shows that RPL-FZ gives 7 % better PDR as compared to OF0 and MRHOF when network has 100 nodes and at high packet rate of 10p/min.

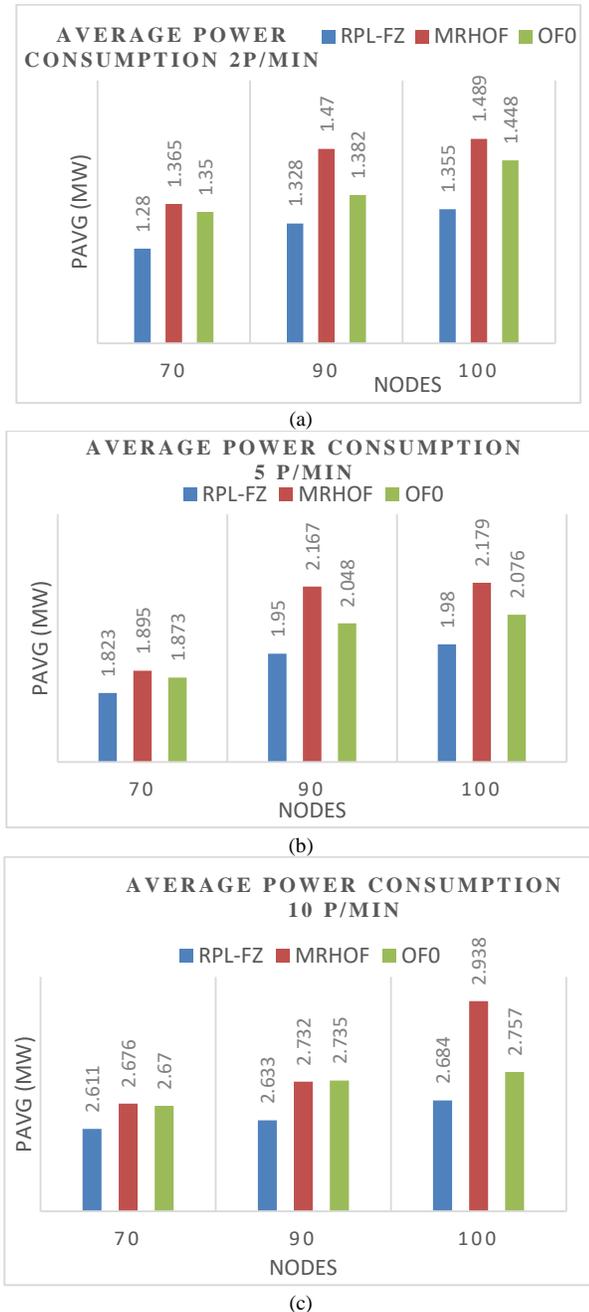


Fig. 6. Average Power Consumption of RPL-FZ, MRHOF and OF0 for (a) 2p/min (b) 5p/min and (c) 10 p/min

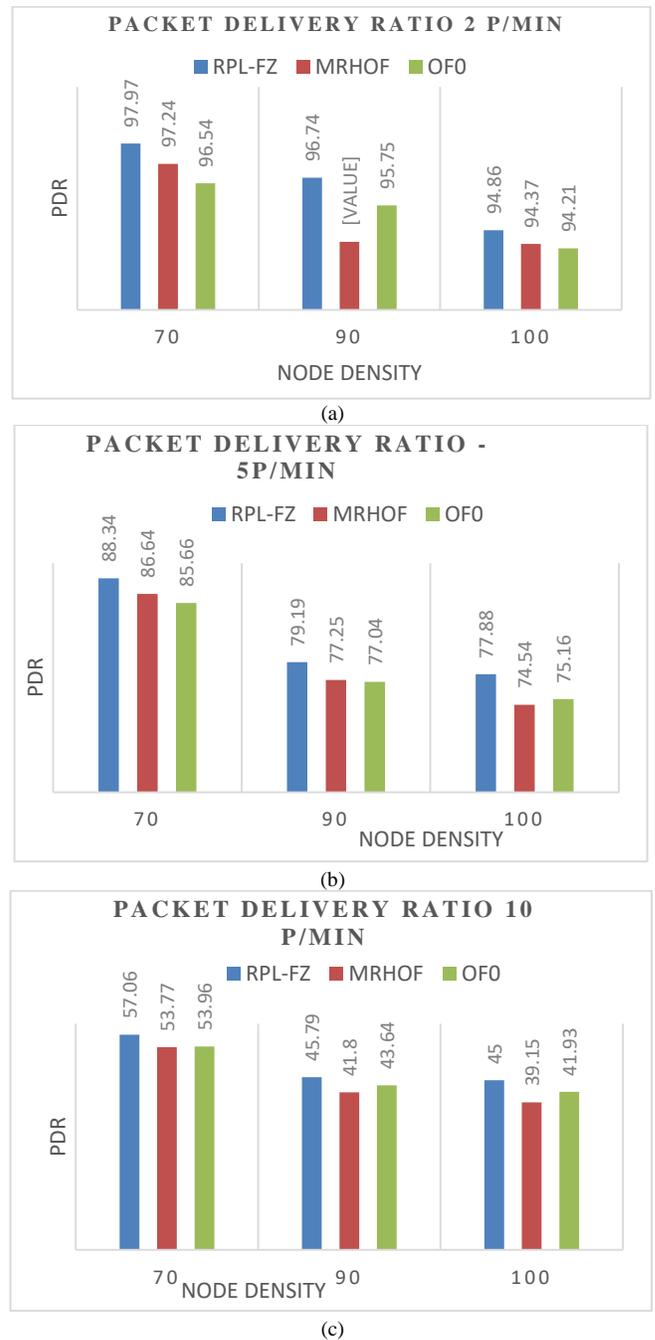


Fig. 7. Packet Delivery Ratio of RPL-FZ, MRHOF and OF0 for a) 2p/min b) 5p/min and c) 10 p/min.

3) *Latency*: The average time taken for a packet to travel from source and reach the destination is termed as latency or end-to-end Delay [23]. Fig. 8 shows the graphs of latency or end to end delay for RPL-FZ, OF0 and MRHOF for varying node densities and packet rates from 2p/min to 10p/min. The three RPL implementations keep the average delay below 2s. MRHOF has highest delay compared to other two. It is seen that the delay increases with increase in node density and packet rates. This is expected since large number of sensors transmitting packets at higher rates will cause more collisions

and packet buffering will increase the delay. RPL-FZ has a delay of 1.8 s, at data rate of 10p/min with 100 nodes, as compared to 2 s for MRHOF and OF0. RPL-FZ uses ETX, residual energy and end to end delay as decision metric for parent selection thus leading to better routing path selection than OF0 and MRHOF. RPL-FZ gives around 8% lower delay than MRHOF and OF0.

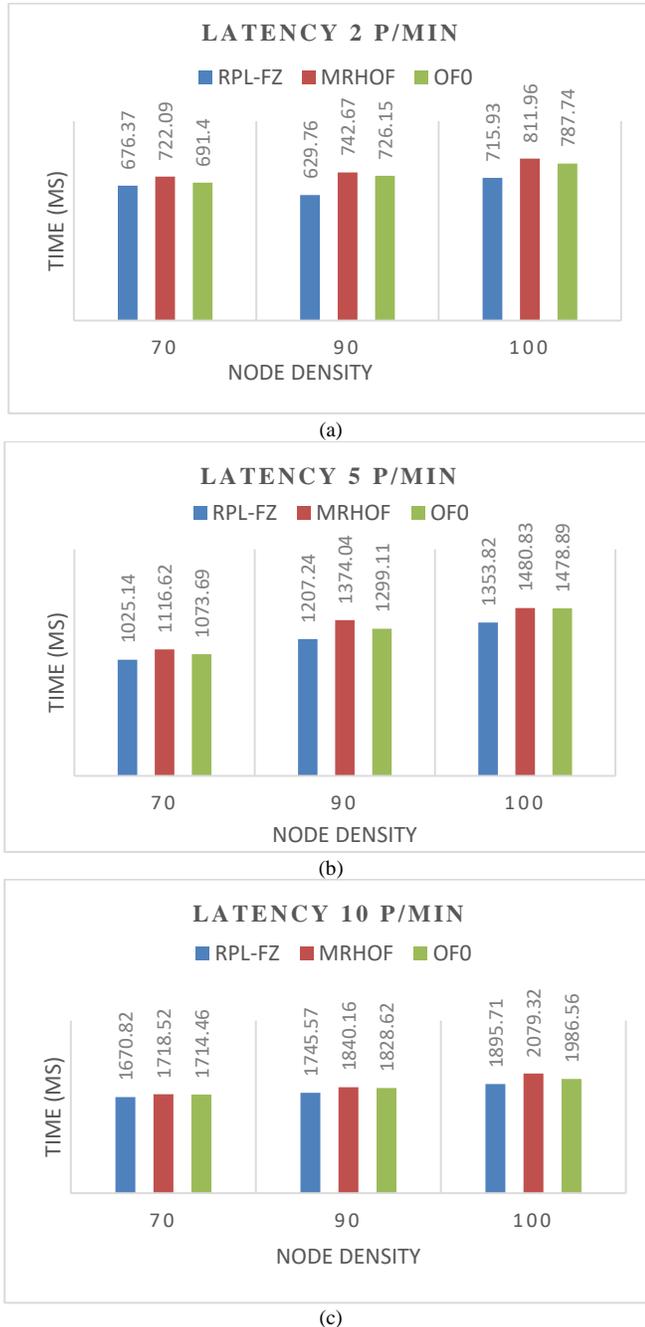


Fig. 8. Latency of RPL-FZ, MRHOF and OF0 for a) 2p/min b) 5p/min and c) 10 p/min.

V. CONCLUSION AND FUTURE SCOPE

In this paper, RPL-FZ, an efficient objective function for RPL-based Low power Lossy Networks is proposed. The existing objective functions defined in RPL standard depend only on a single metric to make the routing decisions and are not suited to meet the quality-of-service requirement of all applications. RPL-FZ combines three metrics (ETX, Delay and Residual Energy) by using the Fuzzy Logic process. The three metrics are combined using fuzzy inference system to evaluate the Quality of each neighbor node. The neighbor with highest Quality value will be selected as preferred parent node to route the data towards Root node. Thus, the sensed data will be efficiently delivered to the collection unit. RPL-FZ has been analyzed and its performance is compared with the objective functions MRHOF and OF0 specified in RPL standard. Simulations have been performed under varying network density and varying data transmission rates. The optimized objective Function RPL-FZ has 8% lower power consumption compared to MRHOF and OF0. Under strained conditions, at data rates of 10p/min, 100 nodes density and single sink node, RPL-FZ gives 7% higher PDR than MRHOF and OF0. RPL-FZ results in 8% lower Latency as compared to OF0 and MRHOF. The results show that RPL-FZ based RPL-implementation performs better in terms of Packet Delivery Ratio, Power Consumption, Network lifetime and Network Latency as compared to OF0 and MRHOF based networks.

This work provides scope for further research in which the Objective Function can take into account more metrics like Hop Count and Received Signal Strength Indicator to make the routing decisions. The performance of optimized Objective function can be tested on real-life test motes like Zolertia Z1 platform or TelosB motes.

REFERENCES

- [1] Akyildiz, I., Su W., Sankarasubramaniam Y. and Cayirci E. "A Survey on Sensor Networks", IEEE Communications Magazine, 2002.
- [2] Ed Winter, P. Thubert et al. "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks". In: Internet Engineering Task Force (IETF) 6550 (2012).
- [3] Ed P. Thubert. "Objective Function Zero for the Routing Protocol for Low-Power and Lossy Networks (RPL)". In: Internet Engineering Task Force (IETF) 6552 (2012).
- [4] O. Gnawali. "The Minimum Rank with Hysteresis Objective Function". In: Internet Engineering Task Force (IETF) 6719 (2012).
- [5] Qusai Q. Abuein, Muneer Bani Yassein, Mohammed Q. Shatnawi, Laith Bani-Yaseen, Omar Al-Omari, Moutaz Mehdawi and Hussien Altawssi, "Performance Evaluation of Routing Protocol (RPL) for Internet of Things", International Journal of Advanced Computer Science and Applications, Vol. 7, No. 7, 2016, pg 17-20.
- [6] S. Kuwelkar and H. G. Virani, "Performance Analysis of the Objective Functions in IPv6 Routing Protocol for LLN's used in IOT Applications," 2019 IEEE Pune Section International Conference, (PuneCon), Pune, India, 2019, pp. 1-5, doi: 10.1109/PuneCon46936.2019.9105896.
- [7] Puput Dani Prasetyo Adi, Akio Kitagawa, "Quality of Service and Power Consumption Optimization on the IEEE 802.15.4 Pulse Sensor Node based on Internet of Things", International Journal of Advanced Computer Science and Applications, Vol. 10, No. 5, 2019, pg 144-154.

- [8] Emilio Ancillotti, Raffaele Bruno, Marco Conti, "Reliable Data Delivery with the IETF Routing Protocol for Low-Power and Lossy Networks", IEEE Transactions on Industrial Informatics, Vol. 10, No. 3, August 2014.
- [9] Shubhangi Kharche, Sanjay Pawar, "Optimizing network lifetime and QoS in 6LoWPANs using deep neural networks", Computers & Electrical Engineering, Volume 87,2020,106775, ISSN 0045-7906, <https://doi.org/10.1016/j.compeleceng.2020.106775>.
- [10] Adeeb Saaidah, Omar Almomani, "An Efficient Design of RPL Objective Function for Routing in Internet of Things using Fuzzy Logic", International Journal of Advanced Computer Science and Applications, Vol. 10, No. 8, 2019, pg 184-190.
- [11] Mah Zaib Jamil, Danista Khan, Adeel Saleem, Kashif Mehmood, Atif Iqbal, "Comparative Performance Analysis of RPL for Low Power and Lossy Networks based on Different Objective Functions", International Journal of Advanced Computer Science and Applications, Vol. 10, No. 5, 2019, pg 183-190.
- [12] Hanane Lamaazia, Nabil Benamarb, "OF-EC: A novel energy consumption aware objective function for RPL based on fuzzy logic" Journal of Network and Computer Applications, Elsevier, Sept 2018. <https://doi.org/10.1016/j.jnca.2018.05.015>.
- [13] Bardia Safaei, Amir Mahdi Hosseini Monazzah, Alireza Ejlali, "ELITE: An Elaborated Cross-Layer RPL Objective Function to Achieve Energy Efficiency in Internet of Things Devices" IEEE INTERNET OF THINGS JOURNAL, July 2020.
- [14] Olfa Gaddour et al. "OF-FL: QoS-Aware Fuzzy Logic Objective Function for the RPL Routing Protocol". In: 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt) (2014).
- [15] Kechiche Ines, Bousnina Ines, S Abdelaziz, "A novel opportunistic Fuzzy logic based objective function for the Routing Protocol for Low-Power and Lossy Networks" 15th International Wireless Communications and Mobile Computing Conference, IWCMC 2019.
- [16] Shridhar Sanshi, Jaidhar CD, "Fuzzy optimised routing metric with mobility support for RPL", IET Communications, 2019, Vol. 13 Iss. 9, pp. 1253-1261.
- [17] A. S. Mamdani, E. H., "An experiment in linguistic synthesis with a fuzzy logic controller", International journal of man-machine studies, vol. 7, pp. 1-13, 1975.
- [18] Thomas Heide Clausen, Ulrich Herberg, Matthias Philipp. A Critical Evaluation of the "IPv6 Routing Protocol for Low Power and Lossy Networks" (RPL). [Research Report] RR-7633, INRIA. 2011.
- [19] "Contiki OS". In: <http://www.contiki-os.org/>.
- [20] "ContikiRPL". In: <https://github.com/contiki-os/contiki/tree/master/core/>.
- [21] Ainaz Bahramlou, Reza Javidan "Adaptive Timing Model for improving routing and data aggregation in IOT Networks using RPL" IET Journal, Jan 2018.
- [22] Ajay Kumar, Narayanan Hariharan, "DCRL-RPL: Dual context-based routing and load balancing in RPL for IoT networks", IET Communications, 2020, Vol. 14 Iss. 12, pp. 1869-1882.
- [23] Yad Tahir, Shusen Yang, Julie McCann, "BRPL: Backpressure RPL for High-Throughput and Mobile IoTs", IEEE Transactions on Mobile Computing, Vol. 17, NO. 1, Jan 2018.

Enhance Risks Management of Software Development Projects in Concurrent Multi-Projects Environment to Optimize Resources Allocation Decisions

Ibraheem M Alharbi¹, Adel A Alyoubi², Majid Altuwairiqi³, Mahmoud Abd Ellatif⁴

College of Business, University of Jeddah, SA^{1,2,4}

College of Computer and IT, Taif University, SA³

Faculty of Computers and AI, Helwan University, Egypt⁴

Abstract—In software development project management, Risk management represents critical knowledge and skills at the level of a single software project and at the enterprise level, which executes multiple software projects concurrently. The best decision of Risk management contributes to optimizing resource allocation at the enterprise level for achieving its goals. Therefore, the issue needs centralized risk management at the enterprise level as a whole and not for each project. Risk management is implemented through several stages and using different methods. Various studies deal with multiple aspects of software management. This research provides an analytical view of risk assessment in multi-environment software development projects that take place simultaneously. The study uses a public dataset previously used in previous research for several simultaneous projects in one organization. It describes the multi-software project's risks through 12 variables. A comparative analysis uses classification methods (Random Forest- TreesJ48 - REP Tree - Simple Logistic) to assess risks and put them in central view. The research experiment has proven high accuracy in determining risk levels in a multi-project environment, reaching approximately 98%, using the REP tree technique.

Keywords—Risk management; multiple software projects; risk assessment; software development projects

I. INTRODUCTION

The study, analysis, and management of software project risks remain essential in research and practice, despite all the attempts to provide solutions. Still, the reports monitor high failure rates in software development projects. Risk Management (RM) is one of the critical tasks facing project managers. The importance of risk management is due to labor laws and legislation as well as industry standards and internal guidelines for enterprises. The first principle of RM according to ISO 31000 standard is that “RM creates and protects value” [14].

The Standish Group Report presented a comparative analysis on projects from 2013 until the end of 2017. In this period, the three-class of project results (success, challenge, and failure) are illustrated in Fig. 1 [33]. It displays the differences between waterfall and agile. It assures that an agile project succeeds approximately 2X more than a waterfall and

1/3 less likely to fail. According to the Standish Group, 2018 Chaos Reports, an agile project has a 60% bigger success rate than waterfall projects. The main idea learned is that all software development projects face risks regardless of the development methodology.

Therefore, IEC 31010:2019 guides the selection and application of risk assessment techniques in a wide range of situations. Many standards focus on how to manage the activities of risk management in IT organizations. These standards illustrate how to address risks in organizations and companies. Various ISO standards target management systems such as quality perspectives in ISO 9001, IT Service Management (ITSM) in ISO/IEC 20000 - 1, project management in ISO 21500, and information security in ISO/IEC 27001. These IT - related and non - IT standards are significant for many companies [6].

Risk assessment and forecasting in the early stage of software project development is essential for risk management and improved success rate of software development projects. Therefore, organizations must assess and analyze the risks of software projects to be planned by more accurately identifying potential risks and adopting scientific risk mitigation methods. Software risk factors may be diverse and complex, and historical data may be uncertain and unregulated [19].

The International Organization for Standardization (ISO) has issued a standard risk management framework [14]. The main objective of ISO 31000 is to support enterprises in managing risk. ISO 31000 displays the principles, framework, and process of Risk Management. An organization can use it regardless of its level of size, activity, or sector. The definition of risk assessment in ISO 31000 consists of three sub-processes: identifying risk, analyzing risk, and evaluating risk. The sub-process identifying risk is used to determine, recognize, and define the affected risks on the project's objectives. The main aim of analyzing risk is to know the nature, causes, and sources of the identified risks for evaluation risk level. Evaluating risk aims to determine the acceptable level of risk and what should be tolerated, based on comparing risk analysis results with risk criteria [29] [14].

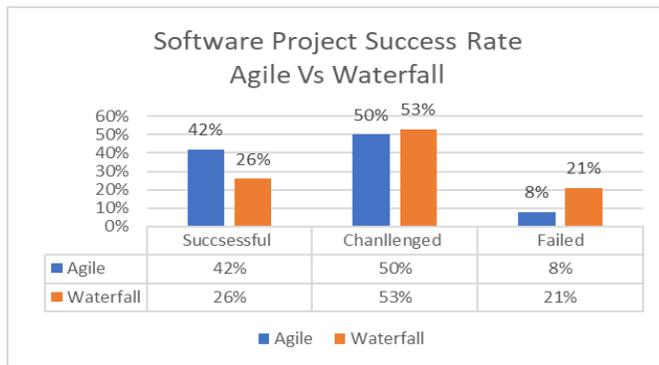


Fig. 1. Projects Success Rates (Adapted from [33]).

In NASA, enhance the safety decision process by using a Probabilistic Risk Assessment (PRA) methodology. It identifies Risks through the answer of three basic questions: the first is What can go wrong? , the second is, How likely is it? And the third is What are the consequences? These answers are prepared by systematically identifying, model, and quantify scenarios. These answers may be lead to knowing an undesired result. Qualitative and risk assessment requires quantitative semantic models. In addition, the probability theory is very descriptive and does not fit the calculation [37].

A simultaneous multi projects development environment is the common environment in software companies. Whereas rarely, software development company executes one project within the timeframe. However, various software projects are developed within the same time frame. Some of these projects consider developing government service portals, e-commerce portals, point-of-sale software, management information system software, enterprise resource planning systems, or other types of different software projects with various characteristics and features. So, the various projects are developed concurrently, and they are concurrent and sharing resources. Therefore, the management of the software development company needs to be concerned with each project restriction – scope, cost, schedule, quality, and relationships among these projects. In short, the main objective of project management, at the company level, should be expanded to improve resource allocation within the entire organization environment, not focus on each project individually.

Based on the previous paragraph, risk management in a multi-project environment has become more complex. Managers face difficult decisions on reducing the average delay per project or the time needed to complete the full range of projects using effectively allocated resources [16]. Risk management in the simultaneous software development environment will be more complex by the diversity of methodologies, policies, models, and procedures in the development process. Dealing with risk management in a single project is very different from a multi-project environment simultaneously. However, many IT organizations do not give much importance to risk management. Risk management is still subjective and often not uniform and does not take place with a systematic methodology.

Companies may need to adopt a centralized risk management methodology that helps maintain their market share through optimal resource management. They may lose if there is no proactive risk management. A supportive regulatory structure is also needed to achieve the central management of risky institutions to control risks throughout the organization.

The main objective of this paper is to clarify the relationships between the risk features and the level of risk and how risks are classified and assessed in simultaneous multiple software development projects. Therefore, this research tries to answer the following questions:

- 1) What are the relationships between risk level and risk features?
- 2) How to identify high-level risks across Simultaneous Multi-Projects?
- 3) How can the relationship between objective and subjective variables be determined and their impact on measuring the level of risk?

After the introduction section, the reminder structure of this paper consists of the following: Section 2 presents background and literature review to explain the main concepts and principles of risk Assessment in a Simultaneous Multi-Projects Development Environment. Section 3 displays the research experiment to answer the questions of the paper and explain the results. Finally, Section 4 consists of the conclusions extracted from research and future work representing expanded trends for current work.

II. BACKGROUND AND LITERATURE REVIEW

Many different studies have addressed risk management in software development projects. It has been subjected to various titles such as software development risks, software project risks, and IT projects risks. Some focus on completing a list of detailed points and risk points. Others include methods of analyzing and assessing levels of risks, including what deals with studies, statistical analysis, and methods of data mining.

Despite more than 30 years of extensive research into the analysis and management of the risks of software development projects, this has led to unified guidance on software project risk management. Adopting these risk management methods in practice still needs more attention so that software projects can be submitted in accordance with the required performance standards. Therefore, this still faces huge challenges [32].

There are six foundations for risk management. It includes risk management as modeling; rational process, factor analysis, capability, social process; and data analysis [5].

By studying more than 100 projects by [23] they found relationships between risk management practice and success in achieving projects in a timely and budgetary manner. They conclude that higher levels of risk management in project management achieve higher levels of project success [23].

In short, risk management is a critical success factor in the implementation of a successful project by identifying and eliminating project risk levels. Senior management supports risk management, the actual practice of risk management practices, and regular risk monitoring [35]. Therefore, the success of the project can occur more frequently [36].

The success of software projects is a subjective process that may vary from project to project and is linked to many risks. Project managers may differ because of these risks depending on the demographic characteristics of managers and the project's nature and features. This study attempted to identify and measure software risks' dimensions and analyze differences in perception among project managers about software risks [26].

Based on a research case study, de Packer and others concluded that risk management activities contribute to the project's success by creating a shared vision of the project with beneficial communication and direct effects [4].

The [17]study has shown a relationship between risk management, success, and self-performance of IT projects. Good risk management can lead to project success only if risks are identified and controlled before the project begins [17].

In a study conducted for measuring risk management impact on IT projects' success, researchers collected data from 200 IT project managers and used statistical analytical methods to measure this effect. The study found that there was a positive impact of risk management on IT projects' success [21].

High-tech and complex projects are more vulnerable to resource risks. Scope and resource risks are also vital risk categories for high-end, high-tech, and medium-complex projects [43].

According to [16] study, one of the complex challenges facing the multi-project environment is allocating resources between projects, especially if resources are minimal. Therefore, Project managers have to make resource allocation decisions to complete the project package in time or reduce project delays. Therefore, multi-project management is a growing area of research [16].

In a recent study, the researcher presented a viable model for monitoring software development projects in international companies. This solution is based on earned value management EVM and Monte Carlo as a simulation method. The proposed model in this study has provided proactive risk management that meets the project's objectives and adapts to changing resource needs [20].

The researchers presented a study on the analysis of risk assessment elements in a multi-project environment. Researchers used some Data Mining methods to reach the elements influencing risk assessment in the multi-project environment and suggested conducting further study and analysis on this topic [38].

A. Risk in Software Development Projects

According to the PMI definition, there are two levels for each project: individual Risks and overall risk. Individual risks can affect project objectives positively or negatively if they occur [22]. Risk refers to the factors that influence the project's success or failure. Also, agile principles include some of the principles to support risk management [18]:

- 1) Customer satisfaction has the highest priority through all phases of the project to deliver valuable software.
- 2) They welcome changing requirements until late in development, as agile's operations contribute to change in favor of customer's competitive advantage.
- 3) Developers and Businesspeople should work as a team throughout all phases of the project.

Previous principles and any best practice supporting applying those principles eliminate or mitigate the negative impact of risks (threats) or promote positive risks (opportunities) that often lead to project success, challenges, or failure [19].

Although one of the new development methodology goals, such as agile, is to reduce the harmful risks and maximize the positive opportunities for software projects' success. In contrast, three principles of agile mentioned in the previous paragraph prevent the occurrence of causes and causes that affect customer satisfaction and the quality of software products. But one of the researchers observed that the final elimination of risks is complicated. Still, there are clear differences between development software projects by a traditional methodology and projects with an agile methodology.

Fig. 2 displays a comparison of risk through the project between waterfall and agile projects. The figure illustrates that all projects have a risk regardless of the development methodology of the software project. Besides, the figure displays that agile projects have a declining risk model wherein agile projects risk declines as the project progresses, whereas in waterfall project has increased risk model. with agile project management, the days of catastrophic failure decline with the project's progress. The elimination of large-scale failure is the biggest difference between risk on a waterfall and agile projects [33].

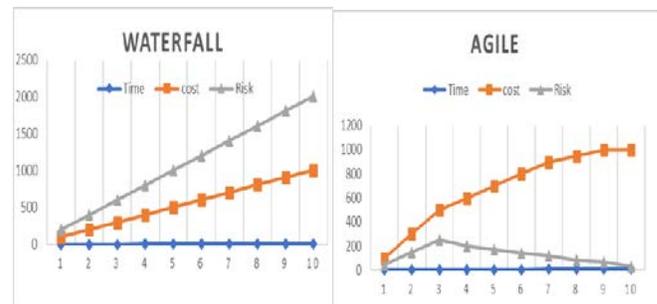


Fig. 2. (a) Comparison of Risk and Cost between Waterfall and Agile Projects (Adapted from [33]).

B. Risk Management in Simultaneous Multi-Projects Development Environment

In most software companies, multiple software development projects are executing concurrently. The cause of existing projects is related to achieve the strategic goals of the enterprise. And there is a need to manage multiple projects together; therefore, critical success factors must be analyzed and worked in an isolated manner and an entire environment, including all parallel projects. Resource management is one of the most important aspects of a multi-project management environment because it requires a participatory vision that contributes to resource planning and coordination, allocation, and balance of projects [12].

Simultaneous projects share human and material resources, so project managers must consider the constraints imposed on each project and between projects to succeed. Risk management is addressed at all stages of the project and explicitly addresses scope, schedule, budget, and quality. In short, if resources are not managed perfectly, projects may face some challenges that could lead to project failure [39].

The author in [11] acknowledged that the comprehensive framework for risk management in multi-project environments (MPE) is still missing. This framework captures and manages the overall characteristics of MPE. The framework identifies adaptation variables that affect relationships and their risk outcomes and a core group of mediators and supervisors. [11] Risk management during the life cycle of the project has become more important and necessary. It is the best way to visualize project critical success factors entirely, support resource allocation decisions, and solve problems more efficiently. Therefore, academic researchers and industrial practitioners have still had more interest in studying risk management in recent years [39].

Multi-project management is currently a reality in software development environments. Continuous product changes or range levels characterize software development projects. The program development process is complex, particularly in human resources management, such as knowledge and technical expertise. These characteristics may be risk factors that need to be managed. Tactical management is therefore required to use information in an orderly manner, which leads us to consider using a metric-based strategy as a support tool for multiple project managers focusing on risk factors. The researchers, therefore, suggested applying the "risk points" scale and its diversity in the multi-software development project environment as a tool to support the decision and monitor risks during the project life cycle [3].

It is common for many companies to manage project risk separately and not to monitor and analyze the risk impact between projects across the organization. However, the prevailing situation is a simultaneous and shared project resource environment, so the main essence of the decision-making process is communication between partners involved in the organization's shared resources. Communication must be organized vertically and horizontally, and this is the meaning of participatory management [30].

Much of the literature ignores that the risks can be shared among different simultaneous development projects in the enterprise. At the same time, the risks in all levels of management and the relationships among risks are high. So, it would be appropriate to manage the risk from a holistic perspective. So, to better understand, consider the following toe levels of Risk Management [39]:

1) *Risk management of one project:* The enterprise has one project; the aim is to manage related risks to ensure the project's success.

2) *Risk management of multiple projects:* The enterprise has the risk management of a single project in addition to the common risks in the multiple projects' environment. So, the enterprise has a matrix of project risks. The aim is to ensure project success, system optimization, and the allocation of resources and analyze the impact of resource allocation on multiple projects.

The author in [10] illustrate those risks must be managed when managing multiple projects. With the organizational structure and the appropriate software, you can also share and monitor risks in all projects. At any time, the large organization implements different types of multiple projects at the same time. These projects have different sizes and stages of completion. In this environment, risk management can face a variety of challenges and bring significant benefits, some of which are [10]:

- 1) Cost savings by identifying overall risk mitigation.
- 2) Risk vision varied in multiple projects.
- 3) Reducing emergency budgets.
- 4) Effective distribution of resources.

C. Assessing the Impact of Risks and Prioritizing Software Development Projects

As shown in Fig. 3, assessing the impact of risks and prioritizing risks to be dealt with are outlined in the second and third steps of the risk management process. This assessment usually shows how cost or technical performance can affect the achievement of targets. However, the implications of these standards are not limited; It is also necessary to consider the political or economic consequences. The likelihood of any threat that may adversely or positively affect the project is also assessed. This assessment often involves the use of self-assessment techniques, especially if it is difficult to use objective methods to assess the possibilities directly. (i.e., engineering, analysis, modelling, and simulation) [8].

According to IEEE standard 729, a software requirement includes three types that display in Fig. 4:

Functional requirements: these are requirements that are requested by the end-user specifically standard should be provided by the system. Since non-functional requirements including scalability, security, maintainability, reliability, scalability, performance, reusability, and flexibility. Area requirements: area requirements are those of a particular category or area of the project [13]. Table I displays some software requirements standards and studies.

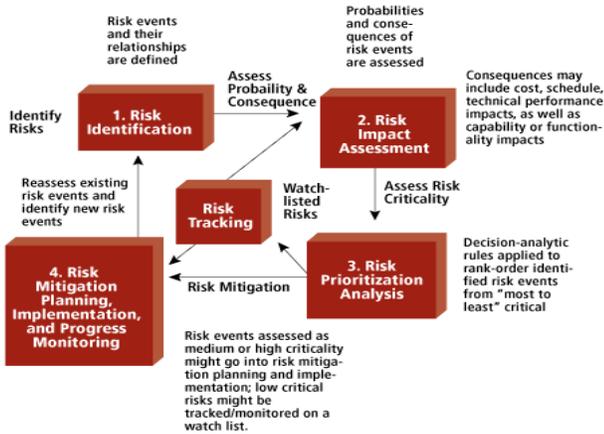


Fig. 3. The Fundamental Steps of Risk Management [8].

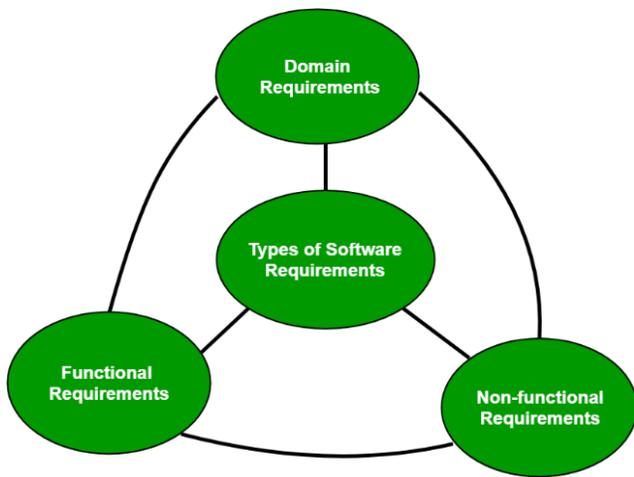


Fig. 4. Software Requirements [13].

TABLE I. SOFTWARE REQUIREMENTS

Study/standard	Software Requirements
[45] ISO/IEC 9126	"Portability" has "co-existence", "Maintainability", "Functionality", "Security", "Efficiency", "Reliability", "Usability", and "Compatibility".
[44] IEEE standard 729	Functional Requirements, Non-functional requirements (Flexibility, Scalability, Reusability Portability, Reliability Security, Performance Maintainability, and. Domain requirements)
[41]	Functional requirements (FR), Operational requirements (OR), Technical requirements (TCR), and Transitional requirements (TSR)
[27]	Security, Interfaces, Functional, Supportability, Usability, Standards, Reliability & Availability, Constraints, Performance, Safety.
[15] ISO/IEC/IEEE29148:2018	Business Requirements Specification (BRS), System Operational Concept (OpsCon), Stakeholder Requirements Specification (StRS), Software Requirements Specification (SRS) System Requirements Specification (SyRS),

According to the study conducted by [1], he found that the most frequent software risks are planning and controlling user requirements, risks associated with the project team, risks related to the organizational environment, and project complexity [1]. Another previous study by [25] showed that the risks of the software project were interrelated. the risk dimensions include technology experience, project size, and project structure used in the system requirement, planning and control, project complexity, team, and organizational environment [25]. Table II displays some software risk dimensions studies.

TABLE II. SUMMARIZES SOME OF THE STUDIES ON SOFTWARE RISK DIMENSIONS

Study	Year	No	Software Risks Dimensions
[40]	1993	3	Development Environment Product Engineering Program Constraints
[28]	1995	3	Technical, Organizational, And Environmental,
[9]	2002	3	Management (Customer & Stakeholders, Corporate), Technology (Performance, Application, Requirements) External (Culture, Natural Environment, Economics),
[34]	2004	4	Requirements - User - Planning & Control - Project Complexity
[1]	2011	5	Organizational Environment, Team, Planning and Control, User Requirements, and Project Complexity
[25]	2011	9	Success Risk, External Risk, Financial Risk, Technology Risk, Management Risk, Business Processes Risk, Security Risk, People Risk, Information Risk, And
[7]	2013	9	Organizational Environment, Project Size, Planning, And Control, Team, Technology Experience, Project Complexity, System Requirement, Project Structure, User,
[39]	2013	11	Cost, Time, Quality, Team, Risk, Size, Changes, Complexity, Design, Organizational
[31]	2014	6	Resource Risk, Technical Risk, Personnel Risk, Legal Risk, Financial Risk, Management Risk, Contractual and Other.
[22]	2017	4	External Risk- Technical Risk -Commercial Risk- Management Risk-
[27]	2018	14	Schedule, Estimations, Organizational Requirements, Requirements, Project Cost, Complexity, User, Project Complexity, Software Requirement., Team, Planning and Control, Organizational Environment,

D. Research Experiment

In this part of the research, the aim is to answer research questions through the results that will be reached in this experiment. One of the data mining tools that will be used as a tool for data mining is Weka 3.8.4. Therefore, this section of the research is divided into five subsections:

- 1) Data set.
- 2) What are the relationships between Risk level and Risk Features?
- 3) How to Identify top risks across Simultaneous Multi-Projects?

- 4) Identify the association between the subjective and objective variables and their impact on measuring risk levels?
- 5) Overall comment on the search experience.

E. Data Set

In this study, the researchers used the data set in [27]. This data set includes 12 variables describing the risks of several projects for software development, some of which are project type, requirements, impact, probabilities, and priority risks where it is fully presented in Table III, The researcher based the challenge of the listed variables on several methods [27]:

- 1) Previous studies, especially studies [34].
- 2) The experts have at least five years of experience in the software industry companies that achieve the second and third levels in the maturity and ability model.
- 3) The number of cases in the dataset consists of 299 instances. the dataset is formatted in Attribute-Relation File Format (ARFF) [27].

F. Answer to Question 1: What are the Relationships between Risk Level and Risk Features?

The first question in this paper is the relationship among the risk characteristics and assess the risk levels. Determining this relationship is very important for identifying the traits with the most significant effect. Recognizing this relationship contributes to forming the risk record and risk dataset to help manage projects centrally because the appropriate volume of data is the target, neither more nor fewer. Collecting too much data represents a cost and effort lost and may distract the decision-maker and collecting too little data is insufficient.

Attribute reduction is an important area of research in machine learning, artificial intelligence, and pattern recognition to improve the performance and economics of predictions. Attribute reduction participates in identifying redundant attributes to remove this redundancy and improve decision-making [24]. So, attribute reduction represents an essential step in the preprocessor of data mining, machine learning tasks, and data pattern recognition [2].

Many factors affect the risk assessment of the software development project. These factors are divided into two types, subjective and objective. Factors that are objectively measured and objectively described are not affected by the personality of the person who counts them. In contrast, other variables describing project risk are subjective values, such as the impact, magnitude, likelihood of occurrence, and frequency of risk, and are influenced by the character of the decision-maker.

The experiments that were conducted in this research where use many methods for feature selecting. Using more than one feature selection method confirms results to solve the feature's problem. The research seeks to identify the characteristics with the most significant impact from all methods. Also, The second sub-objective in this research is what the features don't relate to assessing the risk level based on many methods. Clarifying this issue is of great importance

to decision-makers in designing the software company's risk register. The risk register is the appropriate tool for compiling project risk data, followed by analyzing these data to help the analysis results formulate proper strategies for efficiently managing software development risks. Table IV displays The results.

Six methods were used in the research experiment to determine which features are more important in risk assessment. The six methods then unanimously agreed on the importance of three features: priority, probability, risk size. Fig. 5 shows that although the relationship between the magnitude of the risk and the level of risk is substantial. In some cases, the relationship level may decrease due to other factors that may be requirements and impact, like what's shown in shape.

On the other hand, by the relationship thread, three methods confirmed no significant relationship between two features(Affecting No of Modules, Fix Cost (% of the project), and the risk level assessment.

These features (Affecting No of Modules and Fix Cost) have nothing to do with a risk assessment, which does not mean they are unimportant. However, their importance still exists as necessary information in managing the risks of the software project and not in determining the level of risk.

TABLE III. THE STRUCTURE OF DATA SET

Attributes	Value Sets
Requirements	Text to describe the requirements
Project Category (PC.)	1. Transaction Processing System (TPS) 2. Management Information System (MIS) 3. Enterprise resources planning System (ERP) 4. Safety-Critical System (SCS)
Requirement Category(RC.)	Reliability & Availability— Usability- Security- Safety - Supportability - Functional - Constraints, Interfaces - Standards- Performance
Risk Target Category (RTC)	Budget - Personal -Resource Availability-Quality-Schedule -Performance-Functional Validity-People -User- Requirement - Dimension- Time - Cost Design- Project complexity- Unrealistic Requirements Business - Overdrawn Budget-Process – Software - Planning & Control – Team- Organizational Environment-
Probability(Pro)	0 % -100 %
Magnitude of Risk(MoR)	Extreme - Very High – High - Medium - Very Low – Low- Negligible
Impact(Imp)	Catastrophic, high – moderate- Low- insignificant
Risk Dimensions (RD.)	Requirements- Estimations -Project complexity- planning and control- Project cost - Team, Planning, and Control – User - Organizational Environment- Software Requirement - Schedule - Organizational Requirements - Complexity,
Affecting No Modules(ANM)	Numbers of Modules
Fixing Duration(FD)	Period by Day
Cost of Fixing(CoF)	% project cost
Priority(Pri)	0 % -100 %
Risk Level(RL)	1- 2- 3 - 4 - 5

TABLE IV. THE RESULT OF ATTRIBUTE REDUCTION

Relief F Attribute Eval :	Info Gain Attribute Eval
Ranked attributes:	Ranked attributes:
0.24334 4 Pro	1.9085 11 Pri
0.23648 11 Pri	1.474 4 Pro
0.21339 5 MoR	0.4605 5 MoR
0.12029 1 PC	0.2892 3 RTC
0.07439 3 RTC	0.197 7 DoR
0.05351 2 RC	0.1936 1 PC
0.03673 7 DoR	0.1264 9 FD (Days)
0.02074 9 FD (Days)	0.1145 2 RC
0.01374 8 ANoM	0.0448 6 Impact
0.00289 10 FC (% of Project)	0 8 ANoM
-0.00139 6 Impact	0 10 FC (% of Project)
Correlation Attribute Eval	Gain Ratio Attribute Eval
Ranked attributes:	Ranked attributes:
0.4563 11 Pri	0.9611 11 Pri
0.4518 4 Pro	0.6107 4 Pro
0.1964 5 MoR	0.1994 5 MoR
0.1734 9 FD (Days)	0.1407 9 FD (Days)
0.1328 1 PC	0.1034 1 PC
0.11 8 ANoM	0.0753 3 RTC
0.078 3 RTC	0.0586 7 DoR
0.0672 10 FC (% of Project)	0.0535 2 RC
0.0615 2 RC	0.0228 6 Impact
0.0522 7 DoR	0 8 ANoM
0.0481 6 Impact	0 10 FC (% of Project)
One R Attribute Eval	Symmetrical Uncert Attribute Eval
Ranked attributes:	Ranked attributes:
98.662 11 Pri	0.9636 11 Pri
86.288 4 Pro	0.6717 4 Pro
55.853 5 MoR	0.2149 5 MoR
46.154 8 ANoM	0.1006 1 PC
44.482 6 Impact	0.0994 3 RTC
44.147 9 FD (Days)	0.088 9 FD (Days)
43.813 2 RC	0.0738 7 DoR
43.144 10 FC (% of Project)	0.0557 2 RC
43.144 7 DoR	0.0227 6 Impact
42.809 3 RTC	0 8 ANoM
42.809 1 PC	0 10 FC (% of Project)

On the other hand, the other six characteristics: project category, risk target category, risk requirements, and risk dimensions, duration of determination (days), and impact on various centers impact risk assessment in software projects.

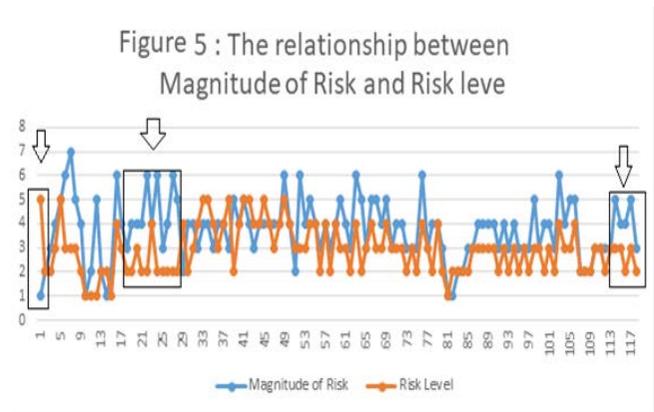


Fig. 5. The Relationship between the Magnitude of Risk and Risk Level.

G. The Answer of Question 2: How to Identify Top Risks Across Simultaneous Multi-Projects?

The systemic risk is the risk that often occurs throughout the organization. The treatment of risk is better to deal with when using a centralized database, although the project manager tries to search through many different project risk records in spreadsheets. Because these risks are in a multi-project environment, the project manager may be allowed to focus on his project. However, preparing the emergency budget for systemic risks is the responsibility of the program manager. Central risk management is a better methodology for reducing overall risk costs [10].

Identifying the "top 10" risks across multiple projects is a common requirement to monitor the most critical risks based on Central risk management methodology. It requires each project manager to provide information on the most significant risks to their project. However, this may lead to an error assessment of risk, as the fourth risk in one project may be more influential than the second risk in another. This situation does not mean the wrong risk assessment only but also may lead to the concealment of a great deal of potential impact, as shown in Fig. 5 [10].

The project results for the project differ from the risk data analysis for simultaneous enterprise-wide projects. Some managers may extract the results of a single project on its scale and then combine the results of project risk analysis in a single scenario [10]. This approach may face some criticism when dealing with this problem such as:

- 1) It is not commensurate with the dynamic nature and management of projects.
- 2) It may lack integrated centralized management to achieve a unified enterprise-wide vision.

Therefore, the art of research focuses on the central management of all project data together.

The graphic forms illustrate this Fig. 6 representing a static image at a particular moment in the life of the simultaneous development environment for the four projects included in the data upon which the research relied. The state

of the risks with the highest priority may differ from the view of Collective centralization from individual mono view at the level of a single project. The development of software in the environment of software companies does not take place in isolated islands; but instead, there is a sharing of resources and goals.

Fig. 7 shows the relationship between some subjective measures of risk and the level of risk. Self-assessment traits are affected by a person's assessment, such as the probability and magnitude of the risk. The figure shows that in some cases, as shown in the form, the estimated value is that the risk is high, while at the company level, this risk is at a lower level than other risks to projects that coincide with the same project.

The unique view of the project manager who focuses on a single project may differ, as shown in Fig. 7, from the holistic view, which deals with multiple simultaneous projects as explained in previous Fig. 6. Fig. 7 focuses on Transaction Processing System Projects, where it deals with it independently and then seeks to deal with the most significant threat affecting the project level and allocates resources to it that contribute to its treatment. This evaluation is a minor view because it is not the only project in the organization. After all, there may be a risk associated with another project with a more significant impact on the organization level. If it is not addressed and allocated resources, it may cause more harm to the software company's reputation.

The overall view of simultaneous projects contributes to formulating a unified picture of project risks at the enterprise level. Integrated risk management brings many benefits based on addressing threats that significantly impact the entire company for cost savings. It also delivers competitive advantages, enhances opportunities to achieve strategic and operational objectives, thereby reaping the greatest benefits and improving operational efficiency.

By conducting a comparative analysis of the classification of risk levels in a multi-project environment using the following four methods. Table V presents the results of the comparison process, showing the levels of accuracy achieved:

- 1) Random Forest
- 2) REP Tree
- 3) Trees J48
- 4) Simple Logistic.

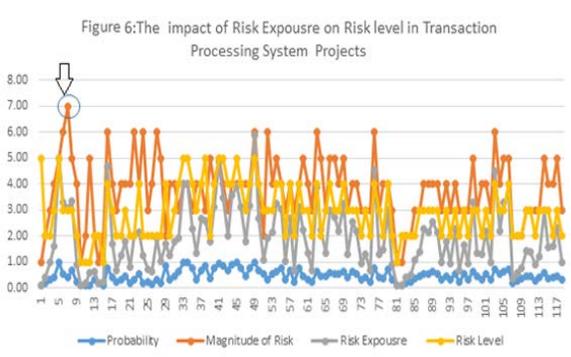


Fig. 6. The Impact of Risk Exposure on Risk Level in Transaction Processing System Projects(Concurrent).

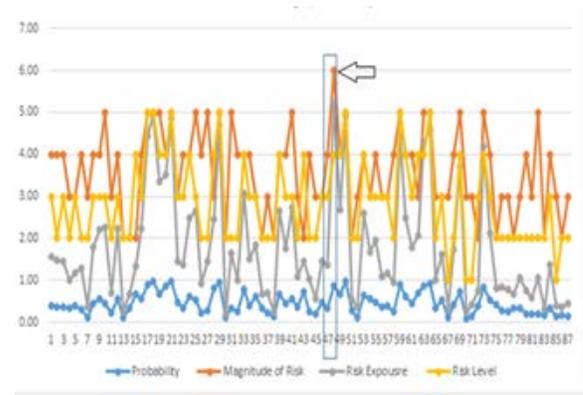


Fig. 7. The Impact of Risk Exposure on Risk Level in Transaction Processing System Projects.

TABLE V. SIMPLE LOGISTIC: CLASSES

Class 1:
207.79 + [RTC=Quality] * 1.28 + [MoR=Negligible] * 0.44 + [Pri] * -10.07
Class 2:
21.72 + [RTC=User] * -1.09 + [Pro] * -0.05 + [MoR=Low] * 0.64 + [MoR =High] * 1.05 + [Impact=Low] * -0.57 + [Pri] * -0.51
Class 3:
-0.49 + [RC=Performance] * -0.86 + [RTC=Budget] * 0.76 + [RTC=Performance] * 0.39 + [RTC=User] * 0.65 + [RTC=Requirement] * -0.67 + [RTC=Organizational Environment] * 3.03 + [RTC=Design] * -0.94 + [RTC=Unrealistic Requirements] * -1.31 + [MoR=Very Low] * -1.35 + [MoR =Low] * 0.36 + [MoR =Medium] * 0.77 + [MoR =Extreme] * 1.17 + [Impact=Low] * 0.41 + [DoR=Requirements] * -0.79 + [DoR =User] * -0.82 + [DoR =Software Requirement] * -0.57 + [DoR =Planning and Control] * -0.6 + [DoR =Cost] * 0.78
Class 4:
-28.32 + [RC=Constraints] * -0.72 + [RTC=Performance] * -0.64 + [RTC=Functional Validity] * 0.88 + [Pro] * 0.04 + [MoR=Medium] * 0.65 + [Impact=insignificant] * -1.14 + [FD (Days)] * 0.11 + [Pri] * 0.42
Class 5:
-154.64 + [RC=Security] * -1.69 + [RC=Constraints] * 1.34 + [RTC=Functional Validity] * -1.69 + [Impact=Low] * -3 + [Pri] * 2.04

There are many ways to classify risk levels in individual and multiple enterprise environments as shown in Table V. Initially, kappa's statistical value is between 0.9034 and 0.9713, and Kappa is used to test the reliability between the rate of the model. Classification methods have an almost degree of accuracy, with accuracy coming from 93% using a simple logistics method to approximately 98% using the REP tree. While error metrics are at very low levels, ranging from 0.0371 to 0.0126, to measure the absolute average error (MAE) and to measure the proximity of final results to projections and the value of 0.1488 to 0.0888 for the average root square error scale (RMSE). RMSE represents the standard deviation of differences between expected values and observed values.

Table V cells display a forecast model to predict risk levels according to a simple logistical method. Each of the five cells describes the characteristics of each category of risk levels from level 1 to level 5. the extract results from Wika 3.8.4.

After review, the simple logistical forecast model achieved the lowest level of accuracy among the classification methods used in the experiment. The best predictive models were the two methods of forecasting (J48 and REP Tree) in the experiment of this research by displaying the classification decision tree in these two ways. J48 algorithm deals with noise, Not much affected by lost data, and deals with continuous features. See Fig. 8, 9,10, and 11.

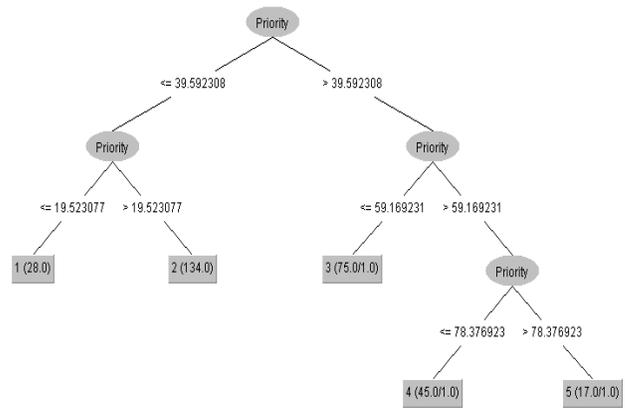


Fig. 9. J48 Tree of Classification.

```

==== Run information ====
REP Tree
=====
Priority < 39.75
| Priority < 19.83: 1 (18/0) [10/0]
| Priority >= 19.83: 2 (90/0) [44/0]
Priority >= 39.75
| Priority < 59.73: 3 (49/0) [26/1]
| Priority >= 59.73
|| Priority < 79.34: 4 (31/1) [14/0]
|| Priority >= 79.34: 5 (11/0) [6/1]
Size of the tree: 9
    
```

Fig. 10. REP Tree.

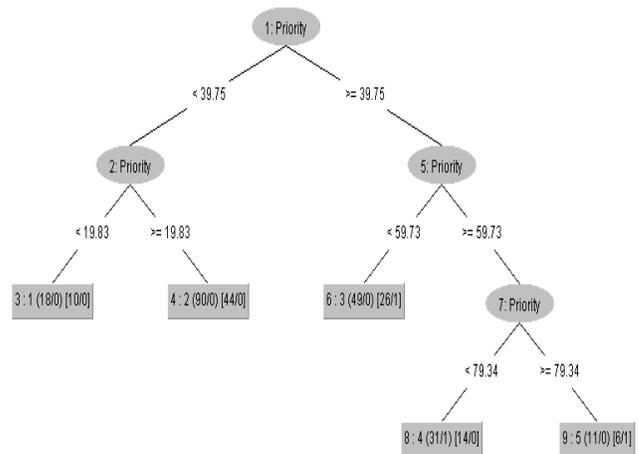


Fig. 11. REP Tree Structure Graph.

```

J48 pruned tree
-----
Priority <= 39.592308
| Priority <= 19.523077: 1 (28.0)
| Priority > 19.523077: 2 (134.0)
Priority > 39.592308
| Priority <= 59.169231: 3 (75.0/1.0)
| Priority > 59.169231
|| Priority <= 78.376923: 4 (45.0/1.0)
|| Priority > 78.376923: 5 (17.0/1.0)
Number of Leaves: 5
Size of the tree: 9
    
```

Fig. 8. J48 Pruned Tree.

Authors extract from the decision tree in the two methods; see Fig. 9,10 and 11. The main feature in the risk level classification is the priority, with rating values varying between evaluation methods. The risk priority feature represents a personal judgment of the project manager, who, according to the characteristics of each risk, determines the risk priority among the risks of simultaneous projects run by the company. In this regard, the experience and skills of project managers are critical to risk management in a multi-project environment. Therefore This requires a more in-depth and detailed study.

According to [42], J48 and REP Tree produce high classification and simple tree structure. The Reduced Error Pruning (REP) Tree in Fig. 10 and 11 show that the main factor for risk assessment is the priority of risk, which is the same result reached using the J48 tree displayed in Fig. 9, but there are some differences in digital values.

H. The Answer of Question 3: Identify the Association between the Objective and Subjective Variables and their Effect on Measuring the Level of Risk?

It is important to analyze the relationship between subjective and objective variables in assessing the risks of software development projects. Table VI shows the degree of correlation between these variables according to the data available at this research stage. It is noted from the results of the relationship analysis that there is a strong correlation between Priority, 'Risk Level', and Probability. At the same time, there is a weak correlation between 'Fixing Duration (Days), and others.

TABLE VI. PEARSON'S CORRELATIONS

Variable		Probability	'Fixing Duration (Days)'	Priority	'Risk Level'	Affecting No of Modules'
1. Probability	Pearson's r	—				
	p-value	—				
2. 'Fixing Duration (Days)'	Pearson's r	0.301	—			
	p-value	< .001	—			
3. Priority	Pearson's r	0.962	0.367	—		
	p-value	< .001	< .001	—		
4. 'Risk Level'	Pearson's r	0.923	0.356	0.958	—	
	p-value	< .001	< .001	< .001	—	
5. Affecting No of Modules'	Pearson's r	0.116	0.127	0.229	0.229	
	p-value	0.046	0.028	< .001	< .001	

I. Overall Comment on the Search Experience

It is clear from the research experience generally that the usual environment in software companies is a multi-project environment at the same time. Therefore, it requires an integrated management methodology to manage the risks of these projects together to achieve the company's general goals and then maintain its competitive capabilities and preserve its market share. This management methodology should include quantitative methods that help collect, analyze, and measure risk levels individually at the level of one project and collectively at the level of several projects that are developed simultaneously.

Project managers should also use data mining methods and statistical analysis techniques to reach knowledge levels that help allocate resources more efficiently. They must also use visual data presentation methods to determine the strengths and weaknesses that help them achieve critical success factors for software projects.

III. CONCLUSION AND FUTURE WORKS

At the end of this work, the following paragraphs summarize the findings of the results obtained and present some trends that need further study and analysis.

A. Conclusion

The central management of synchronous software projects gives a clear, integrated picture of the risk levels at the enterprise level. Data mining methods help identify the most important characteristics that must be recorded in the risk records of the enterprise database, where the study showed that the 12 characteristics included in the research data are important. This study has concluded with several important points, the most important of which are the following:

1) Priority, Probability, and Magnitude of Risk represent the main features to classify risk levels, where the six feature reduction methods in experimental ensure that.

2) The central unified risk vision of multiple simultaneous projects gives management an integrated picture to help it

allocate resources in a more efficient way that contributes to the company's overall objectives.

3) The classification methods (Simple Logistics, REP tree, trees, J48, and Random Forest) used in the research experiment contributed to high results of reliability and accuracy reached 98% and low levels of classification error measures.

4) The priority of risk is the most important factor in classification according to the best methods in the research experiment, with different percentages.

a) Relying on an integrated management methodology for risk management in a concurrent project environment helps identify the most influential risks at the company level rather than at the individual project level. Therefore, This contributes to the more efficient allocation of resources to achieve the company's overall objectives.

B. Future Work

The following points represent some of the proposed future trends for research and study in this area:

1) Conduct a more in-depth analysis of the relationship between subjective and objective variables considering the reciprocated impact on risk measurement.

2) Risk level analysis in light of multiple software development methodologies.

3) Measure the level of individual and group experiences of the project managers in measuring risk levels.

4) Analyze and study the relationship between risk characteristics in a multi-project environment and company characteristics such as size, age, organizational structure, etc.

5) Identify a model to measure the impact of knowledge Management on Risk Management in software development companies.

6) What are the impacts of Blockchain technology on concurrent project management in an overseas company? Does it reduce the risks, support fixing Mistakes, to allocate resources efficiently?

REFERENCES

- [1] Amuphaptrairong, T.,2011, Top Ten Lists of Software Project Risks: Evidence from the Literature Survey. Proceedings of the International MultiConference of Engineers and Computer Scientists, 1. Hong Kong.
- [2] Bai, J., Xia, K., Lin, Y., & Wu, P, 2017, Attribute Reduction Based on Consistent Covering Rough Set and Its Application, Hindawi Complexity Volume 2017, Article ID 8986917, 9 pages.
- [3] Jr, J., Wanderley M., Gusmão C. & Moura H., 2016, Application of Metrics for Risk Management in Environment of Multiple Software Development Projects. In Proceedings of the 18th International Conference on Enterprise Information Systems - Volume 1: ICEIS, ISBN 978-989-758-187-8, pages 504-511. DOI: 10.5220/0005859705040511.
- [4] De Bakker, K. d., Boonstra, A., & Wortmann H.,2010, Does risk management contribute to IT project success? A meta-analysis of empirical evidence, International Journal of Project Management 28 (2010) 493–503.
- [5] Bannerman, P. L. (2015). A reassessment of risk management in software projects. In: Handbook on Project Management and scheduling, vol. 2 (pp. 1119–1134). Springer International Publishing.
- [6] Barafort, B., Mesquida, A., & Mas A.,2019 "ISO 31000-based integrated risk management process assessment model for IT organizations", Journal of Software: Evolution and Process, 2019.

- [7] Chawan, P. M., Patil, J., & Naik, R., 2013. Software Risk Management. *International Journal of Advances in Engineering Sciences*, 3(1), 17-21.
- [8] Garvey, P.R., 2008, *Analytical Methods for Risk Management: A Systems Engineering Perspective*, Chapman-Hall/CRC Press, Taylor & Francis Group (UK), Boca Raton, London, New York, ISBN: 1584886374.
- [9] Hall, D.C., & Hulett D.T. 2002. Universal Risk Project—Final report. Available from the PMI Risk SIG website www.risksig.com, or www.techriskmgt.com/home2link.html.
- [10] Han, Wen-Ming, & Huang S., 2006. "An empirical analysis of risk components and performance on software projects." *The Journal of Systems and Software*, 2006: 42-50.
- [11] Hashim, N.I., Chileshe, N., and Baroudi, B., 2012, "Management challenges within multiple project environments: Lessons for developing countries, *Australasian Journal of Construction Economics and Building*, Conference Series, 1 (2) 21-31.
- [12] Hopkins, R., Campbell, K., O'Reilly, D., Tarride, J., Bowen, J., Blackhouse, G., & Goerre, R. (2007). *Managing Multiple Projects: A Literature Review of Setting Priorities and a Pilot Survey of Healthcare Researchers in an Academic Setting*. Perspectives in health information management / AHIMA, American Health Information Management Association. 4. 4.
- [13] Geeksforgeeks.org, 2020, <https://www.geeksforgeeks.org/software-engineering-classification-of-software-requirements/> visited on 1/7/2020
- [14] ISO 31000, 2018. ISO 31000:2018. Risk management - Principles and Guidelines, Risk Management. <https://www.iso.org/standard/65694.html>.
- [15] ISO/IEC/IEEE29148:2018, <https://www.reqview.com/doc/iso-iec-ieee-29148-templates>, visited on 1/7/2020.
- [16] Klevanskiy, N.N., Tkachev, S.I., & Voloshchouk, L.A., 2019, Multi-project Scheduling: Multicriteria Time-cost Trade-off Problem, *Procedia Computer Science*, Volume 150, 2019, Pages 237-243.
- [17] Kornelius Irfandhi, 2016, Risk Management In Information Technology Project: An Empirical Study, *ComTech Vol. 7 No. 3 September 2016*: 191-199.
- [18] Layton, M. C., 2020, Retrieved from What's Different about Agile Risk Management? <https://www.dummies.com/careers/project-management/whats-different-agile-risk-management/> 2020, 7 11.
- [19] Li, Xiaoqing, Q., Hsu, M. K., & Chen, Q., 2019, Support or Risk? Software Project, Risk Assessment Model, Based on Rough Set Theory and Backpropagation Neural Network, *Sustainability* 2019, 11, 4513; doi:10.3390/su11174513, <https://www.mdpi.com/journal/sustainability>.
- [20] Marchwicka, Ewa., 2020, A technique for supporting decision process of global software project monitoring and rescheduling based on risk analysis, *Journal of Decision Systems*, DOI: 10.1080/12460125.2020.1790825.
- [21] Pinchangthong, D., & Boonjing, V., 2017. Effects of risk management practices on IT project success. *Manag. Prod. Eng. Rev.* 8, 30–37. <https://doi.org/10.1515/mpcr-2017-0004>.
- [22] PMI, Project Management Institute, 2017., *A Guide to the PROJECT MANAGEMENT BODY OF KNOWLEDGE*, Project Management Institute, Inc., Newtown Square, Pennsylvania 19073-3299 USA, p397.
- [23] Raz, T.; Shenhar, A. J., & Dvir, D., 2002, Risk management, project success, and technological uncertainty. *R&D Manag* 32:101–109.
- [24] Rong, W., & Ruixia, Y., 2017, "An algorithm for attribute reduction based on classification of condition attributes in rough set," 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, 2017, pp. 5534-5537, DOI: 10.1109/CCDC.2017.7979480.
- [25] Sarigiannidis, L., & Chatzoglou, P. D., 2011, Software Development Project Risk Management: A New Conceptual Framework. *Journal of Software Engineering and Applications*, 4(5), 293- 305.
- [26] Sharma, A., Sengupta, S., & Gupta, A., 2011, Exploring Risk Dimensions in the Indian Software Industry, *Project Management Journal*, Vol. 42, No. 5, 78–91.
- [27] Shaukat Z., Naseem, R., & Zubair, M., 2018, A Dataset for Software Requirements Risk Prediction, 2018 IEEE International Conference on Computational Science and Engineering, IEEE computer society pp112-118.
- [28] Sherer, S. A., 1995, "The three dimensions of software risk: technical, organizational, and environmental," *Proceedings of the Twenty-Eighth Annual Hawaii International Conference on System Sciences*, Wailea, HI, USA, 1995, pp. 369-378 vol.3, DOI: 10.1109/HICSS.1995.375618.
- [29] Shortreed, J., 2010, "Enterprise Risk Management and ISO 31000", *The Journal of Policy Engagement*, vol.2, no.3, 2010.
- [30] Tabunshchyk, G., Arras, P., & Merode, D. V., 2015, Risk Management in Multi-National Projects, 2015, The 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications 24-26 September 2015, Warsaw, Poland.
- [31] Talet, A. N., Mat-Zin, R., & Houari, M., 2014, Risk Management and Information Technology Project. *International Journal of Digital Information and Wireless Communications*, 4(1), 1- 9.
- [32] Taylor, H., Artman, E., and Woelfer, J., 2012, Information technology project risk management: Bridging the gap between research and practice *Journal of Information Technology* 27(1):17-34 • March 2012.
- [33] Vitalitychicago, 2020, <https://vitalitychicago.com/blog/agile-projects-are-more-successful-traditional-projects/> visited on 1/7/2020.
- [34] Wallace, L. & Keil, M. 2004, "Software Project Risk and their Effect on Outcomes", *Communication for the ACM*, vol 47 number 4, pp. 68-73, 2004.
- [35] Willumsen, P.; Oehmen, J.; Stingl, V.; Galdi, J., 2019 "Value creation through project risk management", *International Journal of Project Management*, 2019.
- [36] Zwikael O., & Mark A., 2011 *The Effectiveness of Risk Management- An Analysis of Project Risk Planning Across Industries and Countries*, Risk Analysis Volume 31 issue 1 2011.
- [37] Murta, Daniel & Oliveira, José Nuno, 2015, A study of risk-aware program transformation, *Science of Computer Programming*, Volume 110, Pages 51-77.
- [38] Alharbi, Ibraheem M; Alyoubi, Adel A; Altuwairiqi, Majid; and Abd Ellatif, Mahmoud, " Analysis of Risks Assessment in Multi Software Projects Development Environment using Classification Techniques", accepted for publication in *Advanced Machine Learning Technologies and Applications (AMLTA 2021)*, Series of Advances in Intelligent Systems and Computing, Springer, 2021.
- [39] Júlio Menezes Jr; Cristine Gusmão; Hermano Moura, 2013, Defining Indicators for Risk Assessment in Software Development Projects, *Clei Electronic Journal*, Volume 16, Number 1, April 2013.
- [40] Carr, Marvin., Konda, Suresh., Monarch, Ira., Walker, Clay., & Ulrich, F. (1993). *Taxonomy-Based Risk Identification (CMU/SEI-93-TR-006)*. Retrieved February 20, 2021, from the Software Engineering Institute, Carnegie Mellon University website: <http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=11847>.
- [41] Pandey, Dharendra; Suman, Ugrasen; and Ramani, AK, 2011, A Framework for Modelling Software Requirements, *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 3, No. 1, May 2011, www.IJCSI.org.
- [42] WNHW Mohamed, Mohd Najib Mohd Salleh, Abdul Halim Omar , A Comparative Study of Reduced Error Pruning Method in Decision Tree Algorithms, 2012 IEEE International Conference on Control System, Computing and Engineering, 23 - 25 Nov. 2012, Penang, Malaysia.
- [43] Anjali Shishodia, Vijaya Dixit, Priyanka Verma, (2018) "Project risk analysis based on project characteristics", *Benchmarking: An International Journal*, Vol. 25 Issue: 3, pp.893-918, <https://doi.org/10.1108/BIJ-06-2017-0151>.
- [44] IEEE Standard Glossary of Software Engineering Terminology," in *ANSI/ IEEE Std 729-1983* , vol., no., pp.1-40, 18 Feb. 1983, doi: 10.1109/IEEESTD.1983.7435207.
- [45] Desharnais, J.-M., Abran, A., & Suryn, W. (2010). Identification and analysis of attributes and base measures within ISO 9126. *Software Quality Journal*, 19(2), 447–460. doi:10.1007/s11219-010-9124-5.

Fog-based Remote in-Home Health Monitoring Framework

Fatma H. Elgendy¹

Electrical Engineering Dept
Kafrelshiekh Higher Institute for Engineering and
Technology, Kafrelshiekh, Egypt

Amany M. Sarhan², Mahmoud A. M. Alshewimy³

Computer and Control Engineering Dept
Faculty of Engineering, Tanta University
Tanta, Egypt

Abstract—As a result of what happened to the world during the past and current year of the spread of the Covid-19 epidemic, it was necessary to have a reliable health care system for remote observation, especially in care homes for the elderly. There are many research works have been done in this field, but still have limitations in terms of latency, security, response delay, and long execution times. To remove these limitations, this paper introduces a Smart Healthcare Framework called Remote in-Home Health Monitoring (RHHM), which provides architecture and functionalities in order to facilitate the control of patients' conditions when they are at home. The framework exploits the benefits of fog layers with high-level services such as local storage, local real-time data processing, and embedded data mining for taking responsibility for handling some burdens of the sensor network and the cloud and to become a decision maker. In addition to, it incorporates camera with body sensors in diagnosis for more reliability and efficiency with privacy preserving. The performance of the proposed framework was evaluated using the popular iFogSim toolkit. The results show the proposed system's ability to reduce latency, energy consumption, network communications, and overall response time. The efforts of this work will help support the overall goal to establish a high performance, secured and reliable smart Healthcare system.

Keywords—Fog computing; health monitoring; iFogSim; IoT; cloud computing

I. INTRODUCTION

With the increase in number of patients (like elderly people) living home alone under supervision of their doctors, it becomes necessary to add a facility for monitoring them remotely as the patient must be diagnosed for rapid intervention [1]. Moreover, some patients may have difficulties in reporting their pain condition for being impaired or disabled. Also, contagious diseases patients cannot communicate with others directly as Covid-19 patients. For such cases, smart healthcare systems represent urgent need. These systems used the technology to monitor patients at their homes or hospital and notify their doctors, caregivers or family individuals with their health status and demands [2]. These technologies involve IoT, cloud computing and fog computing.

IoT technology integrates hardware, computing devices, physical objects, software, and people via a network enabling them to interact, communicate, collect, and change data and actions are taken. Incorporating IoT technologies and healthcare improves the quality of healthcare systems.

Previous works have decomposed the architecture of IoT-healthcare system into three layers as depicted in Fig. 1 [3,4,5]: 1) body sensors that aggregate the data from the patients and the surrounding environment to be available for doctors, caregivers or authorized persons anytime and anywhere, 2) Fog gateways to receive and preprocess the sensor data, making some operations on data, and 3) cloud data center for storage, processing and analysis of data [4].

Cloud computing provides on-demand services to its users via internet in any place at any time. However, it is not adequate for meeting the requirements of rapid growth of IoT services, e.g., scalability and latency. It also has many negatives, such as: latency, network congestion, link disconnection and delayed response, in situations where the patient life may depend upon. So, fog computing can be the appropriate solution to collaborate with the cloud to fulfill the needs of real-time IoT services. Fog layer is an intermediate layer between sensors and cloud for data pre-processing, data analysis and storage rather than the cloud [6]. Fog gateway uses network switches, routers, smartphones, or embedded devices near the patient to provide cloud services with minimum latency and response time.

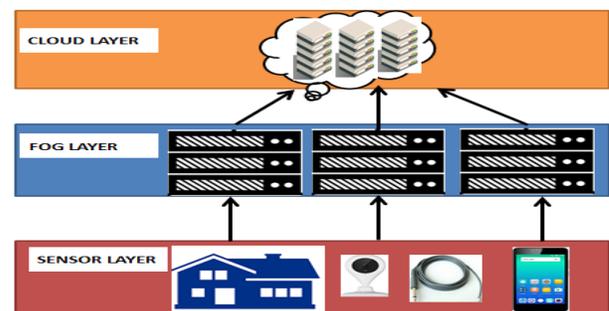


Fig. 1. General Architecture of SMART Healthcare Systems.

Healthcare monitoring systems in the literature comprise to two approaches: sensor-based and camera-based systems. Sensor-based approach uses the body sensor readings to diagnose patient status while camera-based depends upon capturing the patient image to analyze his emotions and body gestures. Each approach has its merits and its demerits according to the application and the objectives. This paper incorporates the two approaches in order to exploit the advantages of each in patient diagnosing whether the pain can be diagnosed by sensors only or not as the diagnosis of internal medicine, teeth pains, bones and muscles pains and

headache. The later health problems can be diagnosed efficiently using pain detection approaches rather than using sensors only.

In this paper, we introduce a smart fog-cloud based healthcare monitoring framework that uses both sensors and cameras for maintaining its various functionalities. The main contributions of our work are summarized as follows:

- Proposing a smart remote in-home healthcare monitoring system.
- Incorporating camera with body sensors in pain diagnosis for more reliability and efficiency, and for diagnosing non detected health problems such as internal medicine, teeth pains and so on.
- Taking advantages of the features of fog computing at which we add the decision maker that uses local data processing.
- Performance evaluation of the proposed fog-based system by using iFogsim toolkit.

This paper is framed in different sections. Section 2 introduces background and related work. The proposed framework is presented and explained in Section 3. Section 4 discusses the experimental results. Finally, Section 5 concludes the paper and suggests future works.

II. BACKGROUND AND RELATED WORK

The fog computing is a new paradigm in healthcare monitoring systems [6]. It can play a major role by its ability to perform data processing which preventing network congestion and reducing the amount of transmitted data. Most of the recent studies [7-17] in remote healthcare systems focus on using body sensors to monitor the patient status and provide the assistance to them. These sensors can be positioned on or embedded inside the body to measure certain signs of the patient. A specialized device is usually used to collect the data received from these sensors and is responsible for communicating them to the monitoring device. However, only few papers exploit using the camera to explain the emotional state of patients and analyzing it. According to that, healthcare related work will be discussed in the following two separated subsections where the first summarizes the usage of fog in healthcare and the second summarizes the usage of cameras in healthcare.

A. Fog Computing in Healthcare

Verma and Sood [7] introduced a fog layer for augmenting health monitoring system. They reduced the amount of data transmitted to the cloud by developing an event triggering mechanism which transfers the patient vital data in unsafe status only. Ahmad et al. [8] proposed a healthcare framework called HealthFog which concentrated on the data privacy in healthcare systems by integrating a cloud access security broker with the system in addition to cryptographic primitives. Nandyala and Kim [9] proposed an architectural view of a healthcare system known as IoT based u-healthcare monitoring system. This architecture exploits the advantages of fog computing which interacts more by serving closer to the edge at Smart Homes and Smart Hospitals.

An enhanced cloud-based Fog computing system is proposed by Gia et al. [10] where bio-signals are analyzed at the fog server side for real time applications. Negash et al. [11] concentrated on a smart e-health gateway implementation by connecting a network to such gateways, both in home and hospital use. Rahmani et al. [3] exploited the smart e-health gateways in healthcare IOT systems where the fog layer is between sensor network and cloud. A prototype of a smart e-health Gateway is presented for implementation and also an IoT-based Early warning score (EWS) health monitoring is implemented. Shree and Padmavathi [12] considered a fog-based scenario where the vital signs are collected from the patient, filtered, processed and take decisions at the fog layer.

B. Camera-based Systems in Healthcare

Mano et al. [4] developed a health smart home system based on the use of face image in patient identification and expression recognition. The authors in [13] observed the agitation in the sedated patients by installing the camera in the intensive care units. Banerjee et al. [14] proposed an approach for patient activity recognition in hospital beds using Kinect sensor. A non-invasive system for monitoring the newborns is presented by extracting and processing of motion signals, from multiple digital cameras, to estimate the periodicity of pathological movements using the Maximum Likelihood criteria [15]. Sun et al. [16] developed a video-based method for automated detection of the infants' discomfort. Chaichulee et al. [17] proposed an approach for detecting the presence of the patient and extracting cardio-respiratory signals from the patient skin with no clinical intervention using a video camera. A hybrid framework, in human behavior modeling in ambient assisted living systems, is proposed by Patel and Shah [18]. They added new events and behaviors and classified them into normal or abnormal human behavior by using camera images.

III. PROPOSED FRAMEWORK ARCHITECTURE

In this section, we describe the main architecture of the proposed framework, called Remote in-Home Health Monitoring Framework (RHHM). Fig. 2 illustrates the proposed framework architecture for remote patient monitoring, that is composed of three main layers: Sensor Network (IoT), fog and cloud layer.

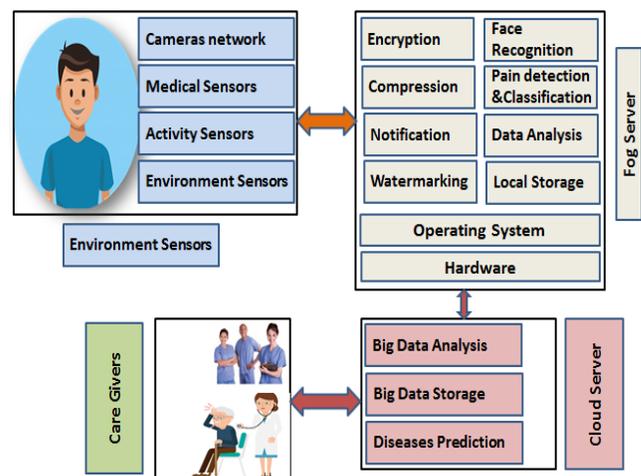


Fig. 2. RHHM Framework Architecture.

The following subsections demonstrate the details of the three main layers of the framework.

A. Sensor Network Layer

To collect the patient information that reflects his activity and medical state, several types of sensors are needed such as medical sensors, activity sensors, environment sensors and smart cameras constituting sensor network layer. Medical sensors include Electro Cardio Gram (ECG) sensor, Electro Encephalo Gram (EEG) sensor, Electro Myo Graphy (EMG) sensor, pulse oximeter sensor, temperature sensor, respiration rate sensor and glucose level sensor. Sensor's elements are responsible for detecting the presence of the patient, capturing his image, collecting readings and sending it to the fog devices via wireless or wired communication protocols. Sensors must be in all positions that the patient can move to it.

B. Fog Layer

Fog layer consists of multiple distributed smart devices or nodes, called gateways. A gateway is a device with computing, storage and network connectivity that is distributed near to the sensors which produce the data. Fog nodes handle four issues: 1) receiving patient data from sensors, 2) analyzing these data to make a decision of the health state of the patient, 3) communicating with caregivers, and finally 4) passing the data to the cloud. Local data processing is added to the fog for adding the intelligence to it in order to enhance the system by increasing its reliability, decreasing latency, overcoming the internet disconnection and speedup the decision especially at emergency situations. Fog layer has many functions and properties as followed:

1) *Data collection module:* This module can collect the physiological and environmental data from various wireless sensors embedded at different locations at home and from elderly patient and convert it to adequate form for analysis. Data filtering, preprocessing and noise removal are implemented in this module.

2) *Patient identification module:* This module is responsible for determining the identification of the patient via biometrics as face recognition, Iris recognition and finger print. However, face recognition is the easiest, fastest, cheapest and most reliable approach whereas it defines the patient identity and his emotions from videos or images, robust in different lightening conditions, able to work with faces from different angles and the camera and mobiles are available in anywhere. Face recognition consists of three steps: face detection, feature extraction and face recognition [19]. Face detection step decides if the captured image contains face or not. If it contains a face, it detects and specifies the face location in the image. Feature extraction includes extracting the features from the detected face by one of the two categories: appearance based or geometric based techniques [20]. Finally, recognition step compares the feature vector with all the stored vectors in the face database to decide the face identity by using machine learning classifiers. Fig. 3 shows the general steps of the face recognition process.

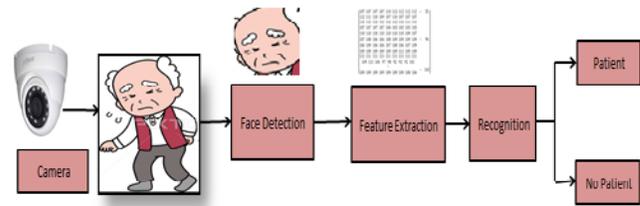


Fig. 3. Face Recognition Process.

3) *Patient pain recognition module:* this module is incorporated in the monitoring system to detect and measure the pain of the monitored patient. The pain can be detected through two methods: the first uses the vital signs as temperature, blood pressure and blood glucose which are collected from the wearable sensors that reflect the vital measures of the body. The second method uses the facial expressions to determine if the patient has pain or not and the degree of pain. Facial expression recognition is beneficial at the healthcare frameworks because it reflects the patient health status and emotions [4, 21]. It also helps the system to define the pain from diseases which do not have a device that reflects it directly such as: headache, teeth pain, and internal medicine pains. Taking a picture of a patient face periodically can give an idea of to what extent the patient is feeling pain and can enable a caregiver to decide whether help is required or not. Also, patient emotions are important factor during diseases treatment and diagnose of various diseases such as: schizophrenia, depression, autism and bipolar disorder. This module classifies the pain degree according to both facial expression and vital signs. The pain classification according to facial expression is depicted in Fig. 4.

4) *Data security module:* As the patient data can be vulnerable through being transmitted from a gateway to the cloud, the security and integrity of these data are essential requirements that must be considered in designing healthcare frameworks. Therefore, encryption and watermarking techniques as important components in secured healthcare frameworks are introduced by this module. Encryption is the process of converting data to an unrecognized form to protect it from unauthorized people. Furthermore, watermarking is the procedure of hiding information within an image without distorting the visibility or credibility of it.

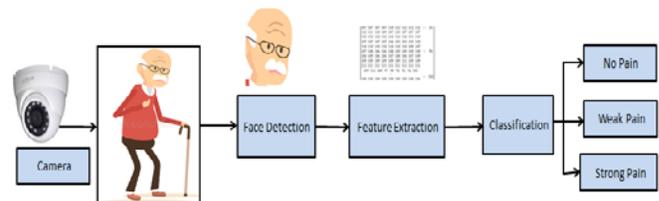


Fig. 4. Pain Detection and Classification from Face Image.

The data security module in our proposed framework aims to hide the patient data when stored in the fog local server or transmitted to the cloud server by first watermarking the sensors data of the patient into the patient face image, then using a proper encryption algorithm the watermarked patient face image is converted into a cipher image that cannot be read except by the authorized parties as depicted in Fig. 5.

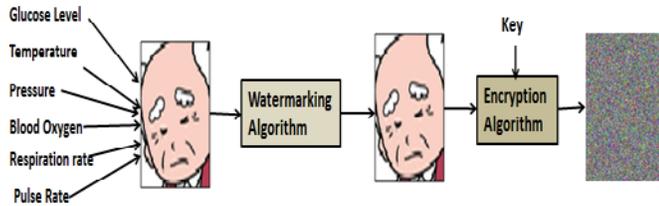


Fig. 5. Watermarking followed by Image Encryption.

5) *Local storage module*: The gateway should have a private local storage which stores the incoming data on a local repository to perform functions as data security and data analysis. Moreover, the data that could not be transmitted to the cloud due to network problems is stored temporally in the gateway.

6) *Notification module*: This module is responsible for sending alarms to caregivers and family persons if an urgent situation has occurred to the patient. It also tells the patient if an action is required; time of medicine, quantity of medicine, a specific food must eat, and any other desired instructions to follow. The notification may be a message, voice telephone call or voice over IP (VoIP). NG112 architecture is an example of VoIP that enables the modernization of emergency communications, allowing for far more data collection (text, video, location or additional data), which will result in a more efficient response [22]. NG112 also helps to ensure equivalent access for all citizens, including people with disabilities.

7) *Data analysis module*: This module is responsible for making decisions at fog layer in situations that need the fast intervention. The decision must be carried out by gathering information about the patient such as: age, weight, height and current diseases. Moreover, it receives the sensor data and applies machine-learning algorithms to decide if an emergency situation has occurred to send a fast alarm and send the collected data to the cloud. If the data does not contradict the patients expected behavior or health, the decision maker module does not take any action and no data will be transmitted to the cloud storage.

A generalized algorithm is proposed for decision-making and notification based on camera face image and vital signs such as body temperature, blood pressure and blood glucose percentage. The proposed algorithm is illustrated in Algorithm 1. The algorithm starts with determining the threshold of each reading such as *pain_score*, temperature threshold, blood pressure, blood glucose percentage. Sensor data is sent to the machine-learning algorithm. This data is then analyzed to determine the action to be taken as shown in the selection structure (IF statement) of Algorithm 1.

Algorithm1: Decision-making and notification algorithm

Input: Real values from Camera and sensors

Output: Notification at emergency situations

Steps: -

- Initialize the threshold of each sensor reading: pain class from face image as *pain_score*, temperature threshold as *T_threshold*, blood pressure threshold as *B_threshold*, blood glucose percentage as *G_threshold*.
- Read the sensors data and camera data.
- Determine *pain_score* using face image and machine learning model
- If (*pain_score*==strong)
 NG112 call to **CareGiver** via mobile app.
- Else If (*blood pressure*<*B_threshold* || *blood pressure*<>*B_threshold*)
 NG112 call to **CareGiver** via mobile app.
- Else If (*temperature threshold* < *T_threshold* || *temperature threshold* > *T_threshold*)
 NG112 call to **CareGiver** via mobile app.
- Else If (*blood glucose percentage* < *G_threshold* || *blood glucose percentage* > *G_threshold*)
 NG112 call to **CareGiver** via mobile app.
- Else
 NG112 call to **patient** via mobile app "Normal situation".
- End if

8) *Data compression module*: The patient data must be sent to the cloud for storage, diseases prediction, risk amount assessment and long-term analysis. Huge amount of data is sent periodically to the cloud which cause network congestion, latency and it needs a large space for storage. Image compression deals with the previous problems. There are two types of image compression algorithms; lossless and lossy methods [23]. Medical image as in our work must use the lossless methods because any loss of the information in the original image can result in improper diagnosis and risk assessment. Lossless compression is decomposed of two steps. The first step: transforms the image to another format. The second step uses an entropy encoder to remove the coding redundancy [23].

C. Cloud Layer

Cloud layer consists of distributed resources, repositories and servers that are located at far place. These devices are managed by a cloud manager that integrates them to receive process and store the patient data. The doctor/caregiver can use these data to fulfill long-term analysis evaluating the status and history of patient health. It is responsible for the following functions:

1) *Big data storage*: The data gathered from sensors, after being analyzed at the fog layer, is transmitted to the cloud storage which provides enormous amount of storage for saving large healthcare data files. The healthcare data is available to doctors, caregivers, hospitals and assurance companies for further various usages.

2) *Big data analysis*: It helps in analyzing all patient data (such as images, symptoms, therapy plans and treatment) and makes all the data available to support and improve long term clinical decision-making processes, treatment and medical research. Machine learning algorithms, predictive algorithms and data visualization algorithms can be conducted on these data to get more insights form these data.

3) *Diseases Prediction*: Based on patient information such as: age, height, weight and historical diseases in his family, this unit is responsible for the expectation of the diseases which the patient can possibly incur in the future. It uses machine-learning algorithms to compute the percentage of the expected diseases based on the correlation of the accepted vital signs as simply shown in Algorithm 2.

Algorithm. 2: Diseases prediction phase

Input: vital signs as: temperature, pressure, blood glucose.
Output: percentage of diseases prediction.

Algorithm:

- Accept each vital signs of a patient.
- Compute the correlation between vital sign and predicted diseases.
- Use machine learning algorithm to specify the percentage of each disease possibly.

IV. DATA FLOW OF THE PROPOSED FRAMEWORK

Fig. 6 depicts the components and dataflow of the proposed framework. Each component and its job are described as follows:

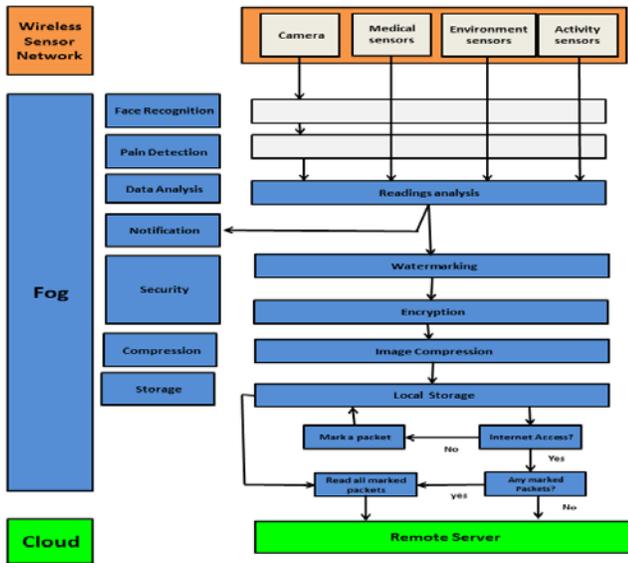


Fig. 6. RHHM Framework Dataflow.

1) The fog device receives the patient face image from the camera or mobile phone. Then, the fog device either continues its work or does nothing according to the decision of recognition phase.

2) If a pain is detected from the face expressions or unusual body activity, it alerts the data analysis unit to integrate with sensor readings.

3) Data analysis is accomplished on sensor readings. Data filtering is applied to the received sensor data to remove the noise, electromagnetic interference and improper attachments of sensors. Then, it analyzes the filtered sensors data to assist the system in diagnose the patient situation.

4) The system notifies responsible persons about the patient health status if an emergency situation is occurred.

5) The patient data is forwarded to be stored at the cloud for future needs. This data is hidden in the patient image. Then, image is encrypted and compressed.

6) The system saves a backup of patient's health data locally in case of a connection problem with the cloud.

V. DISCUSSION

Improvements in using the fog-based framework against cloud based only for remote pain monitoring will be illustrate in the experimental results. To validate the effectiveness of the proposed fog-based framework, the results are compared to the cloud-based framework for health monitoring. The comparison is accomplished in terms of energy consumption, latency, network usage and response time. The results of the simulations are performed on different scales confirmed the effectiveness of the proposed framework compared to the cloud based one. Also, we briefly compare the proposed framework to the existing healthcare frameworks as shown in Table II. To the best of our knowledge, none of the existing systems uses both the face emotions and pain detection in patient identification together.

From Table I, it is obvious that the proposed framework exceeds its predecessors in different aspects.

TABLE I. COMPARISON OF THE PROPOSED FOG FRAMEWORK WITH THE EXISTING FRAMEWORKS

Referenc e	Paradigm	Use of IoT sensors	Use of image processing	Patient identificati on	Face as identificati on
[7]	Fog	√	×	√	×
[8]	Fog	√	×	√	×
[9]	Fog	√	×	×	×
[26]	Cloud	√	×	×	×
[27]	Cloud	√	×	×	×
[28]	Cloud	√	×	×	×
[29]	Cloud	√	×	×	×
[30]	Fog	√	×	×	×
[4]	Fog/cloud	×	√	√	×
[3]	Fog	√	×	×	×
[12]	Fog	√	×	×	×
Proposed framewo rk	Fog	√	√	√	√

VI. EXPERIMENTAL RESULTS

Considering suitability as tool for simulating applications, iFogSim [24] was employed to evaluate the proposed RHHM framework by using the basic functionalities such as datacenters and cloudlets. As first simulator of IoT objects, it connects to the fog devices and cloud in a hierarchical manner [25].

A. Model Building

The simulation is accomplished with four types of fog devices. Four physical topology configurations of a monitored home are conducted to obtain accurate results: config1, config2, config3 and config4 which represent homes composed of one room, 2 rooms, 3 rooms and 4 rooms to be monitored, respectively as in Fig. 7. Each room is equipped with 2 smart cameras monitoring the areas, temperature sensor and pressure sensor, in addition to ISP gateway. The whole home has a WiFi gateway which transmits data to the cloud. Smart cameras transmit live video streams to the mobile phone which recognizes the patient and in order to send to ISP Gateway. Smart phone has 1GB RAM, 1GHz CPU and needs 87.53 W of power while ISP Gateway has 4 GB RAM, 2.8 GHz CPU and needs 107.330W of power as the WiFi Gateway but the cloud has 8 GB RAM, 4 GHz CPU and needs 1648W of power.

Table II shows the latency from each fog device to another in a hierarchy manner. Through the conducted experiments, two scenarios were evaluated as depicted in Fig. 8: the first one is cloud-based RHHM where the sensors including smart cameras send their readings to the cloud for decision making while the second scenario is fog based RHHM where the data is analyzed at the fog gateways to make the decision.

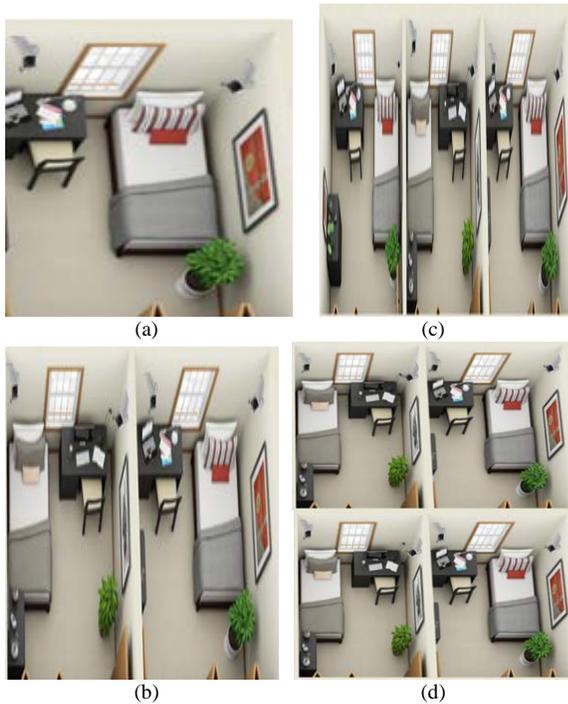


Fig. 7. Physical Topology Configurations: (a) Config1, (b) Config2, (c) Config3 and (d) Config4.

TABLE II. LATENCY BETWEEN FOG DEVICES

Source	Destination	Latency (ms)
Camera	Smart Phone	1
Smart Phone	ISP Gateway	2
Temperature or Pressure sensor	ISP Gateway	5
ISP gateway	WIFI Gateway	2
WIFI Gateway	Cloud	100

These models are evaluated in terms of work delay, energy consumption, network usage and execution time and the simulation charts are given. In the experiments, each scenario is built and evaluated individually using the following steps:

1) The application modules are constructed, which are represented by 11 squares in Fig. 8(a), and 9 squares in Fig. 8(b), as: face recognition, pain detection, watermarking, and so on with the following features: RAM=100MB, MIPS(million instruction per second)=1000, Bandwidth=1000 and VM (virtual machine) is Xen.

2) The application edges connect the application modules with each other; where each application edge is described by its source, destination, tuple CPU length, tuple network length and direction.

3) The application loops, which represent the flow of data, are conducted according to each model. Two application loops were used to represent the delay; the first loop starts from capturing images to caregiver notification while the second loop starts from data analysis model to patient knowledge with the decision.

4) Fog devices are considered.

5) Adding application modules to the fog devices as described in Fig. 8.

6) Finally, the experiment is conducted.

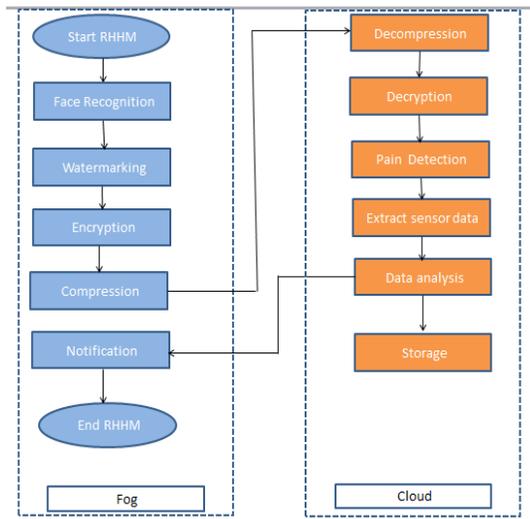
B. Energy Consumption

To prove our framework, we have measured the energy consumed by the different categories of devices (mobile phones, edge devices and cloud) as illustrated in Fig. 9. It is observed that the energy consumed at the mobile phones and edge devices are semi-equal at the two scenarios because the devices work periodically in the two situations. But the energy consumption at the cloud in cloud-based RHHM is higher by 8%-20% than it in fog-based RHHM, because the data analysis and decision making is implemented at the fog in fog based RHHM while they are shifted to the cloud in cloud-based RHHM. The total energy consumed in cloud-based RHHM is 9%-22% higher than that consumed in the fog-based RHHM which means that the proposed fog-based RHHM saves the energy of the system components and meet the performance requirements.

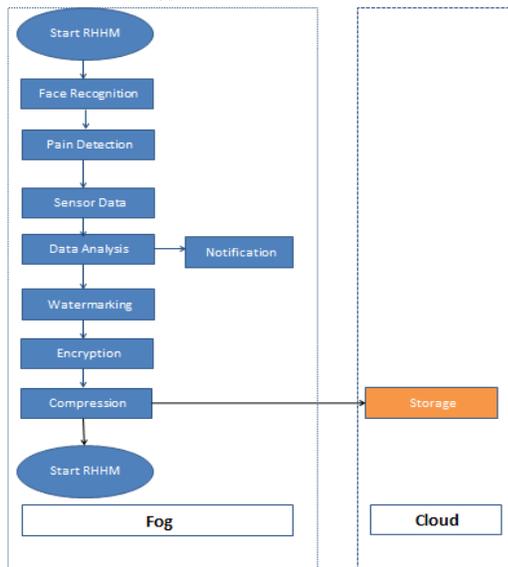
C. End to End Latency

The time from sending the data from camera, or sensors, until the result or the decision appears on the patient smart phone is recognized as the end-to-end latency. It computes the time needed to take the decision according to the patient status. It is noted from Fig. 10 that the processing time on fog devices in fog/cloud-based RHHM is very small compared to the processing on the cloud in cloud-based RHHM. The

latency also increases as data transmission increases resulting from the number of rooms under consideration. The average latency is reduced at fog-based RHHM by 97% than that at the cloud-based RHHM.



(a) Cloud based RHHM.



(b) Fog based RHHM.

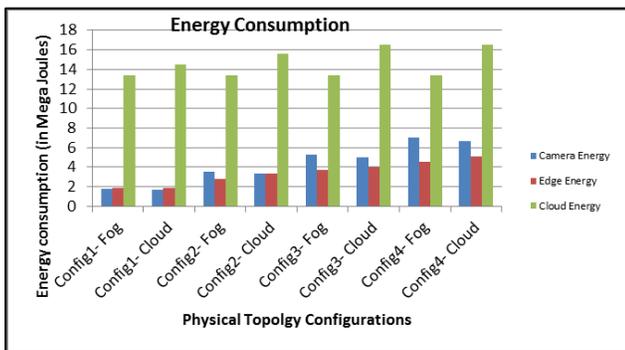


Fig. 9. Energy Consumption of Mobile Phones (Cameras), Edge Devices and Cloud.

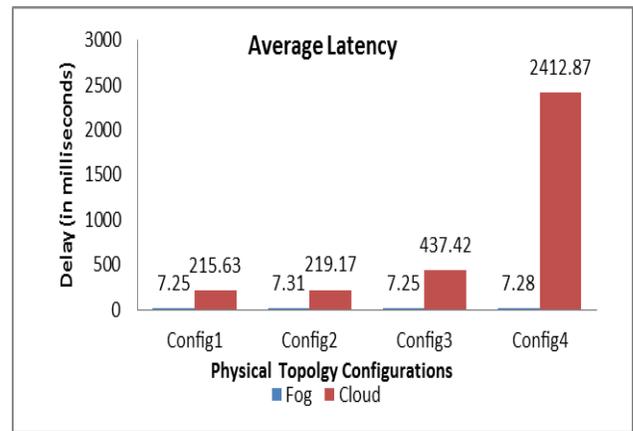


Fig. 10. Average Latency of Control Loops.

D. Network Usage

Fig. 11 shows the network usage at the cloud and the fog. It is obvious that the network time usage and the data transmitted via the network is 15% to 16% less using the fog/cloud based RHHM.

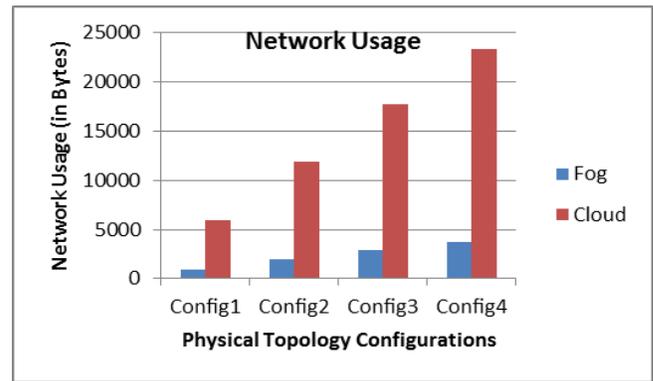


Fig. 11. Network usage.

E. Total Response Time

Fig. 12 shows that the total response time increases when the number of rooms increases and consequently data transmission increases due to the amount of collected data increases and the need for processing increases. The figure also shows the high contrast between the processing at fog layer in fog/cloud based RHHM and cloud in cloud based RHHM.

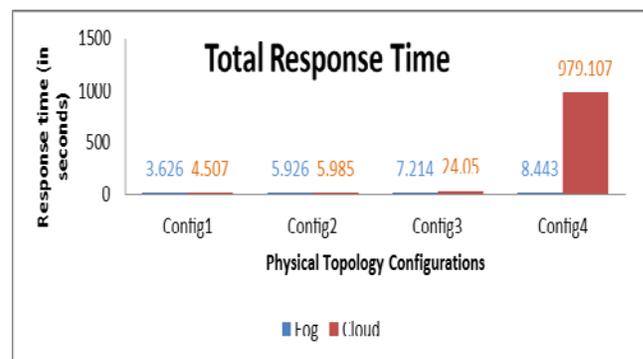


Fig. 12. Total Response Time of the Frameworks.

VII. CONCLUSION AND FUTURE WORK

In this work, we propose a fog/cloud based RHHM framework that moves the decision making to the fog gateways and incorporates a camera with wearable sensors in detecting some pains that cannot be diagnosed or detected through sensors only such as: internal medicine, teeth pain and headache. The framework archives the security of the patient data using watermarking and encryption algorithms. Our conducted experiments illustrate the RHHM system with various merits as enhanced overall system intelligence, energy efficiency, less network usage, and high response time.

So that, RHHM is a good choice to be used at homes, elderly care homes, or hospitals in the near future. The framework can be improved by using alarming systems and incorporating other image analysis techniques as fall detection to improve patient status assessment.

REFERENCES

- [1] Mshali H, Lemlouma T, Moloney M, Magoni D. A Survey on health monitoring systems for health smart homes. *International Journal of Industrial Ergonomics*. 2018.
- [2] Farahani B, Firouzi F, Chang V, Badaroglu M, Constant N, Mankodiya K. Towards fog-driven IoT eHealth: promises and challenges of IoT in medicine and healthcare. *Future Generation Computer Systems*. 2018.
- [3] Rahmani AM, Gia TN, Negash B, Anzanpour A, Azimi I, Jiang M, Liljberg P. Exploiting smart e-Health gateways at the edge of healthcare Internet of Things: a fog computing approach. *Future Generation Computer Systems*. vol. 78, pp. 641-658. 2018.
- [4] Mano LY, Faical BS, Nakamora LHV, Gomes BH, Libralon GL, Meneguete RI, Felho GPR, Giancristofaro GT, Pessin G, Krishnamachari B, Ueyama J. Exploiting IoT technologies for enhancing health smart homes through patient identification and emotion recognition. *Computer Communications*. 2016.
- [5] Soni LN, Datar A, Datar S. Viola-Jones algorithm based approach for face detection of african origin people and newborn infants. *International Journal of Computer Trends and Technology (IJCTT)*. 2017. Mutlag A, Abd Ghani M, Arunkumar N, Mohamed M, Mohd O. Enabling technologies for fog computing in healthcare IoT systems. *Future Generation Computer Systems*. 2018.
- [6] Verma P, Sood SK. Fog assisted IoT enabled patient health monitoring in smart homes. *IEEE Internet of Thing Journal*. 2019.
- [7] Ahmad M, Bilal M, Hussain S, Ho B, Cheong T, Lee S. Health Fog: a novel framework for health and wellness applications. *Journal of Supercomputing*. vol.72, no. 10, pp. 3677-3695. 2016.
- [8] Nandyala CS, Kim H, "From cloud to fog and IoT-based realtime u-Healthcare monitoring for smart homes and hospitals. *International Journal of Smart Home*. vol. 10, no. 2, pp. 187-196. 2016.
- [9] Gia TN, Rahmani MJA, Westerlund T, Liljberg P, Tenhunen H. Fog computing in healthcare Internet-of-Things: A case study on ECG feature extraction. *IEEE International Conference on Computer and Information Technology*. pp. 1-8. 2015.
- [10] Negash B, Anzanpour A, Azimi I, Jiang M, Westerlund T, Rahmani AM, Liljberg P, Tenhunen H. Leveraging fog computing for healthcare IoT. in *Fog computing in the Internet of Things Intelligence at the edge*, Springer. pp. 145-169. 2017.
- [11] Shree GJ, Padmavathi S. A Fog-based approach for real-time analytics of IoT-enabled healthcare. *Internet of Things Use Cases for the Healthcare Industry Book*, Springer. 2020.
- [12] Chase JG, Agogue F, Starfinger S, Lam Z, Shaw GM, Rudge AD, Sirisena H. Quantifying agitation in sedated ICU patients using digital imaging. *Computer Methods and Programs in Biomedicine*. 2004.
- [13] Banerjee T, Enayati M, Keller JM, Skubic M, Popescu M, Rantz M. Monitoring patients in hospital beds using an unobtrusive depth sensors. *International conference of the IEEE Engineering in Medicine and biology Society*. 2014.
- [14] Cattani L, Alinovi D, Ferrari G, Raheli R, Pavilidis E, Spagnoli C, Pisani F. Monitoring infants by automatic video processing: a unified approach to motion analysis. *Computers in Biology and Medicine*. 2017.
- [15] Sun Y, Shan C, Tan T, Long X, Pourtaherian A, Zinger S, de with PHN. Video-based discomfort detection for infants. *Machine Vision and Applications*. 2018.
- [16] Chaichulee S, Villarroel M, Jorge J, Arteta C, McCormick K, Zisserman A, Tarassenko L. Cardio-Respiratory signal extraction from video camera data for continuous non-contact vital sign monitoring using deep learning. *Institute of Physics and Engineering in Medicine (IPEM)*. 2019.
- [17] Patel A, Shah J. Real time human behavior monitoring using hybrid ambient assisted living framework. *Journal of Reliable Intelligent Environments*. 2020.
- [18] Kortli Y, Jeridi M, AlFalou A, Atri M. Face recognition systems: a survey. *Sensors*, MDPI. 2020.
- [19] Shier WA, Yanushkevich S. Pain recognition and intensity classification using facial expression. *The International Joint Conference on Neural networks (IJCNN)*. 2016.
- [20] Elgendy F, Alshewimy M, Sarhan A. Pain detection/classification framework including face recognition based on the analysis of facial expressions for E-Health systems. *International Arab Journal of Information Technology*. 2021.
- [21] Barakat R, Catal F, Tcholtchev N, Rebahi Y. TTCN-3 based NG112 test system and playground for emergency communication. *IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*. 2020.
- [22] Gupta S, Nagar D. Image compression: a review. *International Journal of Advanced Research in Science, Communication and Technology(IJARSCT)*. 2020.
- [23] Gupta H, Dastjerdi A, Ghosh S, Buyya R. iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things. *Edge and Fog computing environments. Software: Practice and Experience*. 2017.
- [24] Mahmoud M, Rodrigues J, Saleem K, Al-Muhtadi J, Kumar N, Korotaev V. Towards energy-aware fog-enabled cloud of things for healthcare. *Computers and Electrical Engineering*. 2018.
- [25] Yang G, Jiang M, Ouyang W, Ji G, Xie H, Rahmani A, Liljberg P, Tenhunen H. IOT based remote pain monitoring system: from device to cloud platform. *IEEE Journal of Biomedical and Health Informatics*. 2018.
- [26] GJ BK. Internet of Things (IOT) and cloud computing based persistent vegetative state patient monitoring system: a remote assessment and management. *International Conference on Computational techniques, Electronics and Mechanical Systems (CTEMS)*, Belagavi, India. 2018.
- [27] Alkhafajiy M, Baker T, Chalmers C, Asim M, Kolivand H, Fahim M, Waraich A. Remote health monitoring of elderly through wearable sensors. *Multimedia Tools and Applications*. 2018.
- [28] Tejaswini S, Sriraam N, Paradeep G. Cloud-based framework for pain scale assesment in NICU- a primitive study with infant cries. *IEEE Third International Conference on Circuits, Control, Communication and Control*. 2018.
- [29] Hassan S, Ahmad I, Ahmad S, Alfaify A, Shafiq M. Remote pain monitoring using fog computing for E-healthcare: an efficient architecture. *Sensors*, MDPI. 2020.

What Drives Airbnb Customers' Satisfaction in Amsterdam? A Sentiment Analysis

Heyam Abdullah Bin Madhi¹, Muna M. Alhammad²

Management Information System Dept
King Saud University, Riyadh
Saudi Arabia

Abstract—The sharing economy is a new socio-economic system that allows individuals to rent out their personal belongings, such as their private car or a room in their home, for a short period. This study aims to investigate the attributes that impact customers' satisfaction when using the sharing economy propriety rentals websites. Large data sets of Airbnb's online reviews and listings in Amsterdam were analyzed using sentiment analysis, word clustering, ordinal logistic regression, and visualization techniques. Findings reveal that the polarity of Airbnb guests reviews in Amsterdam is significantly impacted by property price, value, cleanliness, rate, host communication, easiness of check-in, the accuracy of property description, and whether the host is super host or not. Surprisingly, the property neighborhood was not found to impact customers' sentiment in Amsterdam. In addition, Airbnb guests in Amsterdam tend to positively express their experience satisfaction level mainly based on property exact location and host interaction followed with the facilities surrounding the property, property cleanliness, and room quality. On the other hand, negative online reviews tend to be mainly linked to problems with check-in services followed by aspects related to weak host interaction, location, and room quality. The results indicate that Airbnb hosts need to offer clear and easy check-in services with emphasizing the importance of keeping a good communication channel with their guests to enhance customers' experience and increase customers' satisfaction level. Future studies should investigate the applicability of the findings of this study in the context of other cities.

Keywords—Airbnb; customer satisfaction; customer experience; big data; sentiment analysis; ordinal logistic regression

I. INTRODUCTION

The development of media technology has contributed to the prosperity of a new business concept, which is the sharing economy [1]. A sharing economy is an advanced approach for conducting businesses, where people can share their own resources in a decentralized or peer-to-peer (P2P) platform for an exchange of money [2]. It has generated new business opportunities, especially for those who don't have the essential resources. Additionally, property owners have granted the chance to liquidate and utilize their unused assets across the globe, allowing them to serve customers from different regions and cultures [3]. The sharing economy had provided the necessary services in different qualities and prices 24/7 in order to provide suitable services to fulfill individual needs and cater to a different level of purchasing power. As a result, consumption and production sustainability have been accelerated [4]. The concept of the sharing economy has been

applied in different fields including, shared mobility, hospitality, on-demand staffing, and media streaming. Lime, JustPark, and Uber are some examples of sharing economy applications. In the context of hospitality, several companies have emerged, such as Airbnb, Couchsurfing, HomeAway, Roomorama, and HomeExchange. These organizations act as an intermediary between guests and hosts to facilitate communications between them [5].

Looking at Airbnb as an alternative option for hotels, several existing research have explored this platform from different perspectives. The author in [6] examined the substitution and complementary effects of Airbnb supply on hotel sales performance patterns while [7] studied the extent to which Airbnb is used as a substitute for hotels and [8] studied the price difference between hotel property and the nearby Airbnb listings offers. Additionally, many other researchers have studied the social contact of guests in the Airbnb accommodation. For instance, [9] explored the social contact of Airbnb guests during their stay, taking into account these types of contact: guest to host, guest to the community, and guest to guest communications whereas [10] studied the reciprocal aspect of social interactions in P2P accommodations. Several other studies have examined factors that might affect guests' booking intentions such as studying the impact of gender congruity between guests and hosts on guests' booking [11], and the impact of properties and hosts descriptions on guests' booking and posting reviews [12]. Moreover, some authors recognized the importance of studying online advertising strategies [13], price determination [14], trust-attachment building mechanisms [15], and customers' psychological behavior [16]. Despite the diversity of Airbnb research fields, few studies have examined guests' reviews toward their Airbnb stay. [17] have investigated the attributes that influence Airbnb users' experiences through the use of text mining and sentiment analysis. Similarly, [18] have investigated hidden dimensions in textual reviews through the use of Latent Aspect Rating Analysis (LARA). Analyzing big data such as customer reviews is a key success factor for sharing economy businesses [19]. It enables a business owner to understand the organization's strengths and weaknesses along with customers' habits. Therefore, it helps to predict future trends, improve decision making, detect frauds, and eventually increase business revenues.

However, to the best of the authors' knowledge, there are no previous studies investigating factors that drive customers' satisfaction through analyzing guests' online reviews

simultaneously with the listing features. Therefore, to better understand customers' needs and achieve customers' satisfaction, there is a need to identify factors that makes customers satisfied. Analyzing users' comments and reviews, through the use of various data mining techniques, helps to identify factors impacting users' satisfaction [19]. Therefore, the research question to be answered in this research is:

What makes customers satisfied when using the Airbnb platform in Amsterdam?

Airbnb online reviews dataset in Amsterdam will be analyzed simultaneously with the listings details dataset by using sentiment analysis, topic clustering, regression, and visualization techniques. The paper begins with reviewing relevant literature, clarifying the applied methodology, presenting the results and discussion, and finally concludes the paper.

II. LITERATURE REVIEW

A. Sharing Economy

According to [20], sharing economy concept is the result of a massive shift, not a new phenomenon in itself. This is because sharing economy is an online community that includes economic aspects such as selling, buying, and renting which already existed before. However, the new concept is distinguished by allowing users to participate using different types of resources, in both economic and social aspects, whether it was a human, merchandise, service, or property [21]. A more comprehensive description was provided by [22], where they defined sharing economy as a P2P environment that enables a person to be a client and a provider at the same time. The sharing economy facilitates social participation and cooperative consumption which in turn supports protecting the environment, minimizing resource waste, and enhancing community awareness [4]. The sharing economy covers five main areas: product-service systems, redistribution markets, collaborative lifestyles, access-based consumption, and commercial sharing systems.

As mentioned earlier, sharing economy businesses have witnessed rapid development in the last decade. The author in [5] justify the reason to be related to the fact that sharing economy businesses are less expensive for consumers, especially in the case of accommodation as it is cheaper for travelers to choose a place from Airbnb to stay, rather than renting a room in a hotel. Additionally, [19] have demonstrated that science and technology revolutions are playing a huge role in sharing economy growth through the development of electronic social networking (e-SN) platforms which contributed to the development of society and the economy.

Although business fields differ, online hospitality is one of the most famous fields in sharing economy. The word hospitality was defined by [22] as two associated meanings. From the guests' perspective, hospitality is a process of delivering a service of high quality while from hosts' perspectives; it is about providing services and rooms with a focus on profitability. In terms of hospitality, using the sharing economy, hospitality means allowing individuals to invest their properties to host visitors for a small financial return through

the use of online sharing economy portals such as the Airbnb website.

B. Customers' Satisfaction

The word satisfaction is defined as "a psychological statement with the fulfillment of a need or desire and the pleasure obtained by such fulfillment" [19]. Customers' satisfaction can be interpreted as customers' judgments, opinions, feelings, impressions, or emotional reactions towards their comprehensive experiences of a product or a service [23]. [22] have indicated that it is essential to understand customers' satisfaction and how it is impacted to ensure customers' commitments and obtaining their loyalties. Increasing customer satisfaction will increase customer retention and loyalty. Consequently, organizations' revenues and profits will rise as there is a direct link between customers' satisfaction and profit [24]. However, increasing customers' satisfaction comes at a cost, so organizations must be careful not to reach excessive satisfaction that costs them a lot without returns. To achieve the maximum customer satisfaction level with minimum cost, organizations must understand their customers and what makes them satisfied [24].

Obtaining customers' satisfaction could be a big challenge for many organizations. The first step to increase customers' satisfaction is identifying customers' current satisfaction levels and the reasons that led to this satisfaction level. Capturing and measuring customer experience is an excellent way to achieve that. Analyzing rating scores and comments that were produced by customers after the end of a service, can grant comprehensive knowledge about customers' satisfaction [25]. In this research, Airbnb guests' reviews will be analyzed, to better understand customers' satisfaction along with factors that affect their level of satisfaction.

C. Airbnb

Airbnb, Inc. is an online marketplace that was founded in 2008 for offering accommodates, homestays, or tourism experiences. It is one of the most leading and profitable platforms within the sharing economy [26]. The organization works as a broker that connects real estate owners with hospitality seekers. In particular, it links individuals who need to lease out their homes (hosts) with individuals who are searching for lodging (guests) in that region in exchange for a small commission from each reservation [27]. Hosts are taking part in Airbnb as an approach to obtain some pay from their property even though they are facing the risk that guests may harm their property. On the other hand, guests are taking advantage of having accommodation with relatively lower prices than other places while at the same time facing the risk that the property's quality will not be as expected. Generally, Airbnb provides services for both guests and hosts to provide a better coordinating outcome [28].

Several types of research have been applied to Airbnb, some focusing on user experience and multimedia, others on administration, tourism, and architecture [29]. Covering multiple areas including review bias, hospitality exchanges, price and neighborhood prediction, neighborhoods ranking, socio-economic characteristics, listing recommendation system, rentals' distribution, users' preferences and expectations, image mining, demand mining, grading schema,

matching schema, consumer segmentation, trust evaluation, innovation adoption, and adoption evaluation. In addition to some researches that focus on analyzing customers' satisfaction dimensions [30], analyzing customers feedback using text mining [3], and analyzing public opinions using content analysis [21]. Previous studies have almost exclusively focused on analyzing user reviews separately without linking them to listings data. A more comprehensive analysis is required for analyzing factors affecting customers' satisfaction, in a way that enables non-technical users to understand it without any effort.

III. METHODOLOGY

A. Data Collection and Preparation

Over time, extensive literature has confirmed the importance and usefulness of analyzing online review comments for researchers, business owners, as well as other customers. Since users' comments illustrate their perceptions and feelings formed while using a service or a product, it could be used to measure and analyze users' satisfaction level, along with identifying key factors that led to such a level of satisfaction [31].

On the Airbnb website, guests are posting reviews that reflect their experiences and opinions about the property they stay on. This provides data that can be analyzed, to identify factors that make customers satisfied [32]. The current study obtained online review comments of Airbnb' guests about their accommodations in Amsterdam, as well as detailed listing data from the Inside Airbnb website [33]. Given the fact of the massive availability of Airbnb data as public Datasets available on the Inside Airbnb website, both datasets were obtained easily. The listing dataset contains 107 attributes that describe host, property features, and scored reviews while the review dataset contains six attributes and more than 400 thousand review records. Datasets were then pre-processed or cleaned before the analysis phase, as performing this step usually helps to generate feasible results [34]. For the listings dataset, the cleaning process was done by eliminating attributes containing errors, no values (blank), and useless values. Thereafter, new columns were generated or derived from existing attributes, such as host experience in months derived from the host starting date, weekly and monthly discounts derived from prices. On the other hand, the review dataset was checked for missing data and non-English content to be removed. Google Online Spreadsheet was used to detect comments' language, and to clean non-English data from the set, which represents 18% of the records. Unwanted columns were eliminated including comment ID, reviewer ID, and reviewer name. Additionally, a filter was applied using the date of the comment, choosing comments within 2019 only, to minimize the number of reviews and make it manageable to analyze the data using home devices with limited capabilities. Dummy comments containing characters or numbers only were also removed leaving a total of 110,747 comments eligible for the analysis.

B. Data Analysis

1) *Sentiments analysis*: The first step in the analysis stage began with sentiment analysis. The sentiment analysis has

been conducted using MeaningCloud's Text Analytics add-in in Excel. MeaningCloud sentiment analysis produces two sheets. The first one is for the global sentiment analysis sorting includes the polarity, i.e., positive (P), negative (N), or Neutral (NEW), and a confidence score while the other one is for comments' topics sorting including topics categories and topics types (which were used later on in text mining).

2) *Regression analysis*: The second step in the analysis was identifying factors affecting users' satisfaction. To begin this step, the two datasets were merged. Ordinal logistic regression (OLR) test was chosen as suggested by [35], to answer the research question, by determining which of the independent variables (IV) have a statistically significant effect on our dependent variable (DV), which is polarity.

The reason for selecting the OLR test in this study is because the DV is an ordinal categorical variable where the value is represented with three points scale (N=0, NEW=1, and P=2) and the IVs are a mix of continuous and categorical. Multicollinearity test was conducted to ensure there are no highly correlated IVs, and only two IVs (accommodates and beds) were highly correlated with a result of >0.8; accordingly, the beds variable was removed from the test. Without performing the multicollinearity test, there would be a problem with determining the variable that contributes to the interpretation of the DV [35]. After eliminating the highly correlated variables, the final variables set contained 32 attributes as following: host experience in months, host response time, host response rate, the host is super-host, host total listings count, the host has profile pic, host identity verified, neighborhoods cleansed, is location exact, property type, room type, accommodates, bathrooms, bedrooms, bed type, square feet, price, weakly discount, monthly discount, security deposit, cleaning fee, extra people fee, instant bookable, cancellation policy, require guest phone verification, review scores rating, review scores accuracy, review scores cleanliness, review scores check-in, review scores communication, review scores location, and review scores value.

Taking the results from the OLR test, factors that affect polarity were visualized using Power BI software. Power BI was also used to conduct a text mining test to find the most frequent words within comments. Word's visualization helped to derive additional insights about repeated words in different polarity levels.

IV. FINDINGS AND DISCUSSION

A. Sentiment Analysis and Topic Clustering

The results of the sentiment analysis indicate that Airbnb users were mostly having a positive experience in Amsterdam (see Fig. 1). The topic clustering result shows that users were overwhelmingly positive about two aspects of their experience i.e. location and host (Table I). For instance, the likelihood score of 52% for a location means that reviews with the term location represent 52% of the positive sentiments. Location is an essential indicator for customers' satisfaction since they care about the proximity of different facilities to their residence such as public transportation, restaurants, park, station,

markets, and shops. As a second major topic within the positive sentiments, the term host represents 31.66% of the positive results; this indicates that many users were satisfied with the way their hosts were treating them. In contrast, the topic that consistently received a negative sentiment was check-in (Table II). The term Check-in represents the majority of the negative sentiment with a likelihood score of 95.69%. This result gives an indicator for hosts to improve their check-in services, to make this process quick, efficient, and smooth. This can be achieved by following the authoritative guide to the Airbnb check-in process. Fig. 2 shows word clouds that support the previous findings.

Although Airbnb's comments are mainly positive, taking a closer look at areas where negative sentiments have occurred, allows Airbnb hosts to address these areas by fixing any problems or setting future expectations (e.g., improving check-in services). Moreover, the strong appearance of topics related to the city's environment (such as location, facilities, transportation, restaurants, parks, stations, markets, and shops) among positive sentiments compared to their weak presence among negative ones, indicating that Amsterdam's general environment can play an essential role in shaping the Airbnb guests positive experience.

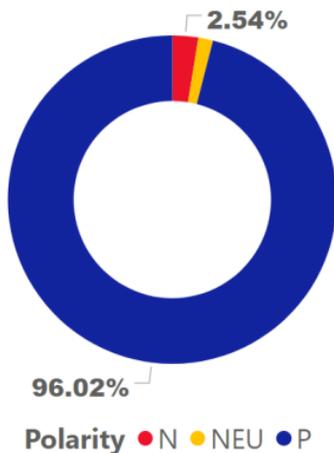


Fig. 1. Polarity Likelihood, Where N is Negative, NEU is Neutral, and P is Positive.



Fig. 2. Word Cloud for Airbnb Positive (Left) and Negative Reviews (Right).

TABLE I. WORDS' LIKELIHOOD SCORES OF SENTIMENT ANALYSIS FOR AIRBNB USERS' ONLINE REVIEWS

Positive			Negative		
Words	Likelihood	Count	Words	Likelihood	Count
Location	52.47%	52554	Check-in	95.69%	41473
Host	31.66%	31708	Host	1.63%	707
Facility	3.51%	3518	Location	0.84%	366
Clean	1.70%	1704	Room	0.35%	151
Room	1.50%	1504	Facility	0.30%	131
Walk	1.23%	1233	Bed	0.19%	81
Transportation	1.21%	1208	Walk	0.13%	58
Bed	1.18%	1185	Transportation	0.12%	51
Great	0.77%	767	Other Entity	0.09%	39
Top	0.57%	571	Service	0.06%	26
Restaurants	0.39%	395	Kitchen	0.06%	25
Process	0.32%	321	Top	0.06%	25
View	0.28%	282	Restaurants	0.05%	23
Other Entity	0.27%	274	Noise	0.05%	22
Park	0.27%	270	Station	0.05%	21
Amenities	0.26%	265	Amenities	0.04%	16
Station	0.26%	262	View	0.04%	16
Ship	0.24%	244	Water Form	0.03%	15
Service	0.20%	198	Shower	0.03%	14
Family	0.19%	190	Not-Clean	0.03%	13
Water Form	0.17%	175	Style	0.02%	10
Couple	0.16%	160	Couple	0.02%	8
Market	0.14%	143	Animal	0.02%	7
Style	0.14%	143	Park	0.02%	7
Check-in	0.14%	142	Market	0.01%	6
Kitchen	0.13%	132	Process	0.01%	5
Breakfast	0.12%	118	Coffee-Shop	0.01%	4
Coffee-Shop	0.10%	104	Family	0.01%	4
Shower	0.10%	99	Ship	0.01%	4
Animal	0.06%	56	Breakfast	0.01%	3
Quiet	0.05%	53	Shops	0.01%	3
House	0.03%	35	Airport	0.00%	2
Shops	0.03%	33	Event	0.00%	2
Staff	0.03%	32	Open	0.00%	1
Airport	0.02%	18	Plant	0.00%	1
Museum	0.01%	15	Sports	0.00%	1
Plant	0.01%	12			
Sports	0.01%	10			
Activities	0.01%	9			
Event	0.01%	7			
Open	0.00%	3			

TABLE II. ORDINAL REGRESSION RESULTS: PARAMETER ESTIMATES

Description	Attribute	Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Sentiment threshold where 0 is neutral	[Polarity = 0]	-16.742	9398.157	.000	1	.999	-18436.79	18403.307
The length of the time since a host registered on Airbnb (#months).	Host_Experience_in Months	-.030	.034	.764	1	.382	-.096	.037
The rate of one host responding to his guests	host_response_rate	.070	.060	1.343	1	.247	-.048	.188
The number of one host total listings	host_total_listings_count	.083	.149	.311	1	.577	-.209	.376
The number of guests accommodated by a given listing.	Accommodates	-.568	.468	1.471	1	.225	-1.485	.350
Number of bathrooms	Bathrooms	-.307	1.012	.092	1	.761	-2.291	1.676
Number of bedrooms	Bedrooms	.705	.683	1.065	1	.302	-.634	2.043
Listing price in dollar	Price	-.018	.007	5.957	1	.015*	-.032	-.003
Discount rate per week	Discount_Rate	.004	.039	.011	1	.918	-.073	.081
Discount rate per month	Discount_Monthly	.003	.029	.014	1	.905	-.054	.061
Some hosts require a security deposit for their listing	security_deposit	.003	.002	1.655	1	.198	-.002	.008
Extra fee for cleaning services	cleaning_fee	.015	.014	1.271	1	.260	-.011	.042
Extra fee for having extra people to reside in the property	extra_people_fee	.004	.013	.079	1	.779	-.022	.030
Property size in f ²	square_feet	1.614E-5	.001	.001	1	.982	-.001	.001
Within an hour = 1, a few hours = 2, hours = 3, a few days = 4	[host_response_time=1]	4.585	5.982	.587	1	.443	-7.140	16.310
	[host_response_time=2]	1.117	.987	1.279	1	.258	-.819	3.052
	[host_response_time=3]	.575	.766	.563	1	.453	-.926	2.076
	[host_response_time=4]	0 ^a	.	.	0	.	.	.
Each code represents a specific neighbourhood area within Amsterdam.	[neighbourhood_cleansed=1]	11.325	3394.217	.000	1	.997	-6641.218	6663.868
	[neighbourhood_cleansed=3]	.967	1.996	.235	1	.628	-2.946	4.880
	[neighbourhood_cleansed=4]	11.666	6271.111	.000	1	.999	-12279.48	12302.817
	[neighbourhood_cleansed=5]	-1.291	1.681	.589	1	.443	-4.586	2.005
	[neighbourhood_cleansed=6]	.726	1.731	.176	1	.675	-2.667	4.120
	[neighbourhood_cleansed=8]	.974	1.731	.316	1	.574	-2.419	4.366
	[neighbourhood_cleansed=9]	.017	1.620	.000	1	.992	-3.159	3.193
	[neighbourhood_cleansed=10]	-3.616	2625.809	.000	1	.999	-5150.107	5142.876
	[neighbourhood_cleansed=11]	17.599	2158.338	.000	1	.993	-4212.666	4247.863
	[neighbourhood_cleansed=12]	1.179	.000	.	1	.	1.179	1.179
	[neighbourhood_cleansed=13]	-.669	2.339	.082	1	.775	-5.254	3.915
	[neighbourhood_cleansed=14]	.286	2.058	.019	1	.889	-3.748	4.320
	[neighbourhood_cleansed=15]	.466	1.943	.057	1	.811	-3.343	4.275
	[neighbourhood_cleansed=16]	13.494	2137.889	.000	1	.995	-4176.692	4203.680
[neighbourhood_cleansed=17]	-1.021	1.821	.314	1	.575	-4.590	2.549	
[neighbourhood_cleansed=18]	0 ^a	.	.	0	.	.	.	
1= Private room in	[property_type=1]	-14.115	9398.155	.000	1	.999	-18434.16	18405.930

Description	Attribute	Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
apartment, 2= Private room in townhouse, 3= Private room in houseboat, 4= Entire guest suite, 5= Boat, 6= Private room in bed and breakfast, 7= Private room in guesthouse, 8= Houseboat, 9= Entire villa, 10= Entire apartment, 11= Entire townhouse, 12= Entire loft, 13= Entire serviced apartment, 14= Entire house	[property_type=2]	-12.641	9398.155	.000	1	.999	-18432.68	18407.405
	[property_type=3]	-14.589	9398.155	.000	1	.999	-18434.63	18405.456
	[property_type=4]	0 ^a	.	.	0	.	.	.
	[property_type=5]	-15.302	9398.155	.000	1	.999	-18435.34	18404.744
	[property_type=6]	-15.791	9398.155	.000	1	.999	-18435.83	18404.255
	[property_type=7]	3.182	9437.052	.000	1	1.00	-18493.10	18499.464
	[property_type=8]	-14.438	9398.155	.000	1	.999	-18434.48	18405.607
	[property_type=9]	2.391	9522.100	.000	1	1.00	-18660.58	18665.365
	[property_type=10]	3.605	11997.929	.000	1	1.00	-23511.90	23519.113
	[property_type=11]	-.888	.000	.	1	.	-.888	-.888
	[property_type=12]	-14.108	9398.155	.000	1	.999	-18434.15	18405.938
	[property_type=13]	-2.049E-6	10602.550	.000	1	1.00	-20780.616	20780.616
	[property_type=14]	0 ^a	.	.	0	.	.	.
	Shared room = 1, private room = 2 and entire home = 3.	[room_type=1]	2.022	1.056	3.667	1	.056	-.048
[room_type=2]		-1.423	1.668	.727	1	.394	-4.692	1.847
[room_type=3]		0 ^a	.	.	0	.	.	.
Airbed=1, Real bed=2, Couch=3	[bed_type=1]	.842	3.396	.061	1	.804	-5.813	7.497
	[bed_type=2]	14.717	3255.619	.000	1	.996	-6366.178	6395.612
	[bed_type=3]	0 ^a	.	.	0	.	.	.
Cancellation policies: flexible = 0, moderate = 1 and strict = 2.	[cancellation_policy=0]	-.474	1.543	.094	1	.759	-3.498	2.551
	[cancellation_policy=1]	-.625	1.450	.186	1	.666	-3.468	2.217
	[cancellation_policy=2]	0 ^a	.	.	0	.	.	.
A host with the superhost badge is labelled 0 and 1 otherwise.	[host_is_superhost=0]	-2.425	.815	8.848	1	.003**	-4.023	-.827
	[host_is_superhost=1]	0 ^a	.	.	0	.	.	.
Host picture presented	[host_has_profile_pic=1]	0 ^a	.	.	0	.	.	.
Verified host = 1, non-verified host = 0	[host_identity_verified=0]	1.076	.862	1.557	1	.212	-.614	2.766
	[host_identity_verified=1]	0 ^a	.	.	0	.	.	.
Exact location=0, not exact location=1	[is_location_exact=0]	1.952	1.919	1.034	1	.309	-1.810	5.714
	[is_location_exact=1]	0 ^a	.	.	0	.	.	.
Instant bookable=0, not instant bookable=1	[instant_bookable_A=0]	-.590	.732	.648	1	.421	-2.025	.846
	[instant_bookable_A=1]	0 ^a	.	.	0	.	.	.
Require guest phone verification=0, not required guest phone verification=1	[require_guest_phone_verification_A=0]	.315	.970	.106	1	.745	-1.585	2.216
	[require_guest_phone_verification_A=1]	0 ^a	.	.	0	.	.	.
Review ranges from 1 to 5 where 1 is the lowest rate and 5 is the highest rate	review_scores_rating	.035	.006	38.512	1	.000***		
	review_scores_accuracy	.168	.040	17.839	1	.000***		
	review_scores_cleanliness	.280	.029	94.662	1	.000***		
	review_scores_checkin	.142	.039	13.321	1	.000***		
	review_scores_communication	.176	.039	20.379	1	.000***		
	review_scores_location	.028	.030	.886	1	.347		
	review_scores_value	.095	.035	7.458	1	.006**		

B. Factors Influencing Comments Polarity

After applying an OLR test, the authors found that only nine out of thirty-two factors were influencing customers' satisfaction with a significant P value of < 0.05 (Table II). These affecting factors are price, room type (Entire Home/Apartment), the host is not a super-host, review scores rating, review scores accuracy, review scores cleanliness, review scores check-in, review scores communication, and review scores value. Surprisingly, the property neighborhood was not found impacting customers sentiment in Amsterdam which might be a case only limited to Amsterdam as it is one of the safest cities in the world.

1) *Price*: Price is the first-factor influences customers' satisfaction ($B= 5.957, p<0.05$). Fig. 3 illustrates the relationship between properties' daily prices and the polarity of comments sentiment. The figure shows that as the price raises, the negative reviews decline. In other words, positive comments are more commonly associated with high price properties comparing with negative ones. According to [14], properties with some features such as luxurious, penthouse, unique, chic & designed, duplex, Sauna, Jacuzzi & Spa, are relatively more expensive than other properties. It is, hence, expected that the benefits of obtaining luxurious characteristics, make customers willing to pay more.

Similarly, [36] indicates that some amenities are considered a price determinant. The availability of these amenities will give Airbnb customers the quality they are seeking. Eventually, they will be more satisfied and happier compared to customers who didn't obtain these amenities. Therefore, property price impacts customers' satisfaction indirectly since prices are determined by quality, which is critical to gain customers' satisfaction.

2) *Room type*: The second affecting factor is the room type, including the entire home/apartment, hotel room, private room, and shared room. Fig. 4 showing the percentage of negative and positive comments among the four property types (calculated as a percentage out of the same polarity). For example, entire home/apt negative comments represent 29% out of the total negative sentiments for the four types. The type of entire home/apt is more likely to get negative reviews than positive ones. After comparing the average of different attributes based on user sentiment (Positive versus Negative), the researchers found some factors that affect the entire home/apt polarity (see Fig. 5). The result indicates that decreasing the average square feet influences making negative feelings. Moreover, as host listing increased, the probability of getting a dissatisfied guest will be increased as well. Hence, it is evidence that there's a negative relationship between host total listings and customers' satisfaction. This could be due to hosts being busy handling all their properties simultaneously, therefore providing poor customer services. Eventually, hosts will fail to deliver the expected level of care to their guests.

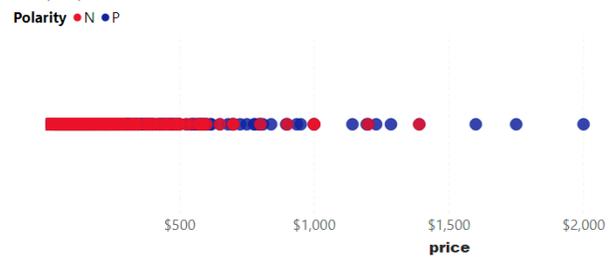


Fig. 3. Price Distribution for Negative and Positive Polarity.

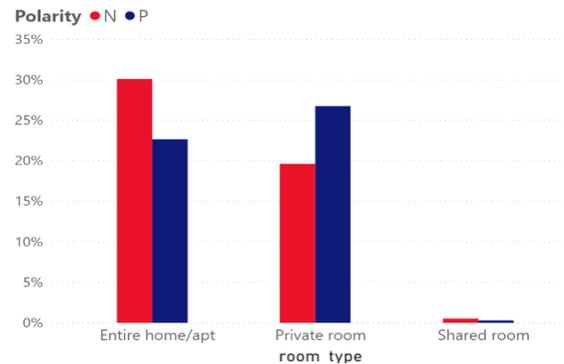


Fig. 4. Likelihood of Polarity, Grouped by Room Type.

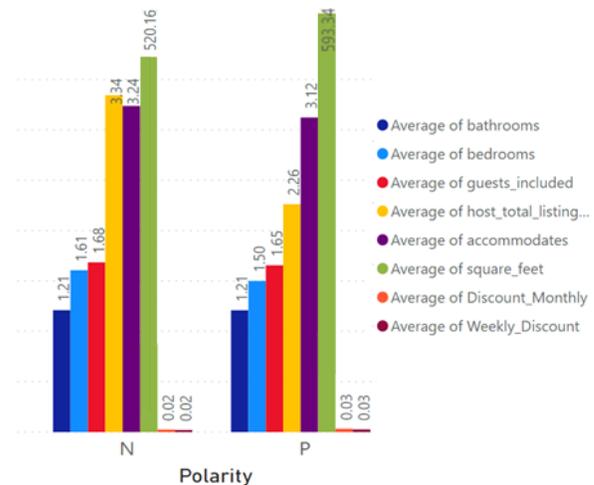


Fig. 5. Factors Influencing users' Polarity when Renting Entire Home/Apt.

3) *Super-host*: The third affecting factor is when the host is not a super-host. Fig. 6 is showing a similar distribution of super-host within the positive comments, which illustrates that whether a host is super-host or not, it will not affect the positivity of a sentiment. In contrast, the negative sentiments are widely affected by the host being not a super-host. Fig. 6 illustrated that 88.8% of unsatisfied customers were hosted by non-super-hosts. This supports [24] findings, who confirmed that poor services have a more significant impact on satisfaction level than good services. People tend to expect a certain level of quality. Obtaining this level will not positively

affect their satisfaction degree, but failing to reach this level will significantly affect their satisfaction negatively. According to [10], what distinguishes P2P accommodation from traditional hotels is to create a sense of place by providing a closer interaction with a resident (host). In such an aspect, online and face-to-face interactions with the host can influence customers' satisfaction, due to the atmosphere host creates. This fact supports our contribution here since low host interactions will lead to customer dissatisfaction.

4) *Neighborhood*: Surprisingly, though location plays an important role as an affecting element on polarity within topic clustering results (see Fig. 2), the OLR test provided evidence that neighborhood wasn't affecting the polarity (Table II). After a careful review of the results, it seems that the neighborhood by itself has no impact, while the exact location inside each neighborhood does. The reason for this situation is that different location within the same neighborhood district in terms of facilities, quietness, transportation, etc. To confirm OLR result, the relationship between polarity and neighborhoods were represented in Fig. 7 using maps and a bar chart. The maps show a similar distribution between the Ps and Ns comments, and the bar chart supported this result by indicating that each neighborhood has almost the same distribution rate of the Ps and Ns (top 15 neighborhoods). This could be due to the characteristic that made Amsterdam a unique city that everyone would like to visit. Such as its high quality of life, the European culture lifestyle, the small village feeling, and its beautiful landscape in all neighborhoods. One of these characteristics is safety, as [37] illustrated that Amsterdam is considered the fourth safest city in the world and the safest among all of Europe. The level of safety of all Amsterdam's neighborhoods is high in terms of personal security, health security, infrastructure security, and digital security. This support our findings that Amsterdam's neighborhoods do not affect Airbnb customer's satisfaction since the visitors (guests) will be positively biased toward this city and all of its neighborhood in general.

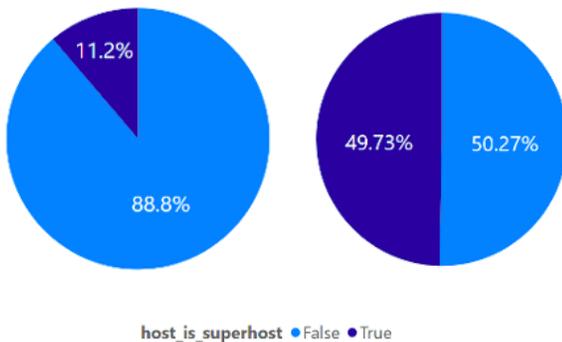


Fig. 6. Superhost Proportion in Positive Comments (Right) and in Negative Comments (Left).

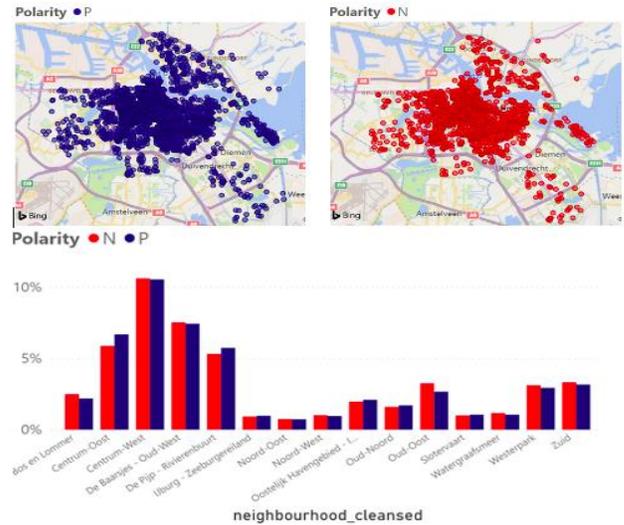


Fig. 7. Polarity Distribution based on Locations.

V. CONCLUSION AND FUTURE WORK

Airbnb represents a flourishing implementation of the sharing economy generally, and hospitality mainly. While Airbnb remains a topic of significant attention within the sector, preliminary literature on Airbnb has examined customers' reviews independently. This study is the first endeavor to analyze Airbnb's listings features as factors impacting review polarity. The authors found that prices, the host being a super-host, and room type are the main impacting factors on customers' satisfaction. Moreover, in the topic clustering test, positive comments contained "Location" as the most frequent word, while negative comments contained "Check-In". However, Amsterdam as a place has been confirmed that it's affecting Airbnb's customers' satisfaction and could be a motive to generate positive comments. This study was attempting to avoid the biased sample effect, by looking at factors as a percentage of each factor polarity out of the same polarity total in visualization. Regarding this attempt, there are some limitations to this study. The use of one sample such as Amsterdam could affect the result since there are limitations with a non-random sample.

This study highlights some possible orientations for future research. It would be beneficial to compare the results of different countries, to distinguish common factors versus cultural factors. Accordingly, this will help to confirm whether customers' satisfaction is affected by a country's general environment or not. Furthermore, it would be interesting to test text attributes such as room description and host "about" against polarity. Lastly, this study introduced a new method of determining factors affecting customers' satisfaction, by examining factors against polarity using sentiment analysis and regression test, then confirm the results using visualization. This approach can be applied to other studies in hospitality and beyond.

REFERENCES

- [1] Georgiadis L, Iosifidis G, Tassioulas L. On the Efficiency of Sharing Economy Networks. *IEEE Trans Netw Sci Eng* 2019;4697:1–1. <https://doi.org/10.1109/tNSE.2019.2904431>.
- [2] Ju Y, Back KJ, Choi Y, Lee JS. Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *Int J Hosp Manag* 2019; 77:342–52. <https://doi.org/10.1016/j.ijhm.2018.07.014>.
- [3] Sigala M, Rahimi R, Thelwall M. Big data and innovation in tourism, travel, and hospitality: Managerial approaches, techniques, and applications. *Big Data Innov Tour Travel Hosp Manag Approaches*, Tech Appl 2019;1–223. <https://doi.org/10.1007/978-981-13-6339-9>.
- [4] Cui L, Hou Y, Gao M, Yang Y. Exploring the influencing factors of sharing economy sustainability based on a two-mode social network analysis. 2019 16th Int Conf Serv Syst Serv Manag ICSSSM 2019 2019;1–5. <https://doi.org/10.1109/ICSSSM.2019.8887647>.
- [5] Martin-Fuentes E, Fernandez C, Mateu C, Marine-Roig E. Modelling a grading scheme for peer-to-peer accommodation: Stars for Airbnb. *Int J Hosp Manag* 2018;69:75–83. <https://doi.org/10.1016/j.ijhm.2017.10.016>.
- [6] Blal I, Singal M, Templin J. Airbnb's effect on hotel sales growth. *Int J Hosp Manag* 2018;73:85–92. <https://doi.org/10.1016/j.ijhm.2018.02.006>.
- [7] Guttentag DA, Smith SLJ. Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations. *Int J Hosp Manag* 2017;64:1–10. <https://doi.org/10.1016/j.ijhm.2017.02.003>.
- [8] Xie KL, Kwok L. The effects of Airbnb's price positioning on hotel performance. *Int J Hosp Manag* 2017;67:174–84. <https://doi.org/10.1016/j.ijhm.2017.08.011>.
- [9] Lin PMC, Fan DXF, Zhang HQ, Lau C. Spend less and experience more: Understanding tourists' social contact in the Airbnb context. *Int J Hosp Manag* 2019;83:65–73. <https://doi.org/10.1016/j.ijhm.2019.04.007>.
- [10] Moon H, Miao L, Hanks L, Line ND. Peer-to-peer interactions: Perspectives of Airbnb guests and hosts. *Int J Hosp Manag* 2019;77:405–14. <https://doi.org/10.1016/j.ijhm.2018.08.004>.
- [11] Su N, Mattila AS. Does gender bias exist? The impact of gender congruity on consumer's Airbnb booking intention and the mediating role of trust. *Int J Hosp Manag* 2019. <https://doi.org/10.1016/j.ijhm.2019.102405>.
- [12] Liang S, Schuckert M, Law R, Chen CC. The importance of marketer-generated content to peer-to-peer property rental platforms: Evidence from Airbnb. *Int J Hosp Manag* 2020;84:102329. <https://doi.org/10.1016/j.ijhm.2019.102329>.
- [13] Liu SQ, Mattila AS. Airbnb: Online targeted advertising, sense of power, and consumer decisions. *Int J Hosp Manag* 2017;60:33–41. <https://doi.org/10.1016/j.ijhm.2016.09.012>.
- [14] Falk M, Larpin B, Scaglione M. The role of specific attributes in determining prices of Airbnb listings in rural and urban locations. *Int J Hosp Manag* 2019; 83:132–40. <https://doi.org/10.1016/j.ijhm.2019.04.023>.
- [15] Yang SB, Lee K, Lee H, Koo C. In Airbnb we trust: Understanding consumers' trust-attachment building mechanisms in the sharing economy. *Int J Hosp Manag* 2019;83:198–209. <https://doi.org/10.1016/j.ijhm.2018.10.016>.
- [16] Wang C (Renee), Jeong M. What makes you choose Airbnb again? An examination of users' perceptions toward the website and their stay. *Int J Hosp Manag* 2018; 74:162–70. <https://doi.org/10.1016/j.ijhm.2018.04.006>.
- [17] Cheng M, Jin X. What do Airbnb users care about? An analysis of online review comments. *Int J Hosp Manag* 2019;76:58–70. <https://doi.org/10.1016/j.ijhm.2018.04.004>.
- [18] Luo Y, Tang R (Liang). Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *Int J Hosp Manag* 2019;80:144–54. <https://doi.org/10.1016/j.ijhm.2019.02.008>.
- [19] Zhang M, Luo X, Cheng X, Fu S. An empirical business study on service providers' satisfaction in sharing economy. *Proc - 2016 IEEE 1st Int Conf Data Sci Cyberspace, DSC 2016 2017*:514–9. <https://doi.org/10.1109/DSC.2016.115>.
- [20] Hawlitschek F, Teubner T, Gimpel H. Understanding the sharing economy - Drivers and impediments for participation in peer-to-peer rental. *Proc Annu Hawaii Int Conf Syst Sci* 2016;2016-March:4782–91. <https://doi.org/10.1109/HICSS.2016.593>.
- [21] Alamsyah A, Rochmah WY, Nugroho DDA. Understanding Public Opinion towards New Sharing Economy Business Model Using Content Analysis. *Proc 2019 Int Conf Inf Manag Technol ICIMTech 2019* 2019;1:300–4. <https://doi.org/10.1109/ICIMTech.2019.8843779>.
- [22] Sari R, Lisanti Y, Luhukay D. The Evaluation of Impact Sharing Economy Critical Success Factor Implementation (case study: Online hospitality). *Proc 2019 Int Conf Inf Manag Technol ICIMTech 2019* 2019;1:455–60. <https://doi.org/10.1109/ICIMTech.2019.8843746>.
- [23] Ch C, Gupta D. Factors influencing customer satisfaction with usage of shopping apps in India. *RTEICT 2017 - 2nd IEEE Int Conf Recent Trends Electron Inf Commun Technol Proc* 2017;2018-Janua:1483–6. <https://doi.org/10.1109/RTEICT.2017.8256844>.
- [24] Kumar V, Reinartz W. Applications of CRM in B2B and B2C Scenarios Part I. 2018. https://doi.org/10.1007/978-3-662-55381-7_16.
- [25] Zhao Y, Xu X, Wang M. Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *Int J Hosp Manag* 2019;76:111–21. <https://doi.org/10.1016/j.ijhm.2018.03.017>.
- [26] Lutz C, Newlands G. Consumer segmentation within the sharing economy: The case of Airbnb. *J Bus Res* 2018;88:187–96. <https://doi.org/10.1016/j.jbusres.2018.03.019>.
- [27] Edelman BG, Luca M. Digital Discrimination: The Case of Airbnb.com. *SSRN Electron J* 2014. <https://doi.org/10.2139/ssrn.2377353>.
- [28] Guo L, Li J, Wu J, Chang W, Wu J. A Novel Airbnb Matching Scheme in Shared Economy Using Confidence and Prediction Uncertainty Analysis. *IEEE Access* 2018;6:10320–31. <https://doi.org/10.1109/ACCESS.2018.2801810>.
- [29] Nguyen LS, Ruiz-Correa S, Mast MS, Gatica-Perez D. Check Out This Place: Inferring Ambiance from Airbnb Photos. *IEEE Trans Multimed* 2018;20:1499–511. <https://doi.org/10.1109/TMM.2017.2769444>.
- [30] Situmorang KM, Hidayanto AN, Wicaksono AF, Yuliatwati A. Analysis on customer satisfaction dimensions in peer-to-peer accommodation using latent dirichlet allocation: A case study of airbnb. *Int Conf Electr Eng Comput Sci Informatics* 2018;2018-October:542–7. <https://doi.org/10.1109/EECSI.2018.8752912>.
- [31] Park E. Motivations for customer revisit behavior in online review comments: Analyzing the role of user experience using big data approaches. *J Retail Consum Serv* 2019;51:14–8. <https://doi.org/10.1016/j.jretconser.2019.05.019>.
- [32] Cuquet M, Fensel A, Bigagli L. A European research roadmap for optimizing societal impact of big data on environment and energy efficiency. *GIoTS 2017 - Glob Internet Things Summit, Proc* 2017. <https://doi.org/10.1109/GIoTS.2017.8016274>.
- [33] Inside Airbnb n.d. <http://insideairbnb.com/>.
- [34] Li J, Lowe D, Wayment L, Huang Q. Text mining datasets of β -hydroxybutyrate (BHB) supplement products' consumer online reviews. *Data Br* 2020;30. <https://doi.org/10.1016/j.dib.2020.105385>.
- [35] Osborne JW. Simple Linear Models With Polytomous Categorical Dependent Variables: Multinomial and Ordinal Logistic Regression. *Regres Linear Model Best Pract Mod Methods* 2020:133–56. <https://doi.org/10.4135/9781071802724.n6>.
- [36] Chattopadhyay M, Mitra SK. Do airbnb host listing attributes influence room pricing homogeneously? *Int J Hosp Manag* 2019;81:54–64. <https://doi.org/10.1016/j.ijhm.2019.03.008>.
- [37] The Economist Intelligence Unit. Safe Cities Index 2019 2019.

Motivational Factors Impacting the Use of Citizen Reporting Applications in Saudi Arabia: The Case of Balagh Application

Muna M. Alhammad¹, Layla Hajar², Sahar Alshathry³, Mashael Alqasabi⁴

Management Information System Department, King Saud University, Riyadh, Saudi Arabia

Abstract—Citizen reporting applications are considered a new approach for interaction between government authorities and citizens. Citizen reporting applications are implemented to collectively gather information from citizens on issues related to public interest such as accidents, traffic violations, and commercial frauds. Through utilizing citizen reporting applications, citizens are able to provide information about incidents efficiently and conveniently to the local authorities via mobile applications that are designed for these specific purposes. For such applications to be successful, citizens' willingness to participate continually and to become daily users of these applications is required. This paper applies the self-determination theory to investigate the factors that encourage citizens to participate in citizen reporting applications. In this study, the factors impacting behavioural intention to use the applications are divided into two categories. First, intrinsic motivation factors that include self-concern, social responsibility, and revenge. Second, extrinsic motivation factors that include output quality and rewards. The study empirically surveyed 297 Saudi citizens from different age groups. The partial least square (PLS) approach validates the research model. Findings reveal that output quality, revenge, and self-concern are significantly associated with citizens' motivations to use the applications, whereas rewards and social responsibility do not significantly influence citizens' motivations to engage with such applications. This study contributes theoretically by enriching literature on the identification of the factors behind user's engagement in citizen reporting applications. It also contributes practically by supporting the developers of citizen reporting applications to consider these factors when designing and marketing this kind of application.

Keywords—Crowdsourcing; self-determination theory; intrinsic motivation; extrinsic motivation; citizens reporting

I. INTRODUCTION

Crowdsourcing is the utilization of various potential external contributors to accomplish a certain task via an online platform [1]. There are two types of users in crowdsourcing [2]. The first type is individuals or organizations who request tasks or assistance in solving problems. The second type is crowd members who offer to contribute to the problem solution.

Due to the unprecedented advancement of the Internet and online technologies, the process of crowdsourcing has become truly a phenomenon. Crowdsourcing applications are being used in several fields around the globe, such as in scientific research, education, business, and medicine [3]. Also, several

countries employed crowdsourcing applications as part of their disaster relief efforts. During the Colorado wildfires in 2012 and 2013, crowdsourcing played a significant role in allocating resources to the affected areas [4]. Such applications are also being used to report health cases and deploy disaster relief services instantaneously, such as during the cholera outbreak in Haiti in 2010, the dengue outbreak in Thailand and Indonesia in 2010, and most recently, during the COVID-19 pandemic in 2020 [4]. Several countries have developed crowdsourcing applications as an emergency response to the spread of COVID-19. For example, to decrease the spread of the disease, data on symptoms of COVID-19 were collected and analysed using smartphones to wirelessly crowd-map positive cases [4].

COVID-19 is a highly contagious disease that has forced several governments to impose quarantine restrictions around countries [5]. Saudi Arabia was one of the affected countries. The Saudi government, like many other governments, implemented several countermeasures to reduce the outbreak of COVID-19 among its citizens. For example, travellers who came directly or indirectly from China were examined in the Saudi airports and any recent travel to China needed to be declared into immigration and passport offices [6]. Despite the significant impact of these countermeasures, studies [6] have shown that these countermeasures would not be fully effective in preventing the spread of the virus. Further countermeasures were applied and a quarantine was imposed by the Saudi government to control the pandemic effectively. Educational institutions at all levels transformed all their activities online, regular face-to-face meetings were switched to online meetings and attendance to workplaces were suspended and telework was applied instead [7]. Saudi authorities emphasized the importance of social distancing and believed that it is a key measure to minimize the outbreak of the virus. Therefore, the Saudi government decided to close restaurants, grocery stores and shopping malls and allowed only home delivery services through websites or mobile applications [7]. As a result, e-commerce in Saudi Arabia increased by 400 per cent in April and May, 2020 causing 30,000 complaints about shipping delays [8]. These complaints were filed by citizens and handled by the Ministry of Commerce through an application called "Balagh App" [9]. Balagh app is a citizen-reporting application that was launched by the Ministry of Commerce to enable citizens to file commercial violations or frauds. The awareness towards citizen-reporting applications in Saudi Arabia, such as "Balagh app", has increased during the curfew due to the increased number of issues related to the delayed or missing

shipments. “Balagh app” is not only used to file complaints about online stores, but is, also, used to handle issues related to physical stores. In order for citizen-reporting applications to work effectively, citizens have to be motivated to utilize and participate in the applications. Therefore, this research aims to answer the following research question:

What are the factors that contribute the most in motivating users to engage in citizen reporting applications to report commercial issues?

Therefore, this research will address the motivational factors behind the use of citizen reporting applications to report commercial issues. The research will apply the self-determination theory to identify the main factors that enhance citizens’ engagement and participation in citizen-reporting applications.

II. LITERATURE REVIEW

A. Citizens Sourcing and Reporting

Zhao and Zhu [10] define crowdsourcing as a “collective intelligence system” that has three components: the organization, the crowd, and the platform. The organization in the context of this research is the government; the crowd component is the citizens; and the platform is the technology [11]. In the context of open government, crowdsourcing is also known as ‘citizen-sourcing’ or ‘citizen-reporting’ [12].

There is evidence that government agencies can utilize crowdsourcing to improve public services with lower costs, produce policy innovations, and engage larger numbers of public participants [13]. Estellés-Arolas and González-Ladrón-de-Guevara [12] believe that ‘crowdsourcing’ itself is sometimes used in the same breadth as many other concepts such as ‘co-creation’, ‘open innovation’, and ‘citizen-sourcing’. Governments use citizen-sourcing to engage citizens in information production that would help the government in its activities [14]. Therefore, governments can improve public services from the information generated by citizens via citizen reporting applications or online platforms. Citizen reporting is considered one of three sub-categories of citizen sourcing [14]. Citizens reporting allows citizens to report issues to governments and be part of the development process [15]. Several governments have developed citizen reporting applications for different purposes. In Switzerland, there is a citizen reporting application called “FixMyStreet”. The aim of the “FixMyStreet” application is to allow citizens to report issues regarding the infrastructure of the city of Zurich to the local government [15]. The Portuguese government also developed an android mobile application called Citizens@City. This application allows citizens to report urban problems such as holes in the pavement, lack of public lighting, or poor access to wheelchairs to the local authority. Citizens can open the application and report the problem expressed by a subject, description, location, and optional picture of the reported spot [16].

The government of Saudi Arabia realized the value of citizen sourcing and citizen reporting applications as well. The Ministry of Interior launched a citizen reporting application called “Kulluna Amn”. The purpose of “Kulluna Amn” is to make citizens part of the country’s security systems by

reporting incidents. Citizens are able to attach photos or videos with their reports and send them immediately to the concerned local authority. In addition, “Kulluna Amn” has a GPS system to determine the location of the incident. Similar to “Kulluna Amn”, the Ministry of Commerce and Investment launched a citizen reporting application called “Balagh”. “Balagh” enables customers to report any commercial violation through their smartphones with the possibility of attaching photos and locations of the violations. Customers can add the name of the stores, and type of the violations committed as well. The application additionally allows the customers to view the latest news and product recalls launched by the Ministry of Commerce and Investment. In order for the “Balagh” application to work efficiently and effectively, there is a need to motivate customers to participate and engage in the application. This paper will address the motivational factors behind user participation in reporting commercial violations using “Balagh”.

B. Motivational Theory

One of the most important topics in psychology over many years has been the studying of human behaviours and the motivations behind such behaviours. Abraham Maslow intensively studied motivation and articulated his first theory in 1943 under the title of ‘A Theory of human Motivation’. He illustrated in detail the human needs that motivate human behaviour. Motivation is the dynamic relationship between the individual and their surrounding environment that incites them to do something [17]. The topic of motivation has been studied from multiple angles. Although motivation is present in the actions of every individual, there are variations in the level and orientation of these motivations. The psychological literature differentiates between motives and motivation. According to Leimeister, et al. [18] “In the field of motivation psychology, a motive is seen as an individual’s psychological disposition”. Motivation is a combination of a person with specific motives and a situation, which gives certain incentives that trigger certain behaviour [19, 20]. Motives are relatively stable over the lifespan of an individual and do not automatically lead to certain actions. Typically, an activator is needed to initiate behaviour.

The question that arises in this context is what motivates individuals to adopt new technologies and accept them. There are many models that already exist in the field of Information Technology Acceptance that investigates how individuals come to accept new adoptions of information system technologies. This line of models and research also looks into ways to motivate potential users of such systems. Each model explains user behaviour using a different set of various factors. Examples of such models include the Theory of Reasoned Action (TRA) [19]; Self-Determination Theory (SDT) [21], the Theory of Planned Behaviour (TPB) [20]; the Technology Acceptance Model (TAM) [22]; and the Extension of the Technology Acceptance Model (TAM2) [23]. One of the most prominent theories that discuss the concept of motivation is the Self-Determination Theory (SDT) [21]. This theory classifies motivation into two categories: intrinsic and extrinsic. From an intrinsic motivation perspective, individual behaviour is driven by the sake of enjoyment for the activity itself rather than the instrumental value of that activity. On the other side, from an

extrinsic motivation perspective, the behaviour is done to attain some separable consequence, external prods, or reward [21]. In this research, the self-determination theory (SDT) will be applied and extended.

C. Factors Behind user Participation

There is a bulk of literature that examines the motivational factors that affect engagement in citizen-sourcing applications. For example, Alshehri, et al. [24] prove that performance expectancy, effort expectancy, and facilitating conditions have positive influences on user intention (i.e. motivation) to use e-government services in general. The study also found that social influence to be insignificant in terms of predicting the behavioural intention to use e-government services.

Lin [25] examines extrinsic motivations (i.e. expected organizational rewards and reciprocal benefits) and intrinsic motivations (i.e. knowledge self-efficacy, and enjoyment in helping others) of employees' attitudes and behaviour towards knowledge sharing. The results demonstrate that employee knowledge-sharing attitudes and behaviours are significantly influenced by mainly three motivational factors: knowledge self-efficacy, reciprocal benefits, and enjoyment from helping others. The study finds that organizational rewards do not have any significant impact but reports that employee attitudes toward knowledge-sharing do significantly influence behavioural intentions.

Another study by Kumar, et al. [26] shows that the adoption of e-Government is directly influenced by both user characteristics and the website design. User characteristics include perceived risks that are related to the user's financial status, performance risk, data privacy and security. Another user characteristic that is found to be significant in the determination of e-government adoption is the perceived control over the process and the degree of the user's Internet literacy (i.e. the amount of time users are exposed to the Internet, the usage frequency, and the length of time spent in each visit). Furthermore, website design factors comprise the perceived ease of use and the perceived usefulness of the website as explained on the technology acceptance model. The study finds that the adoption rate can be enhanced significantly by the user's perception of the usefulness of the services or online information that is provided by the government. Nevertheless, the perceived usefulness is associated with the perceived ease of use [26]. Service delivery quality also has an impact on user satisfaction, which in turn influences the adoption of e-Government.

Abu-Tayeh, et al. [15] determines how local authorities can be greatly benefited from the use of citizen reporting in the context of smart city areas. The study investigates the ZWN cross-channel platform that aims to enhance civic engagement in Switzerland. It shows that factors such as self-concern and other-orientation (i.e. concern towards others' wellbeing) motivate citizens to voluntarily use the platform in order to support the government. Nonetheless, several essential issues should be taken into consideration as well, such as socio-economic characteristics. For example, gender has been shown to influence the actual participation in citizen reporting systems (e.g. the average number of reports submitted by males and females can vary significantly).

III. HYPOTHESES AND MODEL DEVELOPMENT

The self-determination theory is a suitable theory for examining the motivational factors behind a user's adoption and use of citizen reporting applications. It allows for extending the traditional theory to include factors that might trigger both intrinsic and extrinsic motivations which in turn impact the individual's behavioural intention to use such applications. Engagement in citizen reporting applications can be either driven by an extrinsic motive, such as having a good personal image, expecting rewards, or maintaining better service quality for reporting such incidents; or driven by an intrinsic motive, such as having a sense of inner social responsibility or settling an issue related to self-concern. Therefore, it is hypothesis that:

H1: Extrinsic motivation positively influences users' behavioural intention to use citizen reporting applications.

H2: Intrinsic motivation positively influences users' behavioural intention to use citizen reporting applications.

In the following subsections, we define each of the components of the extrinsic and intrinsic motivations and construct the corresponding hypotheses for the proposed model (see Fig. 1).

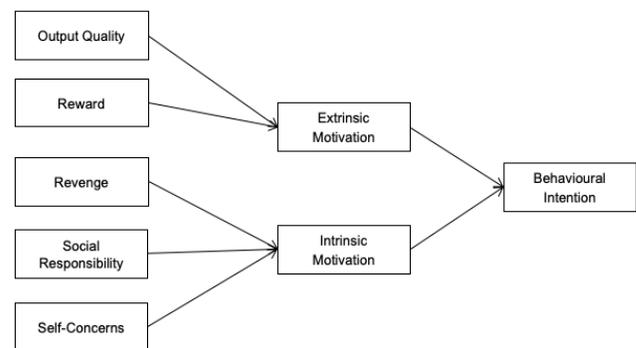


Fig. 1. Proposed Model.

A. Extrinsic Motivation Factors

Extrinsic motivation is closely related to the achievement and increase in value via the means of external benefits, monetary rewards, promotion, as well as other tangible rewards. It is thought of as a significant motivator in the literature relating to information systems. Davis, et al. [27], for instance, suggests that perceived usefulness is reflected by extrinsic motivation whilst intrinsic motivation relates to enjoyment.

1) *Output quality*: Output Quality refers to "an individual's perception about how well the system performs the tasks" [23]. Output quality in this context is the quality of services that are provided by governments to citizens. The government services can be enhanced effectively if citizens collaborate with their governments and participate in citizen reporting applications [15]. Output quality was found to influence the perceived usefulness as explained in TAM2 [23]. The perceived usefulness (PU) can generally be defined as "

measure of the individual's subjective assessment of the utility of an information technology in a specific task-related context" [23]. According to Venkatesh, et al. [28], perceived usefulness is a similar concept to extrinsic motivation. Therefore, we can conclude that the output quality will be one of the reasons that will influence extrinsic motivation toward using citizen reporting applications. In light of the aforementioned, we derive the following hypothesis.

H3: Output Quality positively influences a user's extrinsic motivation towards using citizen reporting.

2) *Reward*: Reward refers to the positive feedback as compensation or remuneration that can motivate citizens to perform the desired behaviours and to encourage them to share their knowledge about any system they used. Rewards can either be verbal or tangible. Verbal rewards, such as explicit positive performance feedback, are found to reinforce intrinsic motivation. Although some studies found that tangible rewards might weaken intrinsic motivation [17], several other studies found that monetary rewards encourage users and motivate them extrinsically to use the system [29, 30].

According to Assegaff, et al. [31], if employees believe they can receive organizational rewards by offering their knowledge, they will develop more positive attitudes towards knowledge sharing at VCoPs (Virtual Communities of Practices). The study, also, found that these rewards significantly affect employees' intentions to participate in VCoPs and share their knowledge. Accordingly, we imply that rewards will affect the extrinsic motivation and will have an indirect impact on behaviour intention to use citizens reporting applications. Therefore, we hypothesise that:

H4: Reward positively influences a user's extrinsic motivation towards using citizen reporting.

B. Intrinsic Motivation Factors

Intrinsic motivation is considered when tackling the success and adoption of information systems. Intrinsic motivation refers to the engagement in behaviour that is inherently satisfying or enjoyable [32]. Enjoyment influences intrinsic motivation positively to the extent that "the activity of using a specific system is perceived to be enjoyable in its own right, aside from any performance consequences resulting from system use" [33].

1) *Revenge*: Definitions of customer revenge in the literature are consistent to a greater or lesser extent. Customer revenge involves a customer exerting some harm to a firm in return for the perceived damages the firm has caused [34]. Previous studies show that a desire for revenge (i.e. a felt need to exert harm) increases the likelihood of "tangible" revenge behaviours [35]. An emphasis on a desire for revenge is important because users are not always able, depending on the context, to transform their desire into actions. This leaves room for the incorporation of moderators (such as power) that could explain when a desire for revenge produces real manifestations that are designed to harm the firm. Hence, we

assume that revenge will have a major effect on individuals' intrinsic motivation and that it will indirectly affect individual's behavioural intention to use citizens reporting applications. Therefore, we hypothesise that:

H5: Revenge positively influences a user's intrinsic motivation towards using citizen reporting.

2) *Self-concern*: Abu-Tayeh, et al. [15] found that self-concern and other-orientation are important factors that impact citizen reporting engagement. They also found that self-concern has the strongest effect among other orientations on engagement. Self-concern can be defined as the inclination to base one's behaviour on the desire to protect and enhance one's self-interest. Citizens can be motivated to engage in citizen reporting because they expect to benefit from it. For example, when the government settles an issue, they had previously reported, because it personally irritated them. This demonstrates that citizens are largely self-concerned. Nevertheless, if citizens report infrastructure issues as they desire to help others, other orientations are the main motive. In a study done by De Dreu and Nauta [36], they state that self-concern can be raised by setting higher aspirations. Research on social dilemmas has shown that people have higher self-concern when potential outcomes are outlined as losses rather than gains. Therefore, we suggest that self-concern impacts an individual's intrinsic motivation which influences individuals' behavioural intention. Hence, we hypothesise that:

H6: Self-concern positively influences intrinsic motivation towards using citizen reporting.

3) *Social responsibility*: Social responsibility is considered as the desire for an individual to help others and participate in a positive way to provide benefits to the society. One of the means of social responsibility is the participation in crowdsourcing applications. Previous studies have shown that social responsibility is positively associated with individuals' engagement in citizen reporting [15].

The study of Razmerita, et al. [37] illustrated that altruism, i.e. the enjoyment of helping others, is the most important factor that impacts the frequency of knowledge sharing. Consequently, we will propose that social responsibility will affect individual's intrinsic motivation to use citizens reporting applications. Therefore, we hypothesise that:

H7: Social responsibility positively influences intrinsic motivation towards using citizen reporting.

IV. RESEARCH METHODOLOGY

A. Sample and Data Collection

The aim of this study is to identify the factors that motivate citizens to engage in citizen reporting applications to report commercial issues. To test the proposed hypotheses, we select a sample of users from the citizen reporting mobile application "Balagh" (translation: "report app") in Saudi Arabia. The Balagh app enables citizens to report commercial violations or frauds. Users' intrinsic and extrinsic motivational factors

toward the usage of this app were collected using a close-ended structured questionnaire. SmartSurvey was used to design the questionnaire then the link was distributed through online channels like emails, WhatsApp, and social media accounts. In order to examine the accuracy, meaningfulness, and clarity of the questionnaire, the questionnaire was pilot-tested and filled by 30 respondents who were chosen randomly. Thereafter, the survey was distributed to 500 individuals, and 297 responses were received. Table I shows the demographic information of the respondents.

B. Measurements

All of the 26 items used in this research were developed using items validated in prior studies and adapted to fit the purpose of this research. Measurements items used in this research and the source of these measurements are presented in Table II. Each factor that might influence the behavioural intention and motivate the citizens to engage with the Balagh App was measured by five-point Likert-scale questions (i.e. from strongly disagree “1” to strongly agree “5”).

C. Data Analytical Procedure

The study uses the Partial Least Squares (PLS) method to analyse the collected data. The PLS path modelling is a variance-based Structural Equation Modelling (SEM) technique that is widely applied in many research fields such as: business and social research for its ability to model composites and factors. This makes it suitable for studying the adoption and engagement with new technology and an ideal method to tackle the higher-level statistical performance of the study. Therefore, this method of analysis was chosen due to its proven validity to test models and relationships as well as its usefulness in test models in the case of a limited number of participants [41].

The analysis was done using SmartPLS software version 3.3.2. for Mac to evaluate the measurements and test the structural models simultaneously. Using SmartPLS, two steps were taken to examine the model. The first step was confirming the reliability and validity. The second step was analysing the structural equation model to evaluate the hypotheses.

TABLE I. RESPONDENTS DEMOGRAPHIC INFORMATION (N = 297)

Variable	Value	Frequency	Percentage
Age	Under 18	7	2.4%
	18-24	28	9.4%
	25-34	76	25.6%
	35-54	153	51.5%
	Over 55	33	11.1%
Gender	Male	40	13.5%
	Female	257	86.5%
Educational level	High school	66	22.2%
	Bachelor degree	187	63.0%
	Master degree	24	8.1%
	PhD.	6	2.0%
	Other	14	4.7%

TABLE II. MEASUREMENT ITEMS

VBs	Items	Source
Behavioural Intention	BI1: I am considering using this app to report incidents. BI2: I would seriously contemplate using this app. BI3: It is likely that I am going to use this app. BI4: I am likely to make future reports using this app.	Adapted from Davis, et al. [27] and Venkatesh, et al. [28]
Extrinsic motivation	EM1: I find using this application useful. EM2: Using this application enables me to report commercial incidents more efficiently. EM3: I can forward my commercial concerns to local government directly.	
Intrinsic motivation	IM1: I participate in this application because I think that this participation is interesting. IM2: I participate in this application because this participation is fun IM3: I participate in this application because I feel good when doing this reporting	
output quality	OQ1: The use of this application will improve the quality of provided commercial services. OQ2: The use of this application will contribute to the development of commercial services. OQ3: Using this application will enhance the overall quality and efficiency of the provided commercial services.	Adapted from Compeau and Higgins [38]
Reward	RW1: My willingness to participate in this application would increase if there were monetary rewards RW2: I will really like to participate in this application if I would receive monetary rewards in return for my knowledge sharing.	Adapted from Wijnhoven, et al. [39]
Revenge	RV1: My feeling of anger towards violators pushes me to use the application and report them RV2: I use this app to publicize the practices of violators and punish them. RV3: I submit a complaint via the application to avenge violators	Adapted from Aquino, et al. [40]
Self-concerned	SC1: I took part in “Balag” because I could report problems that concerned me personally. SC2: I took part in “Balag” because I could report problems that prevented me from fulfilling my needs SC5: I participate in this application because I believe that this kind of reporting is important for me	Adapted from Abu-Tayeh, et al. [15]
Social responsibilities	SR1: I took part in Balag application because it gives me the opportunity to protect others from commercial fraud. SR2: I took part in this application because I could help the community by doing so. SR3: I want to contribute to the development of the commercial services provided in my city by using this application. SR4: I would feel bad about myself if I don’t share information about commercial violations with the relevant authorities. SR5: I participate in this application because I feel that this is something that I have to do it for the society.	Adapted from both Abu-Tayeh, et al. [15] and Schmidhuber, et al. [32]

V. RESULTS

A. Measurement Model

In this step, we present the assessment of the reliability and validity of the measurement model. Cronbach’s alpha and composite reliability (CR) were used to assess the reliability while convergent validity and discriminant validity were used to check the validity of our model. Table III shows 24 out of 26 indicators have factor loadings above 0.7. Only two indicators, rewards and intrinsic motivation, have a lower factor loading. Intrinsic motivation has a factor loading of 0.665 while reward has a factor loading of 0.507 which, according to Hair, et al. [42], are still above the minimum acceptable threshold value of 0.5 for factor loadings. The t-values also demonstrate a satisfactory reliability level of all indicators as all the indicators are significantly linked with their corresponding constructs ($p < 0.001$) [43]. Furthermore, Cronbach’s alpha (α) of all variables exceeds the threshold of 0.7 [44] except for two factors (i.e. intrinsic motivation and revenge). Intrinsic motivation has a Cronbach’s alpha (α) value of 0.669 while revenge has a value of 0.618. According to Hair, et al. [45], any variable with Cronbach’s alpha (α) value above 0.6 still shows acceptable level of internal consistency. Therefore, all of our constructs show a good level of internal consistency. Table III also shows that all composite reliability (CR) values are greater than the recommended value of 0.7 [43]. It also shows that the values of all the average variance extracted (AVE) of our constructs exceed the critical threshold value of 0.5 showing a good convergent validity [46]. Lastly, the correlation of the square root of AVE for all of the constructs, as shown in Table IV, are higher than their correlations with other constructs, proving the discriminant validity of our measurement model.

B. Structural Model Assessment

As the reliability and validity of our measurement model have been confirmed, we proceeded with testing of the structural model. This step is used to examine the relationship between variables and to assess the model’s predictive capabilities. The significance of path coefficients and the R2 are the key criteria for assessing the structural model. The nonparametric bootstrapping procedure in the PLS analysis with 500 samples was performed to examine the structural models and to calculate the path coefficients and their significance levels.

As shown in Table V, all of our hypotheses are supported except two hypotheses that are rejected. Extrinsic motivation has a positive significant relationship with behavioural intention ($\beta = 0.553$, $t = 9.557$, $p < 0.001$) and intrinsic motivation also has a positive significant relationship ($\beta = 0.220$, $t = 4.576$, $p < 0.000$) supporting both H1 and H2. In addition, output quality positively influences extrinsic motivation ($\beta = 0.696$, $t = 18.940$, $p < 0.001$) while reward has no significant relationship with extrinsic motivation ($\beta = -0.084$, $t = 1.215$, $p < 0.225$); supporting H3 and rejecting H4. Intrinsic motivation is also found to be positively impacted by revenge ($\beta = 0.384$, $t = 6.273$, $p < 0.001$) and self-concern ($\beta =$

0.258 , $t = 3.139$, $p < 0.001$) supporting H5 and H6. On the other hand, social responsibility does not significantly impact intrinsic motivation ($\beta = 0.258$, $t = 3.139$, $p > 0.121$); rejecting H7. To determine the predictive power of our model, the value of R² is used. R² is a statistical measure that reports the variance in the dependent variable that can be explained by the independent variables. In general, the model managed to explain the variance in a user’s behavioural intention by 45.6% ($R^2=0.456$). The identified extrinsic motivational factors, i.e. output quality and reward, were able to explain variance in users’ extrinsic motivation by 49.9% ($R^2=0.499$). On the other hand, revenge, self-concern, and social responsibility were able to explain variance in intrinsic motivation by 43% ($R^2=0.431$). All R² values are higher than 0.1 which is the minimum acceptable level as defined by [46]; indicating a good predictive power for our model. Fig. 2 and Table V summarize the analysis results.

TABLE III. THE MEASUREMENT MODEL STATISTICS

Variable	Items	Loading	T-value	α	CR	AVE
Behavioural Intention	BI1	0.809	31.809	0.848	0.897	0.686
	BI2	0.834	25.718			
	BI3	0.819	32.454			
	BI4	0.851	42.551			
Extrinsic motivation	EM1	0.812	31.719	0.729	0.847	0.648
	EM2	0.788	26.573			
	EM3	0.815	29.125			
Intrinsic motivation	IM1	0.665	9.473	0.669	0.797	0.569
	IM2	0.758	15.359			
	IM3	0.831	27.995			
output quality	OQ1	0.845	35.881	0.820	0.893	0.736
	OQ2	0.850	36.877			
	OQ3	0.878	44.257			
Reward	RW1	0.507	1.475	0.816	0.735	0.603
	RW2	0.974	3.007			
Revenge	RV1	0.825	34.076	0.618	0.798	0.569
	RV2	0.723	16.144			
	RV3	0.710	15.628			
Self-concerned	SC1	0.857	38.582	0.717	0.839	0.637
	SC2	0.713	15.983			
	SC3	0.816	22.547			
Social responsibilities	SR1	0.831	39.190	0.863	.900	0.644
	SR2	0.798	24.692			
	SR3	0.778	20.510			
	SR4	0.773	33.179			
	SR5	0.829	35.497			

TABLE IV. DISCRIMINANT VALIDITY: SQUARE ROOT OF AVE

Variables	BI	EM	IM	OQ	RV	RW	SC	SR
Behavioural Intention	0.828							
Extrinsic Motivation	0.645	0.805						
Intrinsic Motivation	0.451	0.416	0.755					
Output Quality	0.672	0.701	0.386	0.858				
Revenge	0.512	0.509	0.605	0.503	0.754			
Reward	-0.071	-0.131	0.114	-0.068	0.082	0.777		
Self-concerned	0.633	0.633	0.567	0.577	0.623	0.054	0.798	
Social Responsibilities	0.758	0.750	0.487	0.728	0.555	-0.143	0.640	0.802

TABLE V. OVERVIEW OF THE HYPOTHESES TEST RESULTS

Hypotheses number	Hypotheses	β	T-value	P-Values	Supported/ not supported
H1	Extrinsic Motivation \rightarrow Behavioural Intention	0.553	9.557	0.000	supported
H2	Intrinsic Motivation \rightarrow Behavioural Intention	0.220	4.576	0.000	supported
H3	Output Quality \rightarrow Extrinsic Motivation	0.696	18.940	0.000	supported
H4	Reward \rightarrow Extrinsic Motivation	-0.084	1.215	0.225	Rejected
H5	Revenge \rightarrow Intrinsic Motivation	0.384	6.273	0.000	supported
H6	Self-concerned \rightarrow Intrinsic Motivation	0.258	3.139	0.002	supported
H7	Social Responsibilities \rightarrow Intrinsic Motivation	0.109	1.553	0.121	Rejected

VI. DISCUSSION AND CONCLUSION

The results indicate that both extrinsic and intrinsic motivational factors influence a citizen’s behavioural intention to report commercial incidents using citizens’ reporting applications. It is found that extrinsic motivation is much more influential than intrinsic motivation, nonetheless. Citizens’ extrinsic motivation is significantly influenced by output quality. However, rewards are found to have no significant influence on a citizen’s extrinsic motivation to report commercial incidents. Mainly, citizens are more concerned about the output quality that results from their engagement with the citizens’ reporting applications for them to be motivated to report commercial incidents. This result is aligned with previous studies, such as [23] and [27], which demonstrated that individuals will be more motivated and encouraged to use a system if they believe the system will help them achieve their desired outcomes. According to [47], it is important to keep citizens informed about the improvements and actions taken based on their reports so that they feel that their reports are taken seriously. On the other hand, though reward was expected to impact a user’s extrinsic motivation to participate in the system, it was not found significant. Previous research studying the role of rewards on impacting users’

extrinsic motivation in several other contexts is conflicted. Some studies, such as [29-31], found that rewards play a significant role in encouraging users to participate, while others such as [17], found no significant role in influencing users’ extrinsic motivation but instead it may weaken users’ intrinsic motivation to use the system. This indicates that government institutions should concentrate on emphasising the effectiveness of citizens’ participation to report commercial incidents on improving the quality of delivering commercial services within the city rather than offering rewards in return for their participation.

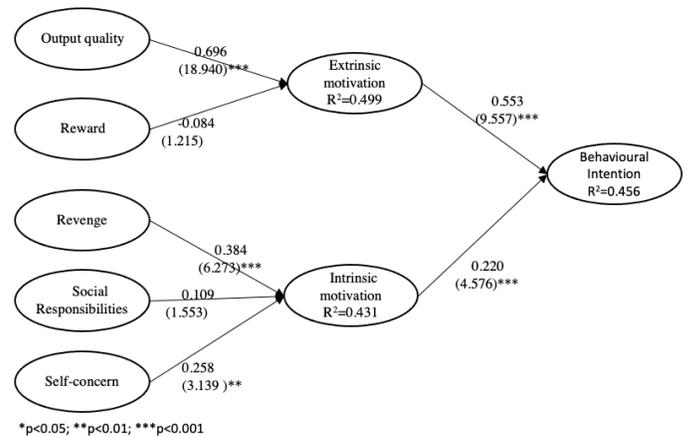


Fig. 2. The Results of the Empirical Study.

On the other hand, citizens’ intrinsic motivation is found to be influenced by self-concern. This is supported by existing literature in the field, such as the work of [15], where they found that self-concern plays a significant role in the actual number of reports filed by an individual on a specific citizen’s reporting application designed to enhance the infrastructure of the city. According to [48], individuals who are highly concerned about their self-gain are found to be more persistent contributors. As this study concentrates on reporting commercial incidents, it makes sense that citizens will report issues that impact them directly to solve their own problems and satisfy their needs. In alignment with this finding, we also found that revenge plays a significant role. When citizens personally face a problem with a certain commercial institution, such as shopping fraud or problems with their purchase, and the institution neglects their complaint, some of them feel anger and experience the urge to seek revenge. Thus, they feel that the use of citizen’s reporting applications to report commercial incidents is a way for punishing the accountable commercial institution and gaining revenge. According to [34], typically customers feel that revenge is an appropriate response, especially after a series of failed amendment requests. Therefore, as citizens’ reporting applications facilitate the process of restoring their rights and forcing the responsible commercial institution to solve the problem, it is justified that revenge plays a significant role in the context of this study. On the other hand, social responsibility is empirically found to have no influence on a user’s intrinsic motivation. This contradicts the findings of [15] where they empirically found that other-orientation is a significant predictor of citizens’ participation in citizens’

reporting applications. This contradiction can be due to the differences in the context of the studies. The authors in [15] studied the motivational factors behind citizens' participation in reporting incidents related to infrastructure, while the context of this study is related to reporting commercial incidents. When reporting commercial incidents, citizens are mostly aware of the issues that impact them directly and are hardly aware of the issues that face others. Therefore, although social responsibility is important, its role in the context of this study is proven to be limited and of no significant impact. Hence, when reporting commercial incidents, citizens participate mainly for self-gain rather than being concerned about the quality of commercial services provided to others. Government institutions running this kind of applications should market their applications by emphasising the personal benefits that users can acquire as a result of using this kind of applications.

As per the above discussion, this study is limited to the context of citizens' reporting applications that are mainly designed to report commercial incidents. Therefore, its result cannot be generalised to other contexts. Future research should concentrate on studying the impact of context in moderating the motivational factors behind citizens' behavioural intention to use citizens' reporting applications. It will also be useful to investigate factors impacting the quality and accuracy of reported incidents and how these factors can be reflected in the design of citizens' reporting applications.

REFERENCES

- [1] S. L. Alam, R. Sun, and J. Campbell, "Helping yourself or others? motivation dynamics for high-performing volunteers in GLAM crowdsourcing," *Australasian Journal of Information Systems*, vol. 24, pp. 1-25, 2020, doi: 10.3127/AJIS.V24I0.2599.
- [2] M. Weiss, "Crowdsourcing Literature Reviews in New Domains," *Technology Innovation Management Review*, vol. 6, no. 2, pp. 5-14, 2016, doi: 10.22215/timreview963.
- [3] D. Smith, M. Mehdi Gharaei Manesh, and A. Alshaikh, "How Can Entrepreneurs Motivate Crowdsourcing Participants?," *Technology Innovation Management Review*, vol. 3, no. 2, pp. 23-30, 2013, doi: 10.22215/timreview657.
- [4] A. Desai et al., "Crowdsourcing a crisis response for COVID-19 in oncology," *Nature Cancer*, vol. 1, no. 5, pp. 473-476, 2020, doi: 10.1038/s43018-020-0065-z.
- [5] "Coronavirus Disease 2019 (COVID-19) Situation Reports. April 1 2020," *World Health Organization*, 2020, vol. 2019. [Online]. Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200324-sitrep-64-covid-19.pdf?sfvrsn=703b2c402%0Ahttps://www.who.int/docs/default-source/coronaviruse/situation-reports/20200401-sitrep-72-covid-19.pdf?sfvrsn=3dd8971b_2.
- [6] M. Barry, M. Al Amri, and Z. A. Memish, "Covid-19 in the shadows of MERS-CoV in the Kingdom of Saudi Arabia," 2020, vol. 10.
- [7] A. Saudi Press, "Kingdom's government decides to suspend attendance at workplaces in all government agencies for period of (16) days except for health, security, military and electronic security centre," ed. 2020.
- [8] A. Bashraheel, "Online shopping in Saudi Arabia jumps 400% during coronavirus pandemic," in *Arabnews*, ed. 2020.
- [9] A. Al-Awsat, "Saudi Arabia Investigates First Stockpiling Attempt," ed. 2020.
- [10] Y. Zhao and Q. Zhu, "Evaluation on crowdsourcing research: Current status and future direction," *Information Systems Frontiers*, vol. 16, no. 3, pp. 417-434, 2014.
- [11] K. Cupido and J. Ophoff, "A Model of Fundamental Components for an e-Government Crowdsourcing Platform," *Electronic Journal of e-Government*, vol. 12, no. 2, pp. 141-156, 2014.
- [12] E. Estellés-Arolas and F. González-Ladrón-de-Guevara, "Towards an integrated crowdsourcing definition," *Journal of Information Science*, vol. 38, no. 2, pp. 189-200, 2012/04/01 2012, doi: 10.1177/0165551512437638.
- [13] P. Dutil, "Crowdsourcing as a new instrument in the government's arsenal: Explorations and considerations," *Canadian Public Administration*, vol. 58, no. 3, pp. 363-383, 2015, doi: <https://doi.org/10.1111/capa.12134>.
- [14] D. Linders, "From e-government to we-government: Defining a typology for citizen coproduction in the age of social media," 2012, doi: 10.1016/j.giq.2012.06.003.
- [15] G. Abu-Tayeh, O. Neumann, and M. Stuermer, "Exploring the Motives of Citizen Reporting Engagement: Self-Concern and Other-Orientation," *Business & Information Systems Engineering*, 2018, doi: 10.1007/s12599-018-0530-8.
- [16] A. M. Ribeiro, R. P. Costa, L. Marcelino, and C. Silva, "Citizens@City Mobile Application for Urban Problem Reporting," in *ENTERprise Information Systems*, Berlin, Heidelberg, M. M. Cruz-Cunha, J. Varajão, P. Powell, and R. Martinho, Eds., 2011// 2011: Springer Berlin Heidelberg, pp. 141-150.
- [17] E. L. Deci, R. Koestner, and R. M. Ryan, "Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again," *Review of Educational Research* Spring, vol. 71, no. 1, pp. 1-27, 2001, doi: 10.3102/00346543071001001.
- [18] J. M. Leimeister, M. Huber, U. Bretschneider, and H. Krcmar, "Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition," *Journal of Management Information Systems*, vol. 26, no. 1, pp. 197-224, 2009/07/01 2009, doi: 10.2753/MIS0742-1222260108.
- [19] M. Fishbein and I. Ajzen, *Belief, attitude, intention and behaviour: An introduction to theory and research*. Reading, Mass: Addison-Wesley, 1975.
- [20] I. Ajzen, "The theory of planned behavior," *Organizational behavior and human decision processes*, vol. 50, no. 2, pp. 179-211, 1991.
- [21] E. L. Deci and R. M. Ryan, *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer, 1985.
- [22] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models," *Management science*, vol. 35, no. 8, pp. 982-1003, 1989.
- [23] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management science*, vol. 46, no. 2, pp. 186-204, 2000.
- [24] M. Alshehri, S. Drew, and R. AlGhamdi, "Analysis of citizens acceptance for e-government services: applying the UTAUT model," *arXiv preprint arXiv:1304.3157*, 2013.
- [25] H. F. Lin, "Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions," *Journal of Information Science*, vol. 33, no. 2, pp. 135-149, 2007, doi: 10.1177/0165551506068174.
- [26] V. Kumar, B. Mukerji, I. Butt, and A. Persaud, "Factors for Successful e-Government Adoption : a Conceptual Framework," *The Electronic Journal of e-Government*, vol. 5, no. 1, pp. 63-76, 2007.
- [27] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "Extrinsic and intrinsic motivation to use computers in the workplace," *Journal of applied social psychology*, vol. 22, no. 14, pp. 1111-1132, 1992.
- [28] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *Management Information Systems Quarterly*, pp. 425-478, 2003.
- [29] B. E. P. Thapa, B. Niehaves, C. E. Seidel, and R. Plattfaut, "Citizen involvement in public sector innovation: Government and citizen perspectives," *Information Polity*, vol. 20, pp. 3-17, 2015, doi: 10.3233/IP-150351.
- [30] N. Shoemaker, "Extrinsic Rewards, Knowledge Sharing, and Self-Determined Motivation," vol. 11, no. 3, pp. 99-114, 2014.
- [31] S. Assegaff, Kurniabudi, and E. Fernando, "Impact of extrinsic and intrinsic motivation on knowledge sharing in virtual communities of practices," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 1, no. 3, pp. 619-626, 2016, doi: 10.11591/ijeecs.v1.i3.pp619-629.

- [32] L. Schmidhuber, D. Hilgers, T. Gegenhuber, and S. Etzelstorfer, "The emergence of local open government: Determinants of citizen participation in online service reporting," *Government Information Quarterly*, vol. 34, no. 3, pp. 457-469, 2017/09/01/ 2017, doi: <https://doi.org/10.1016/j.giq.2017.07.001>.
- [33] C. M. Chao, "Factors determining the behavioral intention to use mobile learning: An application and extension of the UTAUT model," *Frontiers in Psychology*, vol. 10, pp. 1-14, 2019, doi: 10.3389/fpsyg.2019.01652.
- [34] Y. Grégoire, D. Laufer, and T. M. Tripp, "A comprehensive model of customer direct and indirect revenge: Understanding the effects of perceived greed and customer power," *Journal of the Academy of Marketing Science*, vol. 38, no. 6, pp. 738-758, 2010, doi: 10.1007/s11747-009-0186-5.
- [35] H. Zourrig, J.-C. Chebat, and R. Toffoli, "Consumer revenge behavior: A cross-cultural perspective," *Journal of Business Research*, vol. 62, no. 10, pp. 995-1001, 2009, doi: <https://doi.org/10.1016/j.jbusres.2008.08.006>.
- [36] C. K. W. De Dreu and A. Nauta, "Self-Interest and Other-Orientation in Organizational Behavior: Implications for Job Performance, Prosocial Behavior, and Personal Initiative," in *Journal of Applied Psychology* vol. 94, ed, 2009, pp. 913-926.
- [37] L. Razmerita, K. Kirchner, and P. Nielsen, "What factors influence knowledge sharing in organizations? A social dilemma perspective of social media communication," *Journal of Knowledge Management*, vol. 20, no. 6, pp. 1225-1246, 2016, doi: 10.1108/JKM-03-2016-0112.
- [38] D. R. Compeau and C. A. Higgins, "Computer self-efficacy: Development of a measure and initial test," *Mis Quarterly*, vol. 19, no. 2, pp. 189-211, 1995.
- [39] F. Wijnhoven, M. Ehrenhard, and J. Kuhn, "Open government objectives and participation motivations," *Government Information Quarterly*, vol. 32, no. 1, pp. 30-42, 2015, doi: 10.1016/j.giq.2014.10.002.
- [40] K. Aquino, T. M. Tripp, and R. J. Bies, "How employees respond to personal offense: the effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace," (in eng), *The Journal of applied psychology*, vol. 86, no. 1, pp. 52-9, Feb 2001, doi: 10.1037/0021-9010.86.1.52.
- [41] K. K.-K. Wong, "Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS," *Marketing Bulletin*, vol. 24, no. 1, pp. 1-32, 2013. [Online]. Available: [http://marketing-bulletin.massey.ac.nz/v24/mb_v24_t1_wong.pdf%5Chttp://www.researchgate.net/profile/Ken_Wong10/publication/268449353_Partial_Least_Squares_Structural_Equation_Modeling_\(PLS-SEM\)_Techniques_Using_SmartPLS/links/54773b1b0cf293e2da25e3f3.pdf](http://marketing-bulletin.massey.ac.nz/v24/mb_v24_t1_wong.pdf%5Chttp://www.researchgate.net/profile/Ken_Wong10/publication/268449353_Partial_Least_Squares_Structural_Equation_Modeling_(PLS-SEM)_Techniques_Using_SmartPLS/links/54773b1b0cf293e2da25e3f3.pdf).
- [42] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, 7th, Ed. *Multivariate data analysis*. Englewood Cliffs, NJ: Prentice Hall, 2009.
- [43] R. P. Bagozzi, Y. Yi, and L. W. Phillips, "Assessing Construct Validity in Organizational Research," *Administrative Science Quarterly*, vol. 36, no. 3, pp. 421-458, 1991, doi: 10.2307/2393203.
- [44] R. P. Bagozzi and Y. Yi, "On the evaluation of structural equation models," *Journal of the Academy of Marketing Science*, vol. 16, no. 1, pp. 74-94, 1988, doi: 10.1007/BF02723327.
- [45] J. F. Hair, C. William, B. Babin, and R. Anderson, "Multivariate data analysis," ed: Upper Saddle River, NJ: Prentice Hall, 2010.
- [46] R. P. Bagozzi, "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error: A Comment," *Journal of Marketing Research*, vol. 18, no. 3, pp. 375-381, 1981, doi: 10.2307/3150979.
- [47] B. L. Bayus, "Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community," *Management science*, vol. 59, no. 1, pp. 226-244, 2013.
- [48] M. C. Bolino and A. M. Grant, "The Bright Side of Being Prosocial at Work, and the Dark Side, Too: A Review and Agenda for Research on Other-Oriented Motives, Behavior, and Impact in Organizations," *Academy of Management Annals*, vol. 10, no. 1, pp. 599-670, 2016, doi: 10.5465/19416520.2016.1153260.

Securing Student Data Privacy using Modified Snake and Ladder Cryptographic Algorithm

Dr. Kamaladevi Kunkolienker¹

Dept. of Philosophy, P.E.S' R.S.N. College of Arts and
Science, Farmagudi – Ponda – Goa - India

Vaishnavi Kamat²

Computer Engineer
India

Abstract—Transformed by the advent of the Digital Revolution, the world deals with a gold mine of data every day. Along with improvement in processing methods for the data, data security is of utmost importance. Recently, there was a noticeable surge in online learning during the pandemic. Modifying their workflow strategies, the educational institutions provided courses for students designed to suit the need of the hour. This opened up the avenue for a greater number of students to take part in the online learning. With the increase in number of students registered, there exist a substantial repository of data to deal with. Hackers have been targeting student data and using it for illegal purposes. In this research paper, an attempt has been made by modifying the classic Snake and Ladder game to perform encryption on short text student data to ensure data privacy. The Novel algorithm maintains simplicity yet produces a strong cipher text. The algorithm stands strong against the brute-force attack, cipher-only attack, etc. Decryption also uses same key as used for encryption, the key being symmetric in nature. New variable keys are generated every time the algorithm is used.

Keywords—Student data; privacy; encryption; decryption; snake and ladder; variable keys

I. INTRODUCTION

The 'Digital Revolution' is providing us abundant data which has to be dealt with caution. It was evidently noted that since there is lack of awareness of cybersecurity among the students, they tend to ignore the threat and turnout to be the most vulnerable target for the cyber-attacks. Though most of the students in the survey knew about the online risks and threats, but they lacked to understand basic knowledge about privacy [1]. In an empirical study among the students, it was explored that the students expressed concerns about their online data privacy and believed that the teachers or concerned authorities should put in efforts to safeguard it [2]. Hence, Cryptography will play a vital role in securing the data. This paper puts forth a novel algorithm for encryption and decryption for the short text. Board games as we know stimulate brain areas responsible for developing cognitive skills and memory. Here, the board games serve a different purpose altogether. The modified version of the board game will perform encryption and decryption activities for the desired short text placed on the board. The classic snake and ladder game is used with a twist to generate the cipher text.

A. Role of Technology in Education

Educational institutions with the help of technology have made great strides in improving the understanding and

interaction of the students. Students are provided with different online opportunities in terms of courses, tutorials, workshops, seminars to enhance their academic growth. With rise in these opportunities provided, there is enormous rise in the collection of student related data through registrations, personalization preferences and feedbacks.

B. Student Data

The student data usually consist of their personal details, demographic details, student preferences, evaluation reports, faculty observations etc. This data is essential for parents, faculty and policymakers to enable them to streamline their plan of action according to the student requirements.

Most of the organization store the student data and provide it to firms which help them analyze this data using technology like data mining, where several unseen underlying patterns can be highlighted in text.

II. STUDENT DATA PRIVACY

A crucial aspect that needs to be taken care of is to prevent the exploitation of these data-driven systems by the hackers. Hackers can get hold of this data through performing remote attacks on the system, by eavesdropping and can misuse the data, leading to identity theft, creation of fake accounts, modification of student details, etc. [3].

Though the student data has a very good potential and scope for bringing in improvements on various fronts, it is also a matter of concern. Digital learning captures real time information of students and along with their personal information the data captured may be used by hackers for non-educational purposes. It also discusses several legal provisions to safeguard the student data privacy [4].

In order to safeguard the sensitive student data while transferring online or stored on the administrative systems, it should be efficiently encrypted to withstand the possible attack [5].

A. Impact of Student Data Breach

Student data once hacked can open up not only their personal details to the hackers but also the details about their parents and their bank accounts which they use to pay the student fees. The exhaustive student information is sold by hackers through spam mails; a sample snapshot of the email is depicted in Fig. 1 (The snapshot was taken from authors' personal email).

Sr.No	Category / City / State	Database Quantity	Price
01	10 th old 12 th Degree Engg Working professionals		Rs30,000 Discounted Price

Fig. 1. Snapshot of Email Selling Student Data.

Students in their teens are vulnerable to the data breach activities as they are either unaware of the reality or fall prey to the attackers' malicious intents. Data like unique identification number, name, date of birth can be misused leading to identity theft. Identity theft can falsely implicate an innocent student and ruin their life in a big way. Students with weak financial background can be targeted by 'pretext calling' using this stolen data and in the need of the job or suddenly winning a lottery, the students give away more and more private and confidential information with respect to them and their family. 'Phishing' being the most common attack where the attacker posing as authorized person sends out emails to the student email address available from the hacked institution database and requests them for further information and talks about next course of action to take place from the student's side. Students, believing the sender provide the necessary information and are duped, documents are misused, photographs are inappropriately utilized for gaining profits. An instance where a student's photograph gets misused, can draw the student towards the dark world. Harassment, Stalking and human trafficking can take place at a bigger scale if the student data is not handled with utmost care [6] [7].

III. LITERATURE REVIEW

The Snake and Ladder game is directly applied to perform Steganography, which hides the data in a media like images. The algorithm uses concepts on Image Processing like Pixel Value differencing are used along with the development of the snake and ladder game on the data [8].

This research paper consists of a survey of different scenarios from different research papers, for example, storing protected data for platform like Facebook, concerns when mobile systems are used for payments and their key generation. It also surveys a concept called as Software Defined Networking – a new approach in designing, building and managing networks [9].

This paper illustrates how a game of scrabble can be utilized to generate cipher text for big length data. Several permutations and combinations of words are possible making it difficult for the attacker to guess the actual words directly hence providing good security to the plain text [10].

Security to E-Learning System is provided by Elliptic curve. cryptography algorithm and content is filtered using Decision tree algorithm. Cryptography along with Data Mining techniques proves to be a good combination. Several Data Mining Classification Techniques are compared to figure out which yields good results [11].

Another technique to ensure data privacy is the use of Digital signature. They use a different sort of balance cryptography. It is generally used for monetary exchanges, configuring dispersion, and in various circumstances where it is necessary to identify bogus or changing. Other techniques like DES and RSA Algorithm are frequently used cryptographic techniques [12].

IV. CRYPTOLOGY

Heraclitus, one of the most influential thinkers in ancient Greek Philosophy expressed hid deep philosophical insights in aphoristic and cryptic form. His writings are full of riddles which are cryptic in nature. He liked this pungent oracular style as it required a penetration of thoughts which provoked human thinking and it helped him in maintain the confidentiality of his writings [13].

Cryptology as a tool of technology is a culmination of techniques for ensuring the secrecy and/or authenticity of information. Cryptology has two main branches in the form of 'cryptography' and 'cryptanalysis' [14].

A. Components of a Cryptosystem

- Plain Text - It is the message or information that sender intends to send. Example in our case will be the student's name, date of birth, etc.
- Cipher Text – It is the result of transformation on the plain text after encryption algorithm has been applied.
- Key – It is a piece of information that will scramble the plain text as required by the algorithm.
- Encryption algorithm – It consists of sequence of steps of how and when to apply the key to the plain text to produce the cipher text.
- Decryption algorithm – It consists of sequence of steps of how and when to apply the key to the cipher text to get back the plain text [15].

Fig. 2 illustrates the various components in order of their reference.

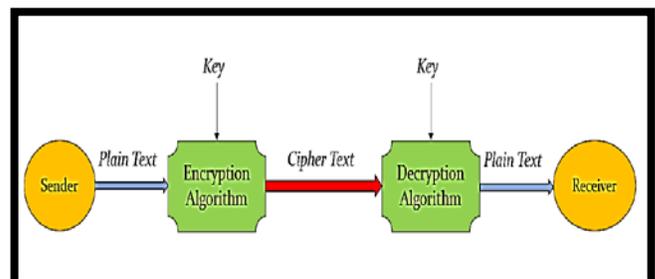


Fig. 2. Components of Cryptosystem.

B. Cryptography

The science and art of secret writing has its inscription since the beginning of human civilization. Ancient civilizations of India, Egypt, China, Japan testify the evidences of use of cryptography. The goal of cryptography is to provide secure communication over insecure channel [16].

C. Cryptanalysis

Working upon the weakness, improves the strength. This is the principal upon which cryptanalysis is based. Analyzing and understanding the cipher text, trying to decipher it without the knowledge of plain text to identify the weakness and work upon it.

D. Goals of Cryptography

- Confidentiality – A cryptosystem ensures that no one other than the sender and the receiver is able to read the input data.
- Integrity – During transmission of data, the cryptosystem should not allow any unauthorized person to modify, alter, change or delete the data.
- Authenticity – The sender and the receiver should be able to identify and validate each other [16].

V. SNAKE AND LADDER ENCRYPTION ALGORITHM

A. Concept of the Proposed Algorithm

The proposed algorithm titled ‘Snake and Ladder Encryption Algorithm (SLEA)’ gets its name from the classic game of Snake and Ladder. ‘SLEA’ is a ‘block cipher’ with variable length, symmetric keys. In the block cipher category, the input plain text is divided into fixed size of blocks and then encrypted. Block ciphers are mainly characterized by the block size and its key. Here, though a fixed size block is used for encryption, still the cells only which contain the data are encrypted, the vacant cells on the board are not considered for the encryption. This speed up the process and key variable key length shorter [17]. Symmetric keys mean same keys are used for encryption and decryption. The algorithm provides three layers of encryption. One through character set mapping and the second layer through placement of the snakes and ladders on the board and third through the snake and ladder game moves. The sender and the receiver should have the same interface setup so that pre-decided objects can be shared before the encryption and decryption process begins through the secured channel.

Sample plain text considered for the purpose of illustration is “SMILE IS JOY”. Length of the plain text including the blank space is 12 characters.

Character Set - The plain text will be first mapped to the characters set and an intermediate cipher text will be generated. There are ‘n’ number of character sets pre-decided between the sender and the receiver available to both when the algorithm interface is installed from the setup file.

Here for the illustration following character set is considered given in Table I. As for the alphabets, numbers can also be mapped to the character set.

Level 1 encryption- the given plain text ‘SMILE IS JOY’ will be mapped to intermediate cipher text as illustrated in Fig. 3.

B. Snake and Ladder Board

The game board is of the size 4x4, consisting of 16 cells. Fig. 4 shows the pattern for inserting the plain text data is always from left to right direction, whereas the traversal of data alternates in direction every time. A minimum of two snakes and two ladders adds sufficient complexity to the data. As the size of input data increases the board size also can be increased and so are the number of snakes and ladders on the board.

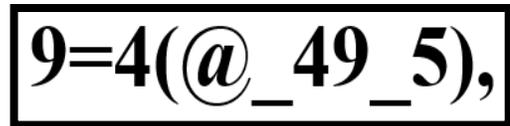


Fig. 3. Level 1 Intermediate Encrypted Text Output.

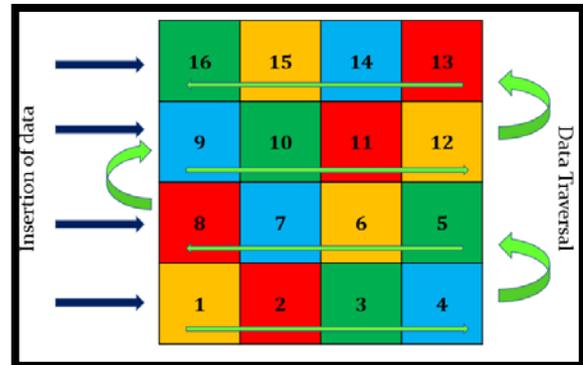


Fig. 4. Modified Board.

TABLE I. CHARACTER SET MAPPING

A	#	I	4	Q	{	Y	,	5	q	“	%]	n	\$	'	+	b
B	1	J	5	R	7	Z	&	6	z	”	<	{]	^	'	-	c
C	2	K	}	S	9		—	7	j	:	>	}	[~	\$	/	e
D	?	L	(T	+	0	a	8	k	'	^	@	o	`	n	*	d
E	@	M	=	U	/	1	s	9	m	'	f	&	p	=	.		
F	;	N	!	V	*	2	y	.	~	(i	%	r		y		
G	3	O)	W	:	3	h	,	“)	g	#	u	<	x		
H	8	P	6	X	-	4	t	;	”	[l	?	v	>	w		

C. Snake and Ladder Encryption Algorithm

Step 1: Create a 4x4 board and select any one arrangement of the snakes and ladders from the predetermined set.

Step 2: Place the level 1 encrypted text onto the 4X4 board.

Step 3: Note the last data entry cell on the board. (This step ensures that the vacant cells are not traversed, hence less time taken for encryption process.).

Step 4: Point the ‘marker’ to the first cell on the board and roll the dice till all of the entries on the board are visited at least once. Simultaneously, frame the data from the visited cells into a sequence, which is our cipher text to be transmitted or stored.

Step 4.1: Travel the cells on the boards depending on the count on the dice.

Step 4.2: At the end of the traversal for that particular move, if a ladder is encountered- climb up and record the data in the cell in to the cipher text sequence. Similarly, if a snake is encountered- drop down and again record the data in the cell in to the cipher text sequence.

Step 4.3: At any point of time if looping takes place, i.e. dice and cell on the board remain same, do not record the data and roll the dice till the number on the face of dice is different. This resolves looping.

Step 5: If the current cell visited on the board is the last cell, check if there exist any unvisited cells on the board. (Unvisited entries may be due to sudden climbing using ladder or dropping down through the snake). Repeat the process till all cells on the board have been visited at least once.

D. Encryption Algorithm Illustrated

Fig. 5 shows the mapped text on the game board with snakes and ladders already placed on the game board. Table II provides stepwise execution of the encryption algorithm.

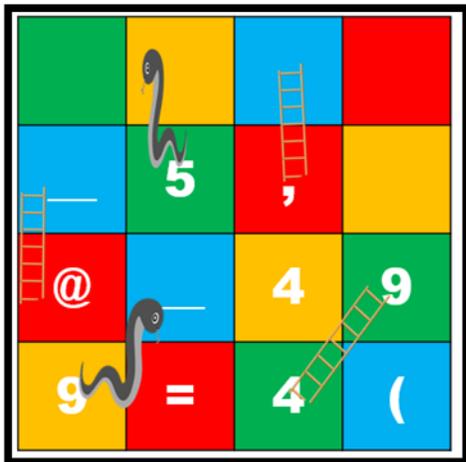


Fig. 5. Mapped Text on the Board.

TABLE II. STEPWISE ILLUSTRATION

Dice Roll Count	Number on the Dice Face	Ladder / Snake at the end of traversal	Cipher Text Generated
1	3	YES	9=4
		Ladder	9=49
2	1	NO	9=494
3	1	YES	9=494_
		Snake	9=494__9
4	6	YES	9=494__9
		SNAKE - LOOPING	No change in the cipher text
5	5	NO	9=494__9=4(94
6	4	NO	9=494__9=4(94_@_5
7	3	NO	9=494__9=4(94_@_5,
		End of Characters on the Board	Extra places not counted
End of Characters on the Board - Check if all entries are visited.			
All entries are visited on the board			

Final Cipher Text Generated is shown in Fig. 6, which consists of 17 characters.

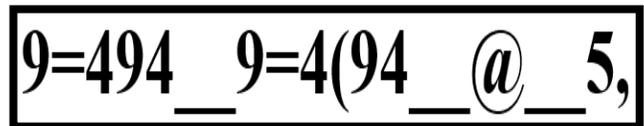


Fig. 6. Final Cipher Text.

Original Plain Text: “SMILE IS JOY”

Length of original plain text: 12

Generated Cipher Text: 9=494__9=4(94_@_5,

Length of Generated cipher text: 17

E. Snake and Ladder Key Generation during Encryption

For the decryption of the cipher text at the receiver side, we need a Key that will help us get back the plain text. As the cipher text is generated, simultaneously the key is also generated. The key is the numbers on the face of dice plus the pattern number for the board arrangement for snake and ladder plus the character set pattern used.

F. Snake and Ladder Decryption Algorithm

The receiver side will use the key transferred through the secured channel and decode the cipher text. Table III provides stepwise decryption.

TABLE III. DECRYPTION USING KEY

Key	Cipher Text	Board Arrangement
3	9=494_9=4(94_@_5,	Figure 7
3 1	9=494_9=4(94_@_5,	Figure 8
3 1 1	9=494_9=4(94_@_5,	Figure 9
3 1 1 5	9=494_9=4(94_@_5,	Figure 10
3 1 1 5 4	9=494_9=4(94_@_5,	Figure 11
3 1 1 5 4 3	9=494_9=4(94_@_5,	Figure 12

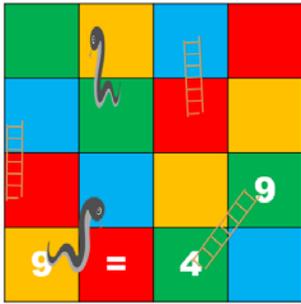


Fig. 7. Key 3.

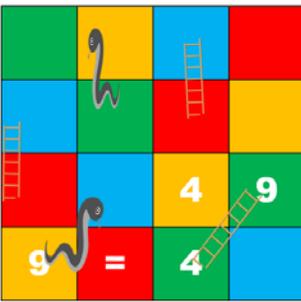


Fig. 8. Key 1.

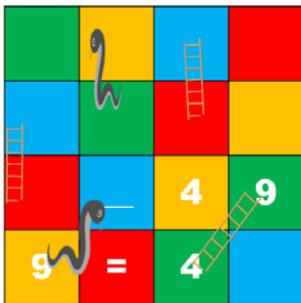


Fig. 9. Key 1.

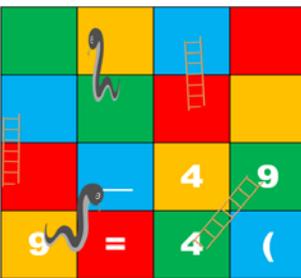


Fig. 10. Key 5.

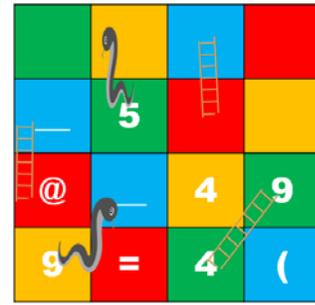


Fig. 11. Key 4.

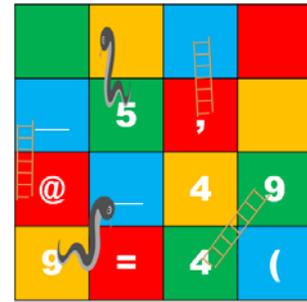


Fig. 12. Key 3.

VI. SECURITY ASPECT

In this section, we will analyze the strength of the snake and ladder algorithm against several types of attacks. Snake and Ladder Cryptographic algorithm provides three levels of security.

Level 1: First, in the form of character mapping set. The set consist of 68 characters mapped uniquely to each other. For every encryption, the character mapping will be changed. If any intruder tries to guess the mapping it will be difficult to select from 24800355424368305996009904185691715810473992013553 6767237171073801822144571218329600000000000000 options, that is 68! (68 factorial). This step considerably increases the security in a simple manner.

Level 2: Second, in terms of the placement of the snakes and ladders on the board. The board is selected of 4x4 as it can hold average length of information like student name, roll number, subject name, etc. Depending upon the data length per student, the size of the board can be increased to 5x5 or 6x6 accordingly. For every encryption, the number of snakes and ladders can be decided and a different arrangement of snakes and ladders can be selected from the predefined pattern.

For the current illustration, a board of 4x4 consists of 2 snakes and 2 ladders, sufficient for adding complexity to the encryption process. When the interface will be setup between the sender and the receiver, these predefined patterns and arrangements will be known to both. So, during encryption, along with the key the unique pattern identification number is to be attached before transfer of the cipher text.

Understanding the complexity provided by the placement of snakes and ladders – consider, the 4x4 board with two snakes and two ladders. Initially, total available cells on the board for the placement were 16. Placing the 1st snake -16

options available, placing 2nd snake – 15 options available, placing 1st ladder 14 options available, placing 2nd ladder – 13 options available. It is difficult to gauge the positions by simple guessing.

Level 3: Third, the algorithm itself. The manner in which the cipher text will be generated by rolling the dice will be different every time. And occurrences of the snakes and ladders in the route will change the course of cipher text generated.

Now, we will slide through several types of attacks possible during encryption and how the algorithm provides resistance to them.

Ciphertext Only Attacks – In this kind of attack, the encrypted message is intercepted by the attacker. This attack will be successful only if the corresponding plain text is retrieved from the encrypted message.

The Snake and Ladder algorithm has a huge initial character base to transform the plain text into unmeaningful message. The cipher text characters are further jumbled by the movements in the game. So, it is impossible for the attacker to know the plain text within a finite time without the authentic keys.

Known Plaintext Attack – The attacker knows the partial plaintext for the generated ciphertext. So, the attacker's task is to study the remaining cipher text, derive the key if possible and get the plain text.

The Snake and Ladder algorithm holds an advantage with respect this attack. Since, the plain text is neither used as direct part of cipher text nor as part of key formation, and even if the attacker knows that it is the student data that the algorithm is dealing with, it cannot conclude anything directly only from the cipher text.

Man in Middle Attack – Mostly, this attack can target the public key cryptosystems. Here, key exchange has to take place before communication begins.

In the Snake and Ladder algorithm, since key is simultaneously generated with the cipher text and also transmitted with it, this attack does not affect the encryption algorithm.

Chosen Plaintext Attack – The attacker has the ability to encrypt plain text of his choice and generate corresponding cipher text in the given system. By performing this act, the attacker tries to study the relation between plain text to cipher text generation.

In the Snake and Ladder algorithm, at level one itself the attacker will have to go through and execute 68! possibilities.

Brute Force Attack – Try several combinations of keys, to decrypt the cipher text.

In the Snake and Ladder algorithm, though the dice has only 6 faces, we don't know how many times it will be rolled, because the dice rolling depends on the length of the text to be encrypted when it is placed on the board. This is the advantage of the variable nature of the key, which makes it difficult to

guess what it is. And if one tries to do so, it might require exponential time [18] [19].

Running Time Analysis – the problem statement is - with how many minimum dice rolling can the text on board be traversed at least once. There are three important factors affecting this, one is the number of filled places on the board, second variable number of snakes and ladders do affect the traversal and the number appearing on the dice face.

For illustration, a 4X4 board was considered, means total number of blocks to be traversed 16. So, running time would be close to $O(16 * C)$, where C is a constant and dice is assumed to have 6 faces. 16 because the block has to be visited at least once. C includes the time including waiting for the dice to roll again and repetitive traversal of the blocks, and climbing up and down the ladders and snakes, respectively. General running time is $O(N * D * C)$ where N is the number of blocks on the board, D is the number of faces on the dice, if dice of a greater number of faces used and C remains the constant.

VII. ADVANTAGES AND DISADVANTAGES OF THE SNAKE AND LADDER ALGORITHM

A. Advantages

- The nature of the game makes the moves during the encryption unpredictable, which strengthens the cipher text.
- Time required for execution is comparatively less as the encryption is not done on entire block but only the part of data present in it.
- Key generation is simultaneous to encryption process.
- Rolling the dice makes the key Variable in length and the Variable key makes it difficult for the attackers to predict the nature of the key.
- Though, there are large number of students, each student's data in a school / college repository is limited to certain details. Hence, encrypting this short length data with this flexible size algorithm avoids unnecessary space and time complexities.
- Considering the best or average time cases (number on the dice face from 6 to 4, every time it rolled) the algorithm can be considerably fast or average during encryption.

B. Disadvantages

- As the board size increases, the complexity for the algorithm's implementation increases.
- Considering the worst case (number on the dice face from 3 to 1, every time it rolled) the algorithm can be a bit slow during encryption.

VIII. CONCLUSION

The 'randomness' of the 'dice', 'movements' on the 'board' add the twist to the generation of the cipher text just as required. Since, student data can be categorized into distinct fields like their first name, last name, date of birth etc. this algorithm on the board of 4x4 is most suitable to encrypt the

data part by part and store it in the same manner. Though there exist several cryptographic algorithms that provide cipher text, the Snake and Ladder algorithm consumes minimum time and generates the cipher text in length close to the length of plain text and key length is maintained less than cipher text length, hence reducing overall processing time.

Future research could examine improving the key transmission from sender to the receiver. Also, improvements in hardware can considerably increase the efficiency of the algorithm.

As mentioned before, the student data breach can have devastating psychological effects on the students. The pillars of the nation have to be necessarily protected against these hidden cyber-attacks. Ensuring the confidentiality of the student data while collecting, proper cryptographic techniques to be used while storing or transmitting, and enforcing strict laws modified according to the emerging cyber-attacks are some of the basic steps that need to be enforced by the concerned authorities.

REFERENCES

- [1] Gabra A. A., Sirat M. B., Hajar S., Dauda I. B. , Cyber Security Awareness Among University Students: A Case Study, International Journal of Advance Science and Technology, Vol. 29 No. 10S, pp. 767-776, 2020.
- [2] Lorenz Birgy, Sousa Sonia, Tomberg Vladimir , Privacy Awareness of Students and Its Impact on On-line Learning Participation –A Case Study. 1st Open and Social Technologies for Networked Learning (OST), Tallinn, Estonia. pp.189-192, 2012.
- [3] Townsend A. (2021), 3 Reasons Higher Education is a Cyberattack Favorite <https://www.onelogin.com/blog/3-reasons-higher-ed-hacked>.
- [4] Stahl W., Karger J., Student Data Privacy, Digital Learning, and Special Education: Challenges at the Intersection of Policy and Practice, Journal of Special Education Leadership, Vol- 29 No. (2), 2016, pp. 79-88.
- [5] Carr R. (2018), The Rise of Education Data Breaches <https://www.zettaset.com/blog/education-data-breaches/>.
- [6] Bandler J., Merzon A., Cybercrime Investigations- A Comprehensive Resource for Everyone, Published by CRC Press, 2020, pp. Chapter 2.
- [7] Clough J., Principles of Cyber Crime, Published by Cambridge University Press, 2015, pp. 417- 419 and 209-212.
- [8] Seth J. R., Snake and Ladder based Algorithm for Steganographic Application of Specific Streamline Bits on Prime Gap Method, International Journal of Computer Applications, Volume 94 – No 3, 2014.
- [9] Lokesh V., Jayaraman S., Guruprasad H. S., A Survey on Network Security and Cryptography, International Journal of Advance Research In Science And Engineering, Vol. No.3, Issue No.10, 2014, pp. 56-54.
- [10] Kamat V.K., Scrabble-O-Graphy: An Encryption Technique for Security Enhancement. In: Sa P., Bakshi S., Hatzilygeroudis I., Sahoo M. (eds), Recent Findings in Intelligent Computing Techniques. Advances in Intelligent Systems and Computing, Vol 707. Springer, Singapore, pp. 2019, pp. 115-124.
- [11] Patil Vijaya, Vedpathak Aditi, Shinde Pratiksha, Vatandar Vishakha, Janrao Surekha, E-learning system using cryptography and data mining techniques, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 01, e-ISSN: 2395-0056, Jan-2018.
- [12] Mangore Anirudh K. , M. Roberts Masillamani, Efficient Cryptographic Encryption Techniques for Data Privacy Preservation, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8, Issue-7S, May 2019.
- [13] Kunkolienker K., Sri Aurobindo's Views on Heraclitus' Philosophy: A Synthesis, Proceedings published by WASET, EISSN: 2010-3778, 2017.
- [14] Stallings, W., Cryptography and Network Security (4th Edition), Published by Pearson Education, 2006, pp. 28 – 35.
- [15] Buchmann, J., Introduction to Cryptography, Published by Springer New York, 2013, 69– 71.
- [16] Padhye, S., Sahu, R., Saraswat, V., Introduction to Cryptography, Published by CRC Press, 2018, pp. Chapter 1.
- [17] Aumasson Jean – Philippe, Serious Cryptography- A Practical Introduction to Modern Encryption, Published by No Starch Press, 2017, pp. 53-55.
- [18] Henk, C.A., van Tilborg, Encyclopaedia of Cryptography and Security, Published by Springer, 2011, pp. 153-156.
- [19] Swenson, C., Modern Cryptanalysis – Techniques for Advanced Code Breaking, Published by Wiley, 2012, pp. Chapter 5.

A Method to Accommodate Backward Compatibility on the Learning Application-based Transliteration to the Balinese Script

Gede Indrawan¹, I Gede Nurhayata³, Sariyasa⁴

Department of Computer Science
Universitas Pendidikan Ganesha (Undiksha)
Singaraja, Indonesia

I Ketut Paramarta²

Department of Balinese Language Education
Universitas Pendidikan Ganesha (Undiksha)
Singaraja, Indonesia

Abstract—This research proposed a method to accommodate backward compatibility on the learning application-based transliteration to the Balinese Script. The objective is to accommodate the standard transliteration rules from the Balinese Language, Script, and Literature Advisory Agency. It is considered as the main contribution since there has not been a workaround in this research area. This multi-discipline collaboration work is one of the efforts to preserve digitally the endangered Balinese local language knowledge in Indonesia. The proposed method covered two aspects, i.e. (1) Its backward compatibility allows for interoperability at a certain level with the older transliteration rules; and (2) Breaking backward compatibility at a certain level is unavoidable since, for the same aspect, there is a contradictory treatment between the standard rule and the old one. This study was conducted on the developed web-based transliteration learning application, BaliScript, where its Latin text input will be converted into the Balinese Script output using the dedicated Balinese Unicode font. Through the experiment, the proposed method gave the expected transliteration results on the accommodation of backward compatibility.

Keywords—Backward compatibility; Balinese Script; learning application; transliteration

I. INTRODUCTION

As one of the diversity of local language knowledge in Indonesia, the endangered Balinese Script transliteration knowledge [1]–[3] raises concerns for the preservation. The Bali Government has already conducted the preservation efforts through the Bali Governor Regulation [4], [5] and strengthen them with the Bali Governor Circular Letter [6]. These efforts make the Balinese Language, including its Balinese Script transliteration knowledge, running as a mandatory local subject from elementary school to senior high school in Bali Province.

Multiple approaches other than the governmental approach should strengthen the preservation effort and should have a greater impact. This research joined the effort through the technological approach by multi-discipline collaboration between Computer Science and Language discipline. It proposed a method to accommodate backward compatibility on the learning application-based transliteration to the Balinese Script. This work has never been conducted yet and applied to the previous works that were still based on the older

transliteration rules (for short, the older rules) from The Balinese Alphabet document¹. It exposes the backward compatibility method to accommodate the standard transliteration rules (for short, the standard rules) from the Balinese Language, Script, and Literature Advisory Agency [7]. This Bali Province government agency [4] carries out guidance and formulates programs for the maintenance, study, development, and preservation of the Balinese Language, Script, and Literature.

This study was conducted on the developed web-based transliteration learning application, BaliScript, for further ubiquitous Balinese Language learning since the proposed method reusable for the mobile application [8], [9]. It also advances the previous work by (1) accommodating special words [1], [10] through a certain table structure in the database rather than hard-coding them in the application code; (2) making use of the more developed and the less bug of Noto Serif Balinese (NSB) font^{2,3} [11] to represent the Balinese Script rather than the Noto Sans Balinese font⁴. The NSB font is a dedicated Balinese Unicode font which makes it recognized on the computer system including mobile devices and makes the proposed method reusable on the mobile application; and (3) improving the learning experience on the application, that uses this method, through the addition of the Indonesian and English translation for the transliterated word (see the next Fig. 3). Overall, all of those advances are considered as the contribution of this work.

This paper is organized into several sections, i.e. Section I (Introduction) states the problem background related to the transliteration to the Balinese Script; Section II (Related Works) describes the related works in the area of the transliteration to the Balinese Script and its backward compatibility aspect; Section III (Research Method) exposes the supporting algorithm, the implementation, and the testing

¹ The Balinese Alphabet, <http://www.babadbali.com/aksarabali/alphabet.htm> (Retrieved June 16, 2021)

² Balinese Unicode Table, <http://unicode.org/charts/PDF/U1B00.pdf> (Retrieved June 16, 2021)

³ Google Noto Serif Balinese, <https://github.com/googlefonts/noto-fonts/blob/master/unhinted/ttf/NotoSerifBalinese/NotoSerifBalinese-Regular.ttf> (Retrieved June 16, 2021)

⁴ Google Noto Fonts, <https://www.google.com/get/noto/#sans-bali> (Retrieved June 16, 2021)

of the proposed method; Section IV (Result and Analysis) covers the analysis of the testing result; and finally, Section V (Conclusion) consists of important conclusion and future work points.

II. RELATED WORKS

Several related works on Latin-to-Balinese Script transliteration were conducted on the previous works [10], [12]–[20]. All of those were still based on the older rules from The Balinese Alphabet document, except [20]. Displaying Balinese Script output on those previous research was done by non-dedicated Balinese Unicode fonts (i.e. Bali Simbar⁵ and Bali Simbar Dwijendra [21]) and dedicated Balinese Unicode font² [11] (i.e. Noto Sans Balinese and Noto Serif Balinese). The Bali Simbar (BS) font was utilized in [12] and gave a relatively good accuracy result on testing cases from The Balinese Alphabet document. It was also utilized in the developed robotic system that writes the Balinese Script from the Latin text input [13], and on the exploration of the line-break handling during the transliteration [14]. The Bali Simbar Dwijendra (BSD) font, as the improvement of the BS font, was utilized in [15] with additional testing cases from the Balinese Script dictionary [7] to the same testing cases on [12]. It was also utilized in the exploration of the mathematical expression transliteration [16]. Ten transliteration lessons were also learned by using this font on the other testing data [17]. The Noto Sans Balinese font was utilized in [10] with the same testing cases in [12] and gave a relatively good accuracy result. It was also utilized in the developed robotic system that writes Balinese Script from the Latin text input [18]. Extensive accuracy analysis on the developed algorithm [10] was done in [19] for future improvement. the Noto Serif Balinese font was utilized in [20] for the unavoidable affixed words that need to be transliterated.

The other side of transliteration related to the Balinese Script-to-Latin transliteration that utilized the GNU Optical Character Recognition (OCR), i.e. Ocrad⁶ [22]. This research was limited only to the basic syllable recognition (see The Balinese Alphabet document) from the Balinese Script image that was based on the glyph shape of the Bali Simbar font. For advancing functionality and mobile adoption for ubiquitous learning, the utilization of the Tesseract⁷ OCR was conducted that needs several future improvements [23].

III. RESEARCH METHOD

The proposed method to accommodate backward compatibility on the transliteration to the Balinese Script covers two aspects related to the older transliteration rules from The Balinese Alphabet document. Those two aspects, i.e. (1) Backward compatibility allows for interoperability at a certain level with the older rules; and (2) Breaking backward compatibility at a certain level is unavoidable since, for the

same aspect, there is a contradictory treatment between the standard rule and the old one.

This section describes (1) the supporting algorithm of the proposed method; (2) the implementation on the BaliScript, which is the web-based transliteration learning application; and (3) the testing by using the updated testing cases of The Balinese Alphabet document to comply with the standard transliteration rules from the Balinese Language, Script, and Literature Advisory Agency [4], [7].

A. The Algorithm

The proposed method involves the NSB font with its dedicated Balinese Unicode Table [20]. The algorithm to accommodate backward compatibility on the transliteration to the Balinese Script covers two aspects, as described previously. The first aspect involves transliteration of the letter set MBC (Maintaining Backward Compatibility), i.e. the vowel “*ē*” (U+011B) with sound [ə] [24], “*ō*” (U+00F6) with the long sound of the vowel “*ē*”, the consonant na rambat “*ṅ*” (U+0146) with sound [ŋa], sa sapa “*ś*” (U+015B) with sound [ʃa], sa saga “*ṣ*” (U+015F) with sound [ea], ta latik “*ṭ*” (U+0163) with sound [ta], or its uppercase letter “*Ē*” (U+011A), “*Ö*” (U+00D6), the consonant “*Ṇ*” (U+0145), “*Ś*” (U+015A), “*Ṣ*” (U+015E), “*Ṭ*” (U+0162). The second aspect involves transliteration of the letter set BBC (Breaking Backward Compatibility), i.e. the vowel “*e*” (U+0065) or its uppercase letter “*E*” (U+0045) that has sound [e]. Noted that the uppercase letters were not the concern, since each of them has the same transliteration result as its counterpart lowercase letter.

Those two aspects should be handled by the proposed method. Fig. 1 shows the flowchart of the algorithm and uses regular expression [25], [26] on the implementation.

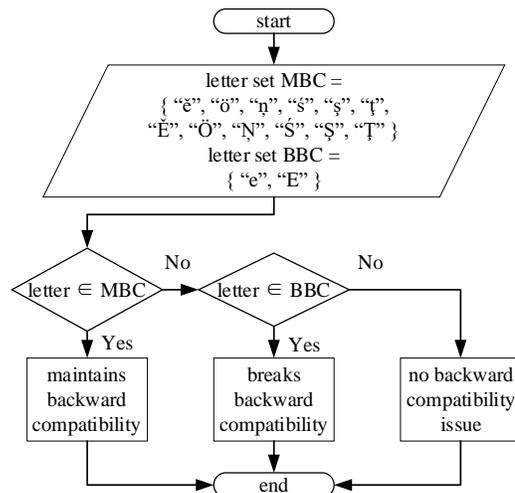


Fig. 1. The Flowchart of the Algorithm.

⁵ Bali Simbar, <http://www.babadbali.com/aksarabali/balisimbar.htm> (Retrieved June 16, 2021)

⁶ The GNU Ocrad OCR, <https://www.gnu.org/software/ocrad/> (Retrieved June 16, 2021)

⁷ Tesseract OCR, <https://github.com/tesseract-ocr/> (Retrieved June 16, 2021)

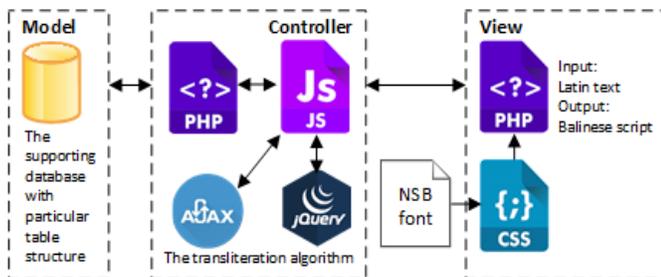
B. The Implementation

Fig. 2 (a) shows the Model-View-Controller (MVC) architecture [27]–[29] of the web-based transliteration learning application, BaliScript, that was used by the proposed method. The supporting database’s table (Fig. 2 b) consists of records from the Balinese Script dictionary [7]. Fig. 3 shows the Indonesian and English translation of the example transliterated word for improving the learning experience on the application. As described previously, this feature is one of several advances as the contribution of this work. The BaliScript was constructed by Apache web server, MySQL database server, and PHP code combined with JavaScript code. This application was also used for the exploration of scriptio continua management in the previous work [30].

Fig. 3 (a) shows the View of the MVC, i.e. (1) the input view that uses the Select box⁸; and (2) output view that displays the transliteration result and other results from the closest similar words in the database where the similarity calculation is based on the Levenshtein distance [31], [32]. Fig. 3 (b) shows the transliteration output from the example homonym word [33], [34] at the similarity list by using AJAX-based switching (clicking on the word “USE” related to the certain word).

C. The Testing

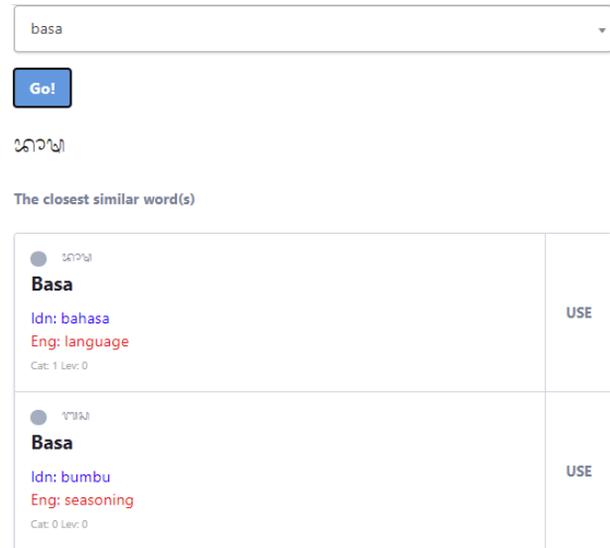
The testing of the proposed method was conducted on the BaliScript, which was run on the Intel Core i7-4600U CPU @2.09GHz platform with 8 GB RAM and Windows 8 64-bit Operating System.



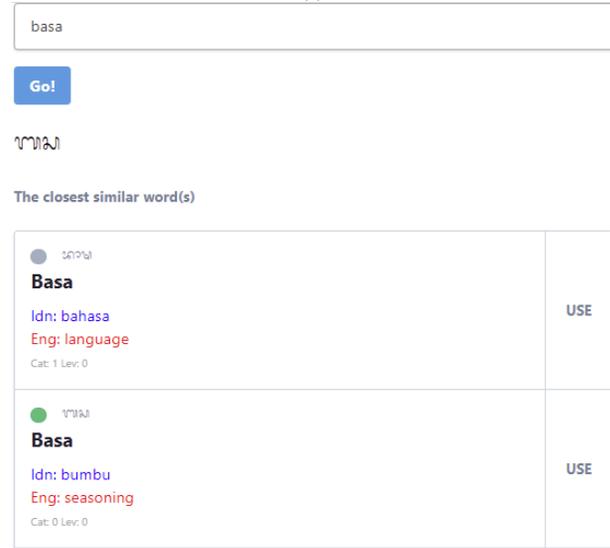
id	word	sword	special	idn	eng	...
...
1952	basa	bhāḷṣa	1	bahasa	language	...
1953	basa	basa	0	bumbu	seasoning	...
...

(b)

Fig. 2. The BaliScript Implementation: (a) MVC Architecture; (b) Supporting Table in the Database



(a)



(b)

Fig. 3. The View of the BaliScript with Transliteration and Translation Result at the Same Time: (a) Output with the Closest Similar Words; (b) Output from the Example Homonym Word at the Similarity List.

IV. RESULT AND ANALYSIS

Table I shows the testing cases consist of sections of interest (the marked sections) related to the result of backward compatibility (see Fig. 4). Noted that the testing used the updated testing cases that comply with the standard transliteration rules from the Balinese Language, Script, and Literature Advisory Agency [7] rather than the original testing cases [10] that refer to The Balinese Alphabet document.

⁸ Select2 box, <https://select2.org> (Retrieved June 16, 2021)

TABLE I. TESTING TRANSLITERATION CASES

No.	Case ^a	Case ^b	Remarks ^c
1	ha na ca ra ka da ta sa wa la ma ga ba nga pa ja ya nya Bakta <u>K</u> ala <u>P</u> aksa Raka Cakra <u>W</u> alaka Krama	ha na ca ra ka da ta sa wa la ma ga ba nga pa ja ya nya Bakta <u>K</u> ala <u>P</u> aksa Raka Cakra <u>W</u> alaka Krama	Basic syllables and examples. Bring Time Force “Bigger brother” Disc “A non-priest” Member
2	<u>K</u> ā <u>d</u> ep <u>J</u> ēro Siya Kayu <u>S</u> ela Angklung Daitya Patūt <u>D</u> wī	<u>K</u> ā <u>d</u> ep <u>J</u> ēro Siya Kayu <u>S</u> ela Angklung Daitya Patūt <u>D</u> wī	Vowel signs examples. Sold House Nine Wood Yam “Musical instrument” Giant “Should be” Two
3	a ā i ī u ū e ai o au Ak <u>ṣ</u> ara <u>I</u> swara Upacāra Eka Airlangga Ong OM	a ā i ī u ū e ai o au Ak <u>ṣ</u> ara <u>I</u> swara Upacāra Eka Airlangga Ong OM	Independent vowels and examples. Alphabet “God’s name” Ceremony One “A Javanese King” “One holy letter” “Symbol of God”
4	rě rō lě lö Talěr Kěrěng	rě rō lě lö Talěr Kěrěng	Illegal combination of syllable - vowel signs and examples. Also “Eat a lot”
5	Pak Raman Pakraman Baglug Rubag lugu Briag	Pak Raman Pakraman Baglug Rubag lugu Briag	Semi vowels examples. Mr. Raman Membership Stupid “Naive Rubag” Laughter
6	<u>n</u> a dha tha <u>ṭ</u> a <u>ṣ</u> a <u>ṣ</u> a gha bha pha Ga <u>n</u> itri Garudha <u>P</u> artha Jaṭayu Bhi <u>ṣ</u> ama <u>Ṣ</u> iwa Laghu	<u>n</u> a dha tha <u>ṭ</u> a <u>ṣ</u> a <u>ṣ</u> a gha bha pha Ga <u>n</u> itri Garudha <u>P</u> artha Jaṭayu Bhi <u>ṣ</u> ama <u>Ṣ</u> iwa Laghu	Ak <u>ṣ</u> ara swalalita and examples. Chain “Big eagle” “Arjuna’s alias” “A bird in Ramayana” Decree “God’s name” “Low tone in singing”
7	C <u>ṅ</u> g <u>ṅ</u> g Bangkung Manah Kar <u>n</u> a Kap <u>l</u>	C <u>ṅ</u> g <u>ṅ</u> g Bangkung Manah Kar <u>n</u> a Kap <u>l</u>	Sound killers examples. “Musical instrument” Pig Mind Ear Ship
8	Mang Siddham	Mang Siddham	Miscellaneous signs examples. “Holy letter” Perfect
9	Om Swastiastu Om <u>Ṣ</u> anti, <u>Ṣ</u> anti, <u>Ṣ</u> anti, Om	Om Swastiastu Om <u>Ṣ</u> anti, <u>Ṣ</u> anti, <u>Ṣ</u> anti, Om	Holy symbol Ongkara examples. “May God blesses you” “May peace be everywhere”
10	cha kha	cha kha	Miscellaneous syllables.
11	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	The digits.
12	, . < .0. > >> : "	, . < .0. > >> : "	Punctuations.
13	i u <u>ḡ</u> o <u>ḡ</u> ö pu phu Sekala seka <u>ḡ</u> Samping Suk <u>ṣ</u> ma K <u>ṣ</u> atria Strī Smerti U <u>ṭ</u> ama Dharma Tamblang	i u <u>ē</u> o <u>ē</u> ö pu phu Sekala seka <u>ḡ</u> Samping Suk <u>ṣ</u> ma K <u>ṣ</u> atria Strī Smerti U <u>ṭ</u> ama Dharma Tamblang	Some variation of usages. Combination of independence vowel a kara with vowel signs Pairing of pa kapal with suku or suku ilut Romanization of the inherent sound: Real real Usage of pangangge ak <u>ṣ</u> ara: Side “Thank you” Warrior Wife “Books of Vedha” Primary Religion “A village’s name”
14	hā nā cā rā kā dā tā sā wā lā mā gā pā yā <u>n</u> ā dhā thā <u>ṭ</u> ā <u>ṣ</u> ā <u>ṣ</u> ā ghā bhā	hā nā cā rā kā dā tā sā wā lā mā gā pā yā <u>n</u> ā dhā thā <u>ṭ</u> ā <u>ṣ</u> ā <u>ṣ</u> ā ghā bhā	Ligatures.
15	Bank <u>P</u> embangunan Daerah Bali <u>B</u> e <u>P</u> e <u>D</u> e Bali Ba <u>P</u> e Da Bali Ba Pa Da Bali	Bank <u>P</u> embangunan Daerah Bali B <u>ē</u> , P <u>ē</u> , D <u>ē</u> , Bali Ba, <u>P</u> e, Da, Bali Ba, Pa, Da, Bali	Abbreviations examples. Regional Development Bank Bali Be Pe De Bali Ba Pe Da Bali Ba Pa Da Bali
16	A <u>ḡ</u> eh ak <u>ṣ</u> aran <u>ḡ</u> , 47, lui <i>r</i> ipun: ak <u>ṣ</u> ara suara, 14, ak <u>ṣ</u> ara wianjana, 33, ak <u>ṣ</u> ara suara punika tal <u>ḡ</u> er dados pangang <u>ḡ</u> e suara, tur mad <u>ḡ</u> ew <u>ḡ</u> e suara kakalih, kawā <u>ṣ</u> tanin: suara hr <u>ḡ</u> swa miwah d <u>ḡ</u> rgha.	A <u>ḡ</u> eh ak <u>ṣ</u> aran <u>ē</u> , 47, lui <i>r</i> ipun: ak <u>ṣ</u> ara suara, 14, ak <u>ṣ</u> ara wianjana, 33, ak <u>ṣ</u> ara suara punika tal <u>ḡ</u> er dados pangang <u>ḡ</u> e suara, tur mad <u>ḡ</u> ew <u>ḡ</u> e suara kakalih, kawā <u>ṣ</u> tanin: suara hr <u>ḡ</u> swa miwah d <u>ḡ</u> rgha.	Word boundaries and line break rules. Many of those letters, 47, i.e.: vowels, 14, consonants, 33, those vowels also become vowel signs, and have two sounds, each is called: sound hr <u>ḡ</u> swa and d <u>ḡ</u> rgha.

^a. The original testing cases

^b. The updated testing cases that comply with the standard transliteration rules

^c. The Balinese Alphabet document

- [4] Bali Government, "Bali Government Regulation No. 3 on Balinese Language, Script, and Literature [Peraturan Pemerintah Bali No. 3 tentang Bahasa, Aksara, dan Sastra Bali]," 1992. https://bphn.go.id/data/documents/perda_3_1992.pdf (Retrieved June 16, 2021).
- [5] Bali Government, "Bali Governor Regulation No. 80 on Protection and Usage of Balinese Language, Script, and Literature, also Organizing Balinese Language Month [Peraturan Gubernur Bali No. 80 tentang Pelindungan dan Penggunaan Bahasa, Aksara, dan Sastra Bali serta Penyelenggaraan Bulan Bahasa Bali]," 2018. <https://jdih.baliprov.go.id/produk-hukum/peraturan/abstrak/24665> (Retrieved June 16, 2021).
- [6] Bali Government, "Bali Governor Circular Letter No. 3172 Year 2019 about The Usage of Balinese Traditional Clothing and Balinese Script [Surat Edaran Gubernur Bali No. 3172 Tahun 2019 tentang Penggunaan Busana Adat Bali dan Aksara Bali]," 2019. <https://jdih.baliprov.go.id/produk-hukum/peraturan/abstrak/24741> (Retrieved June 16, 2021).
- [7] I. G. K. Anom et al., *Balinese - Indonesian Dictionary with its Latin and Balinese Script [Kamus Bali - Indonesia Beraksara Latin dan Bali]*. Denpasar: Bali Province, 2009.
- [8] G. J. Hwang, C. C. Tsai, and S. J. H. Yang, "Criteria, strategies and research issues of context-aware ubiquitous learning," *J. Educ. Technol. Soc.*, vol. 11, no. 2, 2008.
- [9] H. Ogata, Y. Matsuka, M. M. El Bishouty, and Y. Yano, "LORAMS: linking physical objects and videos for capturing and sharing learning experiences towards ubiquitous learning," *Int. J. Mob. Learn. Organ.*, vol. 3, no. 4, pp. 337, 2009.
- [10] G. Indrawan, I. K. Paramarta, and K. Agustini, "A new method of Latin-to-Balinese script transliteration based on noto sans balinese font and dictionary data structure," in *The 2nd International Conference on Software Engineering and Information Management (ICSIM)*, 2019.
- [11] The Unicode Consortium, *The Unicode Standard Version 13.0 – Core Specification*. The Unicode Consortium, 2020.
- [12] G. Indrawan, Sariyasa, and I. K. Paramarta, "A new method of Latin-to-Balinese script transliteration based on bali simbar font," in *The 4th International Conference on Informatics and Computing (ICIC)*, 2019.
- [13] P. N. Crisnapati et al., "Pasang aksara bot: a Balinese script writing robot using finite state automata transliteration method," *J. Phys. Conf. Ser.*, vol. 1175, no. 1, pp. 012108, 2019.
- [14] G. Indrawan, K. Setemen, W. Sutaya, and I. K. Paramarta, "Handling of line breaking on Latin-to-Balinese script transliteration web application as part of Balinese language ubiquitous learning," in *The 6th International Conference on Science in Information Technology (ICSITech)*, 2020.
- [15] G. Indrawan, I. P. E. Swastika, Sariyasa, and I. K. Paramarta, "An improved algorithm and accuracy analysis testing cases of Latin-to-Balinese script transliteration method based on bali simbar dwijendra font," *Test Eng. Manag.*, vol. 83, pp. 7676–7683, 2020.
- [16] G. Indrawan, G. R. Dantes, K. Y. E. Aryanto, and I. K. Paramarta, "Handling of mathematical expression on Latin-to-Balinese script transliteration method on mobile computing," in *The 5th International Conference on Informatics and Computing (ICIC)*, 2020.
- [17] G. Indrawan, I. G. A. Gunadi, M. S. Gitakarma, and I. K. Paramarta, "Latin to Balinese script transliteration: lessons learned from computer-based implementation," in *The 4th International Conference on Software Engineering and Information Management (ICSIM)*, in press.
- [18] G. Indrawan, N. N. H. Puspita, I. K. Paramarta, and Sariyasa, "LBtrans-Bot: a Latin-to-Balinese script transliteration robotic system based on noto sans balinese font," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 3, pp. 1247–1256, Dec. 2018.
- [19] L. H. Loekito, G. Indrawan, Sariyasa, and I. K. Paramarta, "Error analysis of Latin-to-Balinese script transliteration method based on noto sans balinese font," in *Proceedings of the 3rd International Conference on Innovative Research Across Disciplines (ICIRAD)*, 2020.
- [20] G. Indrawan, I. K. Paramarta, D. P. Ramendra, I. G. A. Gunadi, and Sariyasa, "A method for the affixed word transliteration to the Balinese script on the learning web application," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 6, pp. 2849–2857, Apr. 2021.
- [21] I. M. Suatjana, *Bali Simbar Dwijendra*. Denpasar: Yayasan Dwijendra, 2009.
- [22] G. Indrawan, I. G. A. Gunadi, and I. K. Paramarta, "Towards ubiquitous learning of Balinese-to-Latin script transliteration as part of Balinese language education," in *Proceedings of the 4th Asian Education Symposium (AES)*, 2020.
- [23] G. Indrawan, K. U. Ariawan, K. Agustini, and I. K. Paramarta, "Finite-state machine for post-processing method of Balinese script to Latin transliteration," in *The 6th International Conference on Science, Technology, and Interdisciplinary Research (IC-STAR)*, in press.
- [24] J. Esling, *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge University Press, 1999.
- [25] S. Hollos and J. R. Hollos, *Finite Automata and Regular Expressions: Problems and Solutions*. Longmont, CO: Exstrom Laboratories LLC, 2013.
- [26] L. Groner and G. Manricks, *JavaScript Regular Expressions*. Birmingham: Packt Publishing, 2015.
- [27] A. Sunardi and Suharjo, "MVC architecture: a comparative study between laravel framework and slim framework in freelancer project monitoring system web based," *Procedia Computer Science*, 2019.
- [28] A. Moutouakkil and S. Mbarki, "MVC frameworks modernization approach: adding MVC concepts to KDM metamodel," *Int. J. Adv. Comput. Sci. Appl.*, 2019.
- [29] M. S. Singh, "MVC framework: a modern web application development approach and working," *Int. Res. J. Eng. Technol.*, 2020.
- [30] G. Indrawan, I. W. Sutaya, K. U. Ariawan, M. S. Gitakarma, I. G. Nurhayata, and I. K. Paramarta, "A method for scriptio continua management on the transliteration to the Balinese script," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 6, pp. 4182–4191, 2021.
- [31] K. M. M. Aung and A. N. Htwe, "Comparison of levenshtein distance algorithm and needleman-wunsch distance algorithm for string matching," *Natl. J. Parallel Soft Comput.*, 2019.
- [32] C. Zhao and S. Sahni, "String correction using the damerau-levenshtein distance," *BMC Bioinformatics*, 2019.
- [33] U. Roll, R. A. Correia, and O. Berger-Tal, "Using machine learning to disentangle homonyms in large text corpora," *Conserv. Biol.*, 2018.
- [34] J. R. Sánchez-González, "Hemi-and homonyms in the big data era," *Diversity*, 2020.
- [35] A. R. Widiarti and R. Pulungan, "A method for solving scriptio continua in Javanese manuscript transliteration," *Heliyon*, vol. 6, no. 4, p. e03827, Apr. 2020.

Validation: Conceptual versus Activity Diagram Approaches

Sabah Al-Fedaghi
Computer Engineering Department
Kuwait University
Kuwait

Abstract—A conceptual model is used to support development and design within the area of systems and software modeling. The notion of validation refers to representing a domain in a model accurately and generating results using an executable model. In UML specifications, validation verifies the correctness of UML diagrams against any constraints and rules defined within the model. Currently, significant research has been conducted on generating test sets to validate that UML diagrams conform to requirements. UML activity diagrams are a specific focus of such efforts. An activity diagram is a flexible instrument for describing a system's behaviors and the internal logic of complex operations. This paper focuses on the notion of validation using activity diagrams and contrasts that process with a proposed method that involves an informal validation procedure. Accordingly, this informal validation involves comparing requirements to specifications expressed by a diagram of a modeling language called thinging machine (TM) modeling. The informal validation is a type of model checking that requires the model to be small enough for the verification to be done in a limited space or time period. In the proposed method, the model diagram is divided into subdiagrams to achieve this purpose. We claim the TM behavioral model comes with a particular dispositional structure that allows a designer to “carve” a model into smaller components for informal validation, which is shown through two case studies.

Keywords—Validation; conceptual model; activity diagram; thinging machine; informal validation

I. INTRODUCTION

A conceptual model is a mathematical/logical/verbal representation (mimic) of a domain (real or proposed), situation, policy, or phenomenon developed for a particular study [1][2][3]. A conceptual (in contrast to an intentional mental representation such as sensation [4]) model describes “how we conceive of that domain” [5]. It is used to support development and design within the area of systems and software modeling (e.g., databases or business processes).

An example of such a model is a description developed using the Unified Modeling Language (UML) to construct a representation of a domain using primitive constructs and concepts such as the “lens” through which reality is perceived to capture that domain’s meaning [5]. Nonconceptual models such as mathematical models are presented in terms of variables and quantitative relationships (i.e., equations). By contrast, in conceptual models, the variables and relationships between variables are represented visually as a system of icons in a diagram [5]. Visual representations can help to shift the

focus to enhanced qualitative conceptual reasoning, serve as representations of an internal (mental) model, and provide a means for communication and analysis. In the context of conceptual modeling, to ensure a system’s quality, it is critical that the model that represents the domain be semantically correct. This is confirmed by checking that the model satisfies some correctness properties and the system requirements.

A. Validation

Validation is the process of confirming that models are understood, defined well, documented, and based on established fundamentals [6][7]. It conveys a sense that “a scientific effort must be justified in some logical, objective, and algorithmic way” [6]. However, determining whether a particular model fulfills requirements by validating it over the complete domain of its intended applicability often is not possible (e.g., due to cost and time). Instead, tests and logical reasoning are conducted until adequate assurance is achieved that the model can be considered valid for its intended application.

Regarding UML, as a semi-informal notation, significant research efforts have gone into the so-called model-driven testing of UML diagrams. Such efforts mostly involve generating high-level test cases that can be used to validate both specifications and implementations [8]. Specifically, activity diagrams are highly useful for validating requirements with customer representatives [9]. Activity diagrams have become an established modeling notation for various levels of abstraction, ranging from fine-grained descriptions of algorithms to high-level workflow models in business applications [10].

B. Problem: Semantics of the UML Diagram

According to Tariq, Sang, Gulzar, and Xiang [11], the absence of formal semantics for UML activity diagrams makes it difficult to build automated tools for analyzing and validating such diagrams. Recently, UML 2.0 introduced token-driven semantics for activity diagrams inspired by Petri nets. One of the goals of the Foundational UML Subset (fUML [12]) is to provide a well-defined execution of UML activity diagrams. Accordingly, additional validation tools are needed for diagrammatic representation in the context of conceptual modeling. This paper proposes such a tool using a new type of conceptual model.

C. Approach and Limitations

This paper focuses on works that solely examine the notion of validation using activity diagrams, as an example of the current state of research in the validation field. Validation of UML activity diagrams using directed test cases is very promising [8]. The present paper complements such studies by examining validation under “equivalent” representations using a diagrammatic model based on thinging machine (TM) modeling. The aim is to propose a particular technique for model validation based on TM.

II. REVIEW

An UML activity diagram is a semi-formal semantic specification that is intuitive and flexible. It is used to describe a system’s behaviors and the internal logic of complex operations. Therefore, it is widely utilized as a front-end tool for system-level design of software and/or hardware systems.

A. Review: Graphs, Petri Nets, and Event B

The validation literature on developing test cases for activity diagrams is extensive (e.g., [8]). According to Chen and Mishra [8], “Most directed test case generation work is performed by human intervention. Hand-written test cases entail laborious and time-consuming effort of verification engineers who have deep knowledge of the design. Due to the manual development, it is difficult to generate all directed test cases to achieve a coverage goal. The problem is further aggravated due to the lack of comprehensive functional coverage metrics.” Many tools and methods have been developed to support test specifications and test case generation. For example, dSPACE developed a tool that uses activity diagrams for test descriptions and test script generation [13]. Chen, Poon, Tang, and Tse [14] presented a framework with which to construct test cases from specifications by identifying a set of input categories for the activity diagrams as test cases. Hettab, Kerkouche, and Chaoui [15] converted activity diagramming into grammar rules for graphs to capture all the relevant features for test case generation. Shirole, Kommuri, and Kumar [16] transformed activity diagrams into extended control flow graphs. Sunitha [17] incorporated Object Constraint Language (OCL) into activity diagramming for test case generation involving difficulties identifying complete behavior and static changes [18]. Chen et al. [19] matched Java program traces with behavior activity diagramming to identify changes resulting in a failure to identify static changes [18]. Sapna and Mohanty [20] converted UML activity diagrams into tree structures to prioritize scenarios by assigning weights to nodes and edges; however, this approach lacks in-depth code coverage and cannot identify static changes [18]. Some authors have developed frameworks to transform a UML activity diagram into Petri nets automatically using a model checker for analysis (e.g., [21][22]). A different approach involves transforming UML activity diagrams into Event B to specify and verify the distributed and parallel workflow solicitations [23].

B. Approach in this Paper: Informal Validation

In our approach to validation, we first produce an equivalent TM representation of the activity diagram. We consider the complexity of the representation and thus aim for

quick model checking by adopting informal reasoning. All of the reviewed validation methods discussed in the previous subsection can be applied to TM. Informal validation leads to discussing formal validation.

Formal specifications can be used to deliver a precise addition to different descriptions and can be validated, leading to specification faults being detected. Formal validation verifies the correctness of specifications, so it can be used to guarantee the quality of models (e.g., UML [8]). Despite the long interest in formal validation methods, “It seems that practitioners judge formal methods to be insufficiently beneficial to outweigh pragmatic problems” [24]. According to Amey [25], “Customers are often ‘aghast’ at the idea of formal methods being used to develop their products and might say ‘couldn’t you use UML?’” Amey [25] suggests overcoming such prejudices through “formality by stealth” and cites semantically strengthened UML as an example [24].

On the other hand, informal validation techniques rely on the opinions of modelers to draw a conclusion [26]. According to Petty [27], “Informal methods are more qualitative than quantitative and generally rely heavily on subjective human evaluation, rather than detailed mathematical analysis. Experts examine an artifact, for example, a conceptual model expressed as UML diagrams, and assess the model based on that examination and their reasoning and expertise.” Examples of informal methods include inspection, face validation, the Turing test, desk checking, and walkthroughs [27]. According to Banks [26], “In all cases though it is important to note that informal does not mean it is any less of a true testing method. These methods should be performed with the same discipline and structure that one would expect in ‘formal’ methods. When executed in such a way, solid conclusions can be made.”

The purpose of informal validation is to examine the accuracy of a domain’s representation in a conceptual model and in the results produced by the executable model [27]. In this type of validation, the concept of a system is viewed as a group of interacting components, and its desired functionality is articulated by graphical means [28].

Accordingly, in our proposed approach, the validation process involves requirements (e.g., expressed in English) versus specifications (expressed by TM diagrams). The validation here involves showing that the TM model is the correct model for the requirements. Thus, informal validation is a type of model checking that requires “the model to be small enough so that the verification can be done in a limited space or time” [29]. Accordingly, the TM diagram is divided into subdiagrams for this purpose. Our claim is that the TM behavioral model comes with a particular dispositional structure that allows a designer to “carve” a diagram into smaller components for informal validation.

C. Outline

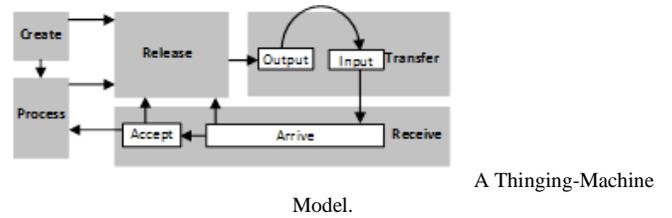
The next section reviews the basic constructs of a TM model. Section 4 presents a case study of validating the process of buying a beverage from a vending machine. Section 5 presents a second case study of validating an online shopping system.

III. THE THINGING MACHINE

TM modeling is a way of understanding how things and processes have come to be structured (see [30] and its TM-related references by the author of the present paper). As shown in Fig. 1, a TM can be described as the following generic (basic) actions:

- Arrive: A thing moves to a machine.
- Accept: A thing enters the machine. For simplification, we assume that all arriving things are accepted; hence, we can combine the arrive and accept stages into one stage: the receive stage.
- Release: A thing is ready for transfer outside the machine.
- Process: A thing is changed, but no new thing results.
- Create: A new thing is born in the machine.
- Transfer: A thing is input into or output from a machine.

Additionally, the TM model includes the mechanism of triggering (denoted by a dashed arrow in this study's figures), which initiates a flow from one machine to another. Multiple machines can interact with each other through the movement of things or triggering. Triggering is a transformation from one series of movements to another.



IV. VALIDATING A VENDING MACHINE

Sapna and Arunkumar [20] considered an example of an activity diagram for the process of buying a beverage from a vending machine (Fig. 2).

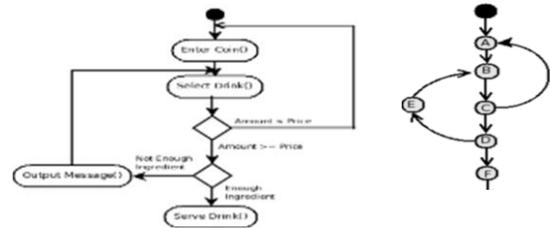


Fig. 2. Partial Views of the Diagrams used to Model the Process of Buying a Vending Machine Beverage Found in Sapna and Arunkumar [20].

In this section, we first produce the corresponding TM models and then apply the validation strategy to this vending-machine example.

A. Static Model

Fig. 3 shows the corresponding static TM model. In Fig. 3,

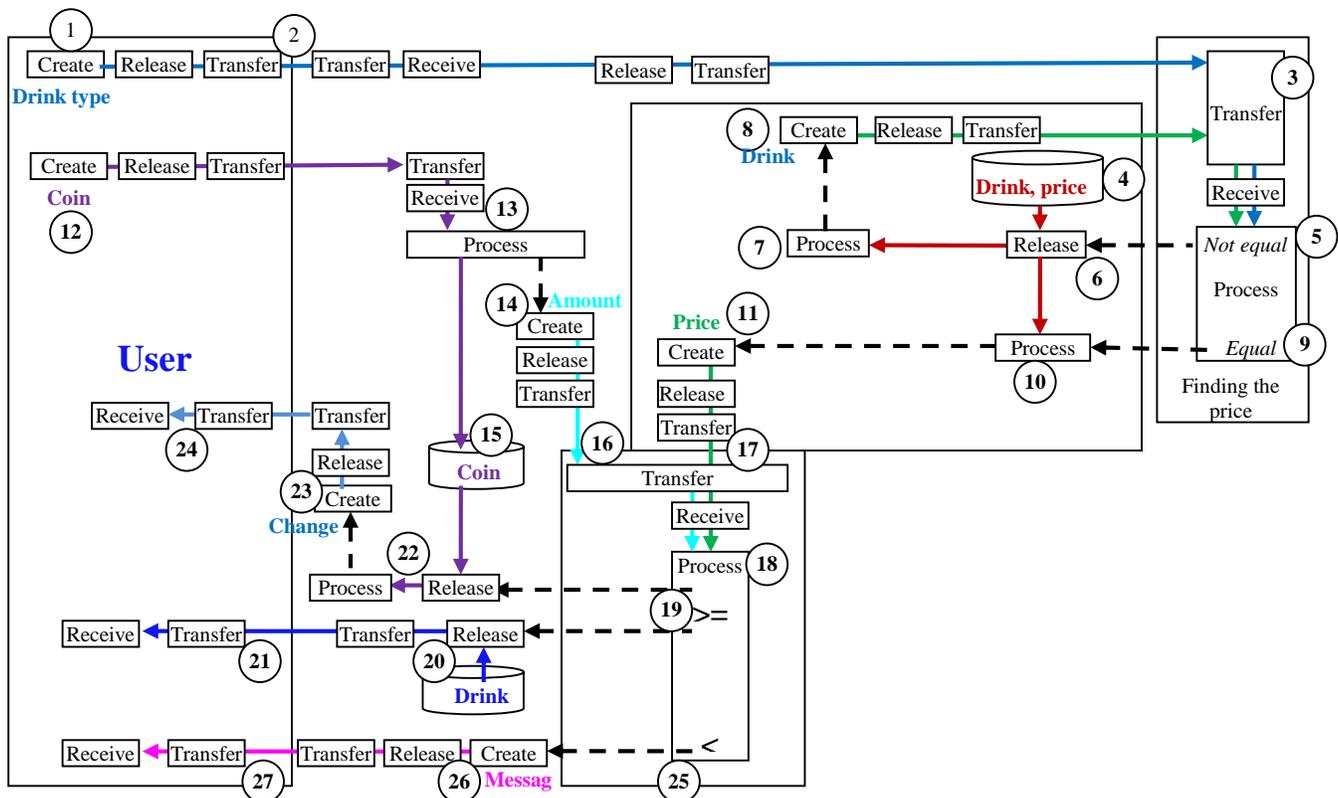


Fig. 3. The Static TM Model of a Vending Machine.

- The user creates (circle 1) a drink selection that flows to the machine (2), where it is sent to a module that finds the (drink, price) record (3). The drink data are found by extracting records (drink, price) (4) one by one and comparing the input drink with the drink data in the matching record.
 - If the drink data inside the record are not equal to the input drink (5), then the next record is released from the set of records and processed (7) to trigger the creation of drink data (8). Note that the drink item is *created* in the sense that it was not known to the machine as an independent thing before it was extracted from the record.
 - If the drink data inside the record match the input drink (9), then this triggers processing of (10) the record to extract the price (11). (Note that, in the activity diagram, the price is a thing “dropped from the sky”. There is no connection between the drink and its price.)
- Meanwhile, the user inputs coins (12), which are processed (13) inside the machine to calculate their value (14). (Note that the activity diagram does not distinguish the coins as physical objects from their amount and value.) Then, the coins are stored in coin boxes (15).
- Both the amount (16) and the price (17) flow to a module that compares them (18).

- If the amount is equal to or greater than the price (19), then the drink is released to the user (20 and 21). Additionally, the coin storage is processed (22) to create change, which flows to the user (23 and 24).
- If the amount is less than the price (25), then a message is created and flows to the user (26 and 27).

B. Behavioral Model

To produce the TM behavioral model, we must identify all events in the vending-machine model. An event in TM modeling is formed from a subset of the static model in addition to a time subthimac. For example, Fig. 4 shows the event *the machine receives a drink selection*.

Identifying a phenomenon’s behavior involves dividing it into component parts and then fitting the behaviors of these parts into a whole. Accordingly, the static model (Fig. 3) can be divided into events as shown in Fig. 5, which shows only the regions of the events for simplification. The resulting list of events is as follows.

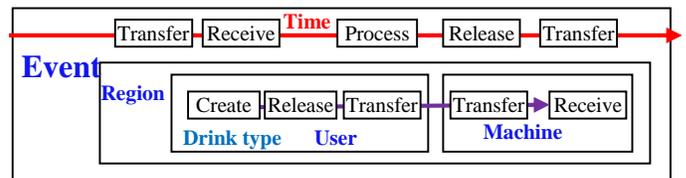


Fig. 4. The Event the Machine Receives a Drink Selection.

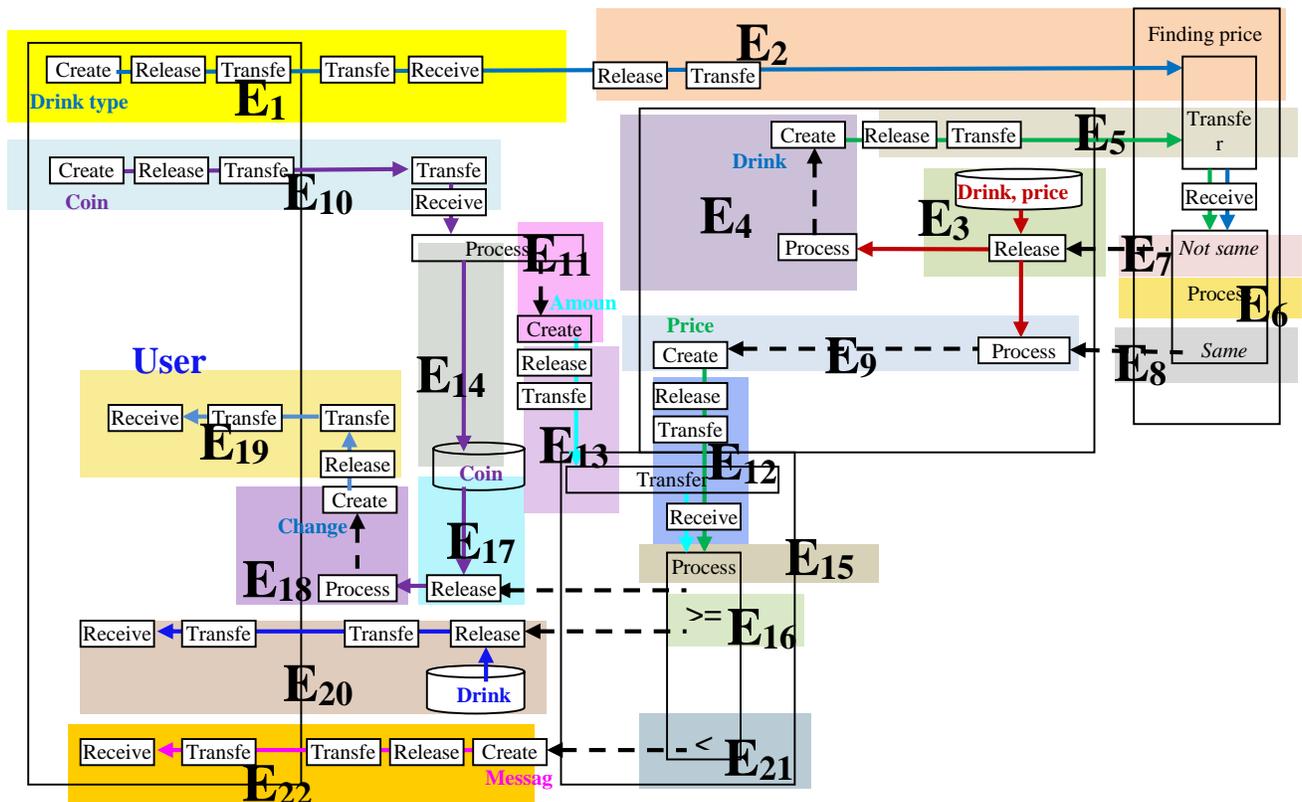


Fig. 5. The TM Events Model of the Vending Machine.

Event 1 (E₁): The machine receives a drink selection.

Event 2 (E₂): The selected drink flows to the price-finding module.

Event 3 (E₃): A record (drink, prices) is retrieved from the list.

Event 4 (E₄): The selected drink is extracted from the record.

Event 5 (E₅): The drink is sent for comparison with the input drink.

Event 6 (E₆): The input drink is compared with the stored drink.

Event 7 (E₇): The input drink is not the same as the stored drink.

Event 8 (E₈): The input drink is the same as the stored drink.

Event 9 (E₉): The price is extracted.

Event 10 (E₁₀): The user inputs coins.

Event 11 (E₁₁): The amount of the coins' value is calculated.

Event 12 (E₁₂): The coins are deposited into the coin boxes.

Event 13 (E₁₃): The amount flows to a comparison with the price.

Event 14 (E₁₄): The price flows to a comparison with the amount.

Event 15 (E₁₅): The amount and the price are compared.

Event 16 (E₁₆): The amount is equal to or greater than the price.

Event 17 (E₁₇): The coin boxes are processed.

Event 18 (E₁₈): The change is extracted from the coin boxes.

Event 19 (E₁₉): The change flows to the user.

Event 20 (E₂₀): The drink is released to the user.

Event 21 (E₂₁): The input amount is less than the price.

Event 22 (E₂₂): A message is sent to the user.

Fig. 6 shows the behavioral model for the vending machine.

C. Validation Strategy

Until this point, we have focused only on the modeling notations. It is time to ask whether the vending-machine blueprint fulfills the requirements. Requirements typically are written in natural language, but the behavioral diagram provides a skeletal structure of events.

Plato famously employed the "carving" metaphor as an analogy for the reality of Forms (Phaedrus 265e): As if we were animals, the world comes to us predivided. Ideally, our best theories will be those that carve nature at its joints [31]. The behavioral model comes with a particular dispositional structure rooted in the five generic types of events: create, processes, release, transfer, and receive.

Validation in TM modeling refers to event validity, which involves the model's events being compared to those of the reality to determine whether they are similar. In TM modeling, a general approach to validating the developed model can be developed using a logical process in which one takes higher-level events produced by the carving process and reduces them to the constituent events, which, in turn are based on generic events. This implies validating each generic action or sequence of these actions. However, because of space limitations and the informal nature of this study, we will not employ such a process but rather simply sketch operational descriptions of events with which to validate the model. This method assumes that the model's validity can be determined from observations.

Fig. 7 shows the decomposition of the behavioral model (Fig. 6) into three parts (super-events), for which the joints suggest division among three super-events as follows.

- Super-event 1: Selecting a drink and finding the price
 - The machine shall accept requests for n types of drinks.
 - The machine determines the price of the selected drink.

Fig. 8 shows the events involved: E₁ through E₉. The figure shows that the verification method involves feeding, internally, all drinks stored in the machine to the machine to verify that the machine performs the two requirements specified above. This verification process covers all legitimate inputs to E₁. An actual verification system can be constructed to input the tuples to E₁.

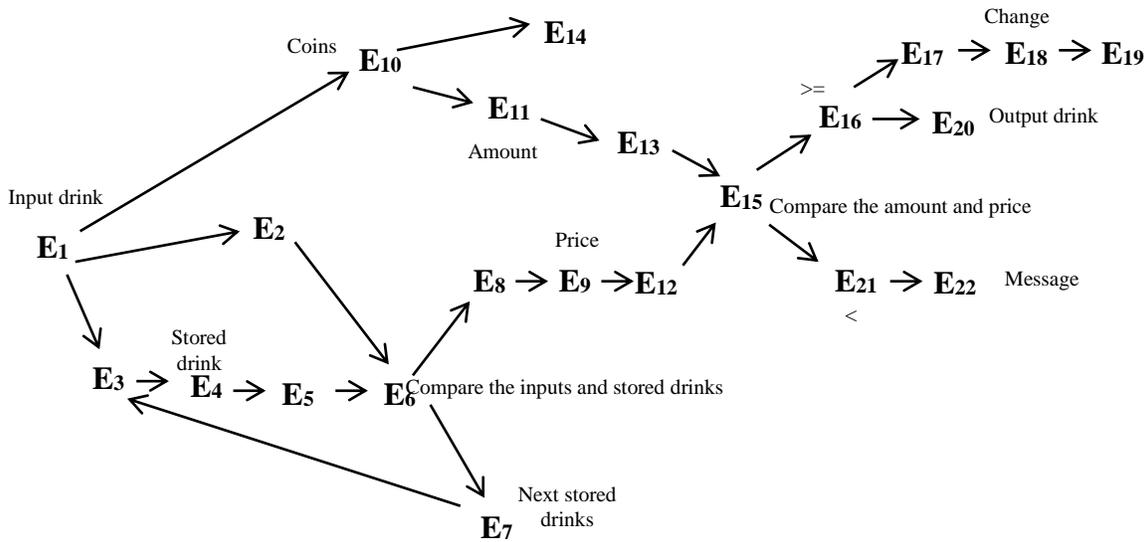


Fig. 6. The Behavioral Model.

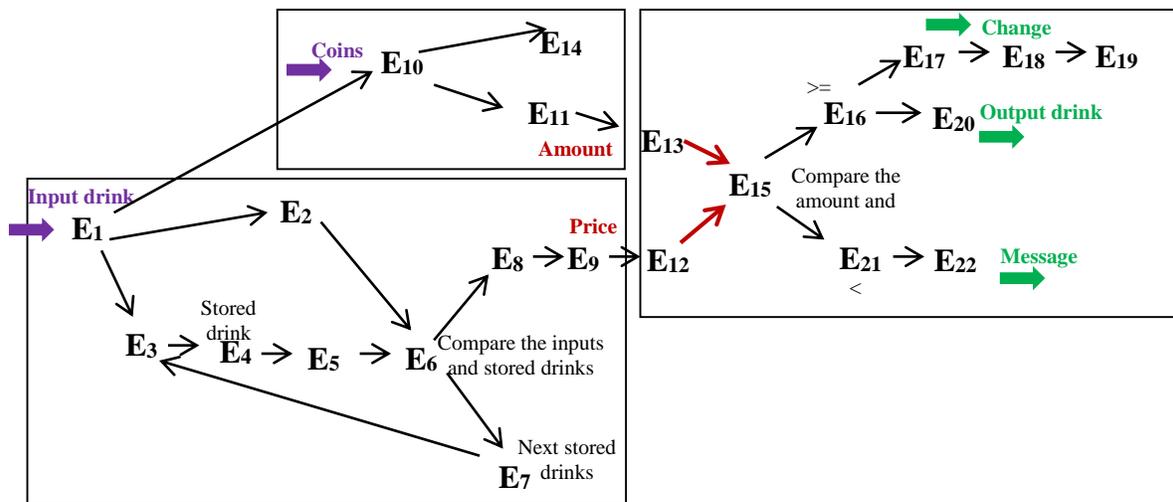


Fig. 7. Three Super-Events in the Behavioral Model.

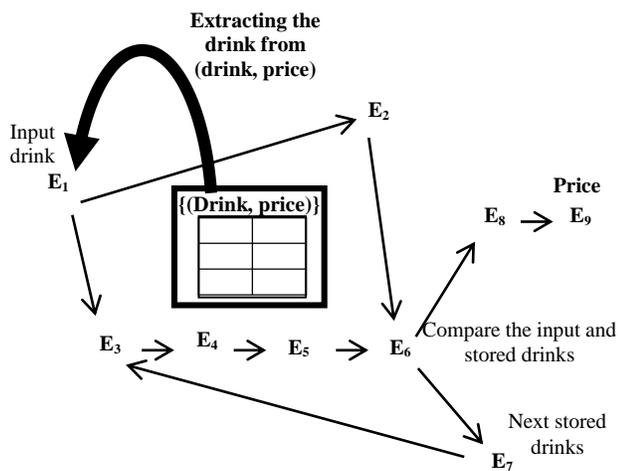


Fig. 8. Verifying that All Drinks have Prices.

Verification of selecting a drink and finding the price

Verification version of E_1 : loop for all drinks:
Read a tuple from the (drink, price)
Extract the drink from the tuple
Perform $E_2, E_3, E_4, E_5, E_6, E_7, E_8,$ and E_9
If (ERROR), then process a report

- upper-event 2: Validating the coins
 - The machine accepts all specified coin types, assuming there are three types.
 - The machine calculates the digital values of these coins.
 - The machine stores the coins in their appropriate places.

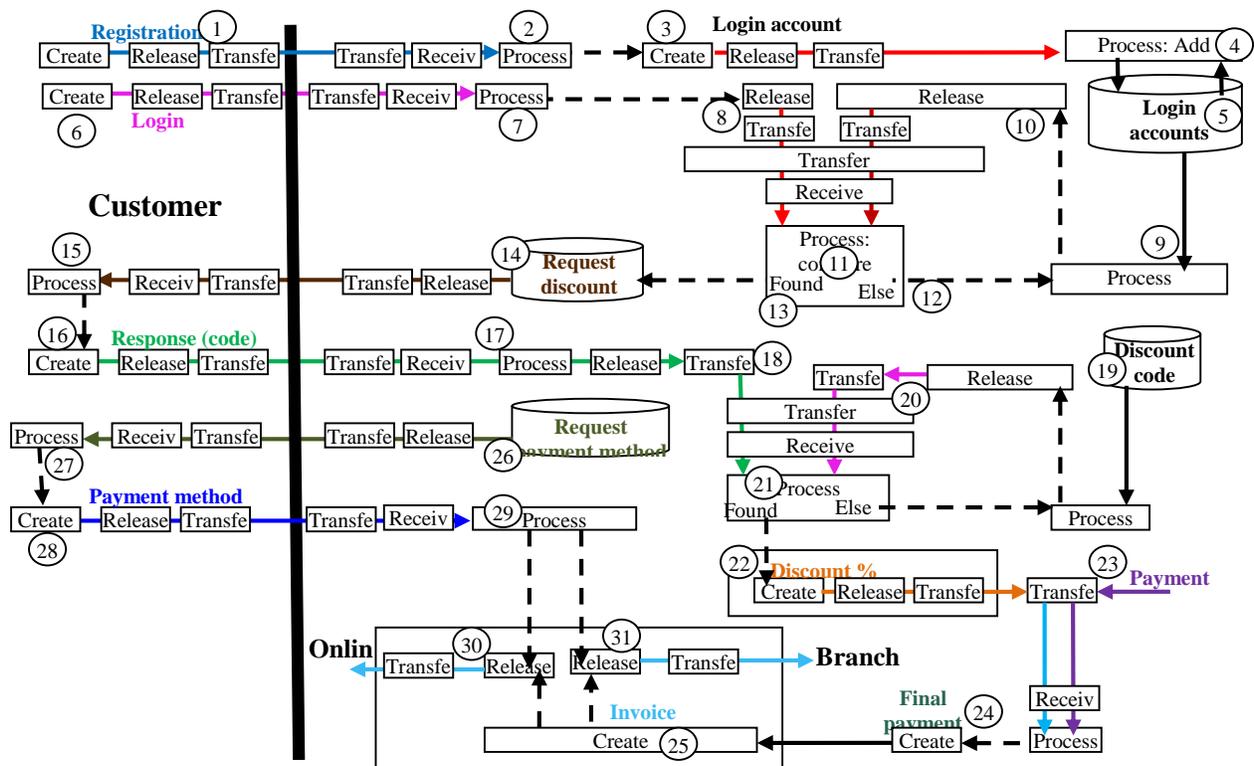


Fig. 12. The Static TM Model of the Online Shopping System.

First, the customer applies for registration (circle 1), which is processed by the system (2) to create a login account (3) that is added to the set of registered accounts (4 and 5). Note how the login account file is processed to add a new account.

- The customer requests to log in (6) and the request is processed (7) to extract the login account from the request (8). Additionally, the accounts file is processed (9) to retrieve an account (10), which is compared with the input account (11). If the two accounts are not the same, then the next account in the accounts file is retrieved for comparison. This process continues until the two accounts are found to be the same (13). Here, we ignore the situation in which the account is not filed in the file because the activity diagram does not mention it. Here, we can add a trigger for an error message to flow to the customer.

- Then, the system sends a request for a discount (14), which is processed by the user (15), to reply (16) with a code (*no discount* is a type of code). The code is processed (17) to be compared with the set of codes (18, 19, 20, and 21). When the code is found, the percentage of discount is calculated (22) and, together with the price (23), is processed to calculate the payment (24) and invoice (25).
- Accordingly, the system requests the method of payment (26), and the customer processes (27) that request to input such a method (28), which flows to the system, where it is processed (29). According to the payment method, the system either sends an invoice online (30) or to the branch (31).

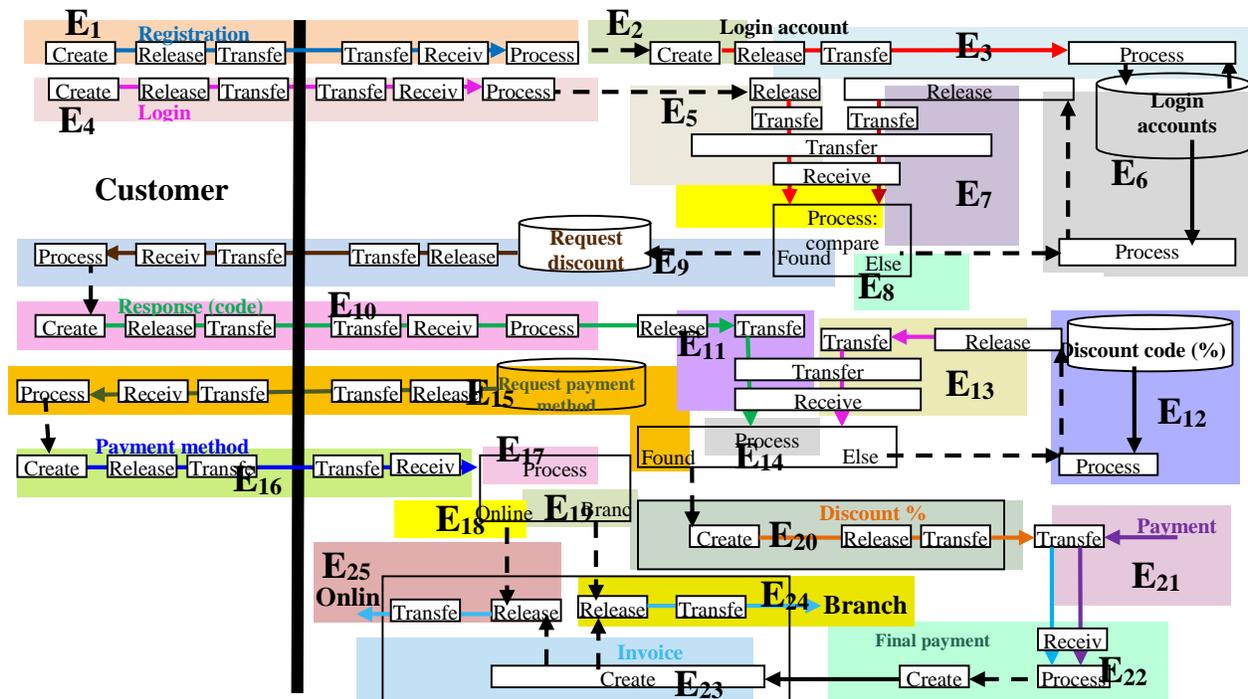


Fig. 13. Events Model. Note that in the Activity Diagram, the Price is considered an Input to the Whole Process.

The events in the model can be specified as follows (see Fig. 13).

- Event 1 (E₁): A customer registers to log in.
- Event 2 (E₂): The system creates a new login account.
- Event 3 (E₃): The system adds the new account to the accounts file.
- Event 4 (E₄): A customer sends a login request.
- Event 5 (E₅): The system extracts the login account from the request and sends it to be checked as a legal account.
- Event 6 (E₆): The accounts file is processed to retrieve an account, which is sent for comparison with the input account.
- Event 7 (E₇): The input account is compared with the account retrieved from the file.
- Event 8 (E₈): The input account is not the same as the account from the file.
- Event 9 (E₉): The input account is found among the legitimate accounts; hence, a request for the discount code is sent to the customer.
- Event 10 (E₁₀): The customer sends a discount code (possibly a code for no discount).
- Event 11 (E₁₁): The code is sent to find its corresponding discount percentage.

- Event 12 (E₁₂): The list of codes is processed to retrieve one code at a time.
- Event 13 (E₁₃): The retrieved code is sent to be processed.
- Event 14 (E₁₄): The code is compared with the list of codes.
- Event 15 (E₁₅): The code is found; thus, a request for the payment method is sent to the customer.
- Event 16 (E₁₆): The customer sends the payment method.
- Event 17 (E₁₇): The payment method is processed.
- Event 18 (E₁₈): The payment method is in the branch.
- Event 19 (E₁₉): The online payment method is chosen.
- Event 20 (E₂₀): The code is found; thus, the discount percentage is extracted.
- Event 21 (E₂₁): The price is received.
- Event 22 (E₂₂): The discount percentage and price are used to calculate the required payment.
- Event 23 (E₂₃): The payment is used in generating the invoice.
- Event 24 (E₂₄): The invoice is sent to the branch.
- Event 25 (E₂₅): The invoice is sent to the online payment system.

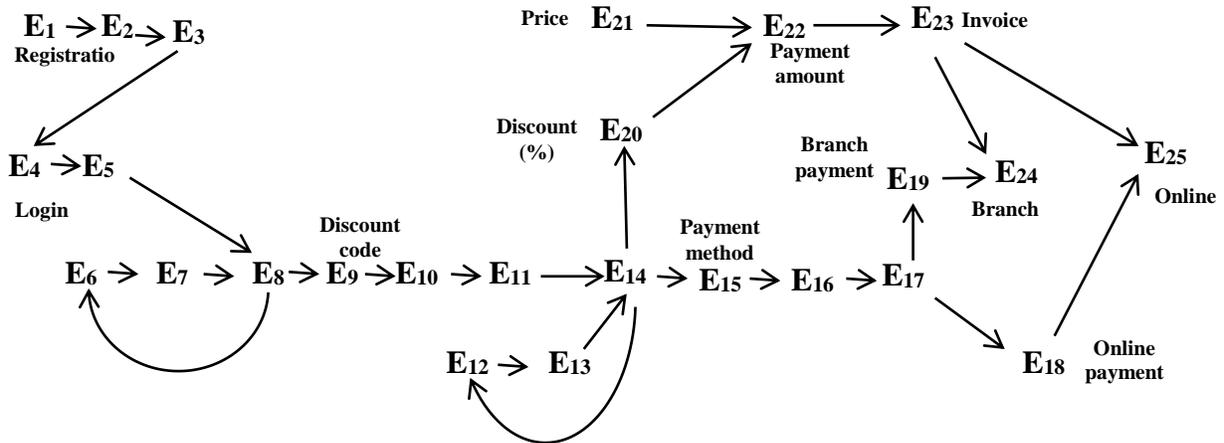


Fig. 14. The Behavioral Model.

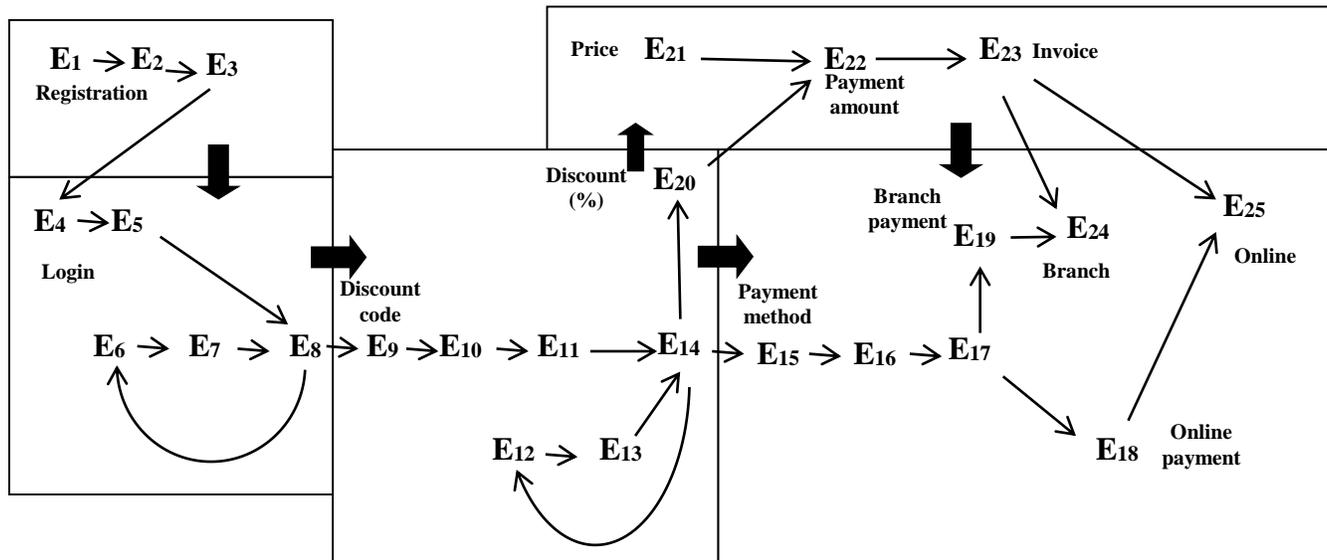


Fig. 15. Carving Small Components from the Behavioral Model.

Fig. 14 shows the behavioral model decomposed into three parts (super-events), in which the joints suggest division among five super-events. As shown in Fig. 15, these super-events are as follows.

- Registration component
- Login component
- Discount percentage component
- Payment component, and
- Method-of-payment component

Apparently, our hypothesis that TM representation would lend itself to such division of high-level events is true for this shopping-system representation. We can apply the same type of informal validation.

However, it is important to point out that the super-events may have relationships with “use cases”; thus, UML use case diagrams are a topic to be investigated in future research.

VI. CONCLUSION

This paper focused on examining the notion of validation using activity diagrams and proposed an informal validation process. This validation process involved requirements, versus specifications expressed by a diagram. Informal validation is a type of model checking that requires the model to be small enough to verify in a limited space or time. Accordingly, the model diagram is divided into subdiagrams for this purpose. We claimed that the TM behavioral model comes with a particular dispositional structure that allows designers to “carve” a diagram into smaller components for informal validation. This was shown through two case studies concerning vending machine and online shopping systems.

This result seems plausible because TM modeling is founded upon five generic actions. Thus, generic events have no subevents, and higher-level events are built from these generic events. Generic events can happen in diverse regions, and they can reoccur. It seems the building structures from the five generic actions “collapse” into smaller structural partitions according to certain aspects such as functionality. The number

(e.g., 7 ± 2) and nature (e.g., basic) of these actions seem to be crucial features that determine the system's overall level of complexity. Additionally, the TM model (Fig. 1) seems to generate nested hierarchies or levels with loosely coupled connections (through only transfers and triggers), which inhibit large structural complexity.

These explanations are still a type of speculation and require further research to be applied in different aspects of modeling systems beyond validation.

REFERENCES

- [1] R. G. Sargent, "Verification and validation of simulation models," in Proc. 1994 Winter Simulation Conference, J. D. Tew, M. S. Manivannan, D. A. Sadowski, and A. F. Seila, Eds. Piscataway, NJ: IEEE, 1994, pp. 77–87.
- [2] H. C. Mayr and B. Thalheim, "The triptych of conceptual modeling," *Software and Systems Modeling*, vol. 20, no. 1, pp. 7–24, 2021.
- [3] J. Fischer, B. Møller-Pedersen, A. Prinz, and B. Thalheim, "Models versus model descriptions," *Modelling to Program: Second International Workshop*, vol. 1401, pp. 67–89, May 2021.
- [4] D. Pitt, "Mental representation," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed. Stanford, CA: Stanford University, 2018.
- [5] N. Guarino, G. Guizzardi, and J. Mylopoulos, "On the philosophical foundations of conceptual models," in *Frontiers in Artificial Intelligence and Applications [e-book]*, vol. 321, Information Modelling and Knowledge Bases XXXI, IOS Press, 2019, pp. 1–15. DOI: 10.3233/FAIA200002
- [6] G. B. Kleindorfer, L. O'Neill, and R. Ganeshan, "Validation in simulation: Various positions in the philosophy of science," *Management Science*, vol. 44, no. 8, pp. 1087–1099, 1998.
- [7] R. G. Sargent, "Verification and validation of simulation models: An advanced tutorial," 2020 Winter Simulation Conference (WSC), Dec. 14–18, 2020. DOI: 10.1109/WSC48552.2020.9384052
- [8] M. Chen, P. Mishra, and D. Kalita, "Efficient test case generation for validation of UML activity diagrams," *Des. Autom. Embed. Syst.*, vol. 14, pp. 105–130, 2010. DOI: 10.1007/s10617-010-9052-4
- [9] L. Fernandez-Sanz and S. Misra, "Practical application of UML activity diagrams for the generation of test cases," *Proc. Romanian Academy, series A*, vol. 13, no. 3/2012, pp. 251–260, 2012.
- [10] H. Gronniger, D. Reiß, and B. Rumpe, "Towards a semantics of activity diagrams with semantic variation points," 13th International Conference, MODELS 2010, Proceedings, Part I. Lecture Notes in Computer Science 6394. Springer, 2010, pp. 331–345. ISBN 978-3-642-16144-5
- [11] O. Tariq, J. Sang, K. Gulzar, and H. Xiang, "Automated analysis of UML activity diagram using CPNs," 8th IEEE International Conference on Software Engineering and Service Science, pp. 24–26, Nov. 2017. DOI: 10.1109/ICSESS.2017.8342881
- [12] OMG, "UML 1.3 specifications," OMG [online], 2017.
- [13] L. Lavagno and W. Muller, "UML as a next-generation language for SoC design," *Electronic Design*, 2006.
- [14] O. T. Chen, P. Poon, S. Tang, and T. Tse, "Identification of categories and choices in activity diagrams," in *International Conference on Software Quality*, pp. 55–63, 2005.
- [15] A. Hettab, E. Kerkouche, and A. Chaoui, "A graph transformation approach for automatic test cases generation from UML activity diagrams," *Proceedings of the Eighth International Conference on Computer Science & Software Engineering*, pp. 88–97, 2015. DOI: 10.1145/2790798.2790801
- [16] M. Shirole, M., Kommuri, and R. Kumar, "Transition sequence exploration of UML activity diagram using evolutionary algorithm," *Proc. 5th India Software Engineering Conference*, pp. 97–100, 2012. DOI: 10.1145/2134254.2134271
- [17] E. V. Sunitha and P. Samuel, "Enhancing UML activity diagrams using OCL," *IEEE International Conference on Computational Intelligence and Computing Research*, 2013.
- [18] P. K. Arora and R. Bhatia, "Agent-based regression test case generation using class diagram, use cases and activity diagram," *Procedia Computer Science*, vol. 125, pp. 747–753, January 2018. DOI: 10.1016/j.procs.2017.12.096
- [19] M. Chen, X. Qiu, W. Xu, L. Wang, J. Zhao, and X. Li, "UML activity diagram-based automatic test case generation for Java programs," *The Computer Journal*, vol. 52, no. 5, pp. 545–556, 2006. DOI: 10.1093/comjnl/bxm057
- [20] P. G. Sapna and B. Arunkumar, "An approach for generating minimal test cases for regression," *Procedia Computer Science*, vol. 47, pp. 188–196, 2015. DOI: 10.1016/j.procs.2015.03.197
- [21] Y. Rahmoune, A. Chaoui, and E. Kerkouche, "A framework for modeling and analysis UML activity diagram using graph transformation," *Procedia Computer Science*, vol. 56, pp. 612–617, 2015.
- [22] T. S. Staines, "Intuitive mapping of UML 2 activity diagrams into fundamental modeling concept Petri net diagrams and colored Petri nets," *Engineering of Computer-Based Systems, ECBS 2008. 15th Annual IEEE International Conference and Workshop*, pp. 191–200, March 2008.
- [23] L. J. Ayed and N. Hamdi, "From UML activities diagrams to event B," *IFAC Proceedings*, vol. 42, no. 4, pp. 420–425, 2009.
- [24] C. Snook and M. Butler, "UML-B: Formal modeling and design aided by UML," *ACM Transactions on Software Engineering and Methodology*, vol. 15, no. 1, pp. 92–122, January 2006. DOI: 10.1145/1125808.1125811
- [25] P. Amey, "Dear sir, yours faithfully: an everyday story of formality," *Practical Elements of Safety, Proceedings of the 12th Safety-Critical Systems Symposium*, pp. 3–15, 2004.
- [26] C. M. Banks, "Introduction to modeling and simulation," in *Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains*, J. Sokolowski and C. Banks, Eds. Wiley, 2010.
- [27] M. D. Petty, "Verification, validation, and accreditation," in *Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains*, J. Sokolowski and C. Banks, Eds. Wiley, April 2010.
- [28] F. D. McKenzie, "Systems modeling: Analysis and operations research," in *Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains*, J. Sokolowski and C. Banks, Eds. Wiley, April 2010.
- [29] S. Dupuy-Chessa and L. du Bousquet, "Validation of UML models thanks to Z and Lustre," *Proc. FME: Formal Methods for Increasing Software Productivity, International Symposium of Formal Methods Europe, 2001*. DOI: 10.1007/3-540-45251-6_14
- [30] S. Al-Fedaghi, "UML sequence diagram: An alternative model," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 12, no. 5, 2021. DOI: 10.14569/IJACSA.2021.0120576
- [31] M. H. Slater and A. Borghini, "Introduction: Lessons from the scientific butchery," in *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science*, J. Keim Campbell, M. O'Rourke, and M. H. Slater, Eds., MIT Press, 2013. DOI: 10.7551/mitpress/9780262015936.003.0001
- [32] M. Bures, B. S. Ahmed, and K. Z. Zamli, "Prioritized process test: An alternative to current process testing strategies," *International Journal of Software Engineering and Knowledge Engineering*, vol. 29, no. 07, pp. 997–1028, 2019. DOI: 10.1142/S0218194019500335.

Identifying Small and Medium Enterprise Smart Entrepreneurship Training Framework Components using Thematic Analysis and Expert Review

Anis Nur Assila Rozmi¹, Puteri N.E. Nohuddin²

Institute of IR4.0
Universiti Kebangsaan Malaysia
Bangi, Malaysia
Universiti Kuala Lumpur
British Malaysian Institute
Gombak, Selangor

Abdul Razak Abdul Hadi³

Universiti Kuala Lumpur Business School
Kuala Lumpur, Malaysia

Mohd Izhar A. Bakar⁴

UniKL British Malaysian Institute
Gombak, Malaysia

Abstract—Small and Medium Enterprises (SMEs) today are facing a competitive business environment, in a complex and rapidly changing environment. For that technology is seen as a mediator capable of transforming SMEs to greater heights in an amid and vigorous pace of a borderless world. The agenda of SMEs to generate national income as well as to create more employment opportunities has made the government much focused in providing improvements in business opportunities to SMEs to boost the country's economic growth. To ensure that the SME owners sustain their business, they should be able to adapt the use of the internet as a key component in designing new business model values, customer experiences and internal capabilities that support the key operations. However, there are still some SME owners who do not leverage on the use of Information and Communication Technology (ICT) in their business operations. This study interviewed eight SME owners who operated their businesses in Kuala Lumpur and Selangor to identify a list of most important business training courses needed for SMEs in Malaysia. The data was analyzed using Thematic Analysis method and it was found that there are five main components of courses in SMEs, namely, Business Management, Sales and Marketing, Accounting and Finance, ICT and Technology, and Production and Operations. As a result of this Thematic Analysis study, researchers have developed a smart entrepreneurship training framework related to the five components and produced a system called, the Malaysian SMEs Psychometric Test or U-PPM which has been reviewed and endorsed by the respective panels of experts. This proposed framework is important for SME owners and management and also the government and stakeholders, when making the correct decisions in selecting business training courses as well as to increase ICT and digital technologies usage in providing a positive impact to all SMEs in Malaysia.

Keywords—Small Medium Enterprise (SME); business owner; thematic analysis method; expert panel; Information and Communication Technology (ICT); course selection system; smart entrepreneurship training framework

I. INTRODUCTION

Entrepreneurship has become an important agenda in order to build a more sustainable national future [38] [77].

Entrepreneurial development approaches improved modern economy [44] and provides spaces in creating employment opportunities [10]. The importance of this field has been applied at the university level [49] with the cooperation of both government and private sectors. There are various programs, training sessions, courses, workshops, seminars, and entrepreneurship education activities that have been implemented to encourage entrepreneurs to obtain formal knowledge and education in the field of entrepreneurship, to face challenges in today's sophisticated business models.

In order to realize the government's expectations to produce more entrepreneurs in the country, various initiatives have been made. Courses and training sessions have been conducted to make SMEs more competitive, provide a simple service and customer support system. SMEs that have been operating for a long time have lots of experience in running their business, but they sometimes face difficulties in several aspects such as internal management, marketing capabilities, technological capabilities, access to knowledge networks, and finance [74].

Thus, it is important for the training providers to implement standards and models of entrepreneurship-based education as well as to develop courses and skills that can be used by the entrepreneurs. This is because the role of training and courses is to help SME entrepreneurs and staff to enhance their performance [35]. In addition, the skills and knowledge are required for them to perform their functions. Entrepreneurs, especially micro-sized businesses are encouraged to find relevant courses to improve their business and their employee's performance. They should be given these trainings and courses so that they can be multi task entrepreneurs.

Among the courses and training activities that are essential to SME owners and staff are those covering management skills [51], leadership [27], communication [63], technical skills [39] [48], according to the field of enterprise, digital marketing [24] [25], and financial management skills [15] [59]. Training is required to ensure that the employees are able to perform their tasks in order to meet the objectives of the enterprise. Hence,

this training needs to be planned carefully in order for it to have an impact on the success of the business.

To ensure that the correct selection is done properly, this study has developed a Psychometric Test model for SMEs using the Thematic Analysis (TA) method. The construction of this model has been evaluated and reviewed by a group of experts from the government and private sectors in the areas of Business Management, Digital Marketing, Accounting, Business Operations, and Language.

II. LITERATURE REVIEW

A. Impact of Courses on Organizations

Entrepreneurs' level of education contributed to the performance and development of the business organization [84]. This happens because they did not emphasize the importance of training and improving knowledge and skills. The lack of skills and access to finance, knowledge of marketing, infrastructure and information had made it difficult to achieve success [40]. Therefore, SME owners need to get training and education at an earlier stage for the training programs to have a greater impact on enterprises [50].

The impact of entrepreneurial training on businesses is through a proactive attitude, innovation, and willingness to take any risks [4]. Additionally, SME entrepreneurs who participate in entrepreneurship training and courses experience a short-term improvement in entrepreneurial self-efficacy [28]. By choosing the right course, SME owners are able to improve their business skills [58]. Employees and SME owners who attended short-term courses can see results of their self-skills training attended by business entrepreneurs and employees [72]. Such impacts had an increase in sales revenue and profits, helped broaden the mind, and inculcated in the owners and employees the ability to take strong personal initiatives and to have perseverance, which further improved business practices.

The positive impacts of such courses were continued to be discussed in the study by [16] where they concurred that the effectiveness of such training is seen to be a valuable investment. In addition, the positive impact of course involvement can also be detected through an increase in the income as well as improvement in the financial performance of the enterprises [41].

The disposition of unskilled and unknowledgeable workers in enterprises is seen as a risk in losing business opportunities [46] other than having the tendency of becoming bankrupt [73]. On the other hand, employees who have gone through self-improvement courses and training provide a positive impact on the job, showing improvements in their performance and job quality, in addition to being motivated and loyal to the enterprise [30]. Therefore, SMEs should seize the available opportunities and take the necessary initiatives to improve and enhance the employees' knowledge and skills to boost competitiveness [69]. In addition, after they have undergone the courses and trainings, it would affect the growth, profitability and job creation in the enterprise [56].

B. Courses Options for Organizations

The Part of SME operations include management, finance, sales and marketing, ICT and technology, and production and

operations, which are also listed as elements that determine the success of an SMEs. Thus, courses and training that act as a medium for knowledge enhancement are seen to improve the performance of the sales force [16]. A decade ago, course selection was made based on performance competency models that built individual competencies [23]. Additionally, the competency model by [75] has been used as an instrument in providing a skilled workforce and meeting job needs.

There are various courses that can be studied by every SME employee in order to boost their skills or knowledge to improve their job performance. These courses include orientation training, internal and external company management training. The designed courses usually depend on the needs of SMEs [22], financial estimates [26] that can be spent and other priorities focused on by the SMEs. Before creating training contents, training providers need to evaluate the training needs so that the program meets the needs of the organization [47].

Other management functions had been prioritized such as accounting, marketing, finance, and production [50]. Therefore, these designed courses need to be holistic in nature, where the knowledge gained can be applied in the daily work of the employees, and it would also help in improving their work performance. Seven key areas are important for SMEs, and among the four courses that can be fitted and have impacts onto the training system in Malaysian SMEs are organizational management, sales and marketing, finance and accounting, and business planning [36].

Business Management courses are very important for owners in training individuals that they believe can lead the organization and other employees [43]. Among the topics covered in this organizational management course are leadership skills, personal skills such as how to motivate and delegate, development of industry and market analysis, operations planning, human resource management, developing sales strategies, and developing financial foundations, analysis and systems for start-ups. The goal of the course is to boost productivity, which is one of the core values of an organization.

According to [83], Human Resource Management (HRM) covers human resource planning, training, performance management, compensation management, safety, and health and employee relations. Good HRM can improve organizational performance by contributing to employees, customers, innovation, and productivity satisfaction. However, for small businesses such as SMEs, the importance of a new management approach is seen as a success if it can manage teamwork, job flexibility, and performance appraisal. SME owners need to understand the various management practices and the importance of staff training on SME performance.

Every enterprise requires the incorporation of digital marketing strategies to enhance sales [32] in today's economy. Therefore, a digital marketing course is seen as a holistic course because it is one of the best transformation platforms for businesses today. SMEs running online businesses need to implement digital marketing strategies that help online users to see them. Digital marketing is important to small businesses because it helps to further sales at a lower cost [71].

There are many benefits of digital marketing such as growing the business, cost effectiveness, ease of getting targeted with larger volume, availability of an exceptional marketing platform, ease of updating products or services, and ease of customer access to impart information on products or services [31] [76]. Digital marketing through social media, search engines, emails, contents, videos, infographics, photos, podcasts, e-books and newsletters can serve as a platform to potential customers who seek for products and services. Information technology has changed our lives and is now transforming traditional marketing methods into digital marketing. This course provides an in-depth focus on the application of marketing principles in the process of implementing marketing functions in new enterprises, including distribution channels, pricing strategies, promotion through social media, location, direct and indirect sales methods, consultation, customer management, and customer service.

This course provides an in-depth focus on the application of principles from accounting and finance courses, including applied activities such as earning, managing, and administering cash. In terms of specific competencies, it explores the use of financial software, as well as income development, income statements, balance sheets, and cash flow statements for commencement, together with their use in evaluating and managing an organization's finances, especially at an early stage. Financial knowledge is related to the managers' level of confidence in managing an enterprise [13].

Finance is the catalyst in economic growth and development in SMEs [2]. Therefore, it is important for every SME owner to participate in financial and accounting training programs, to acquire the knowledge of finance, financial attitudes, and financial behaviour, for it will have a positive impact on every financial step and financial access of the SME. Additionally, SME owners will be able to understand their financial situation better and plan their future finances and make accurate financial decisions as well as improve their level of financial access.

It is essential for employees to acquire the technology-related courses in regard to the technical aspects while working [33]. In terms of retails, it is insightful for employees to have knowledge on computers whilst communicating with customers. From a sales point of view, to be knowledgeable in technology is crucial to get new customers other than maintaining data of the existing customers. As for office management, technology knowledge has a basic usage in managing customers' data [65].

Technology has now changed the way customers access and obtain information. SMEs were also found to lack experts in the human resource field. SME owners should invest and support more activities such as training, expanding human resource development policies, and develop appropriate training methods using the latest technology of mass media such as Facebook or Instagram based on the specific needs for SMEs [19].

Currently, SMEs have been exposed to the use of ICT in making business transactions to compete in highly competitive markets while gaining access at the global level. The

importance of using e-businesses for the economic success and survival of SMEs has forced them to adapt to remain sustainable in the future. SMEs are now more likely to use ICT to improve data accuracy, speed up processes and reduce employees' blunders. The use of technology in financial matters has proven to spur the use of ICT among SMEs [12].

ICT can help SMEs in improving the standard of living. On the contrary, there are still SMEs who do not utilize online business transactions despite having an internet connection [85]. Thus, the role of this course is to encourage SMEs to use ICT in their daily business activities. ICT can help to improve business [11]. Nonetheless, SME owners should invest in relevant courses in training staff to manage the technology.

According to [5], training and taking courses in ICT can make the industry much mature in accordance with the relevant field. SME owners who provide formal courses have been proven to increase the level of readiness among their employees in accepting new technologies as well as creating skilled workers [78]. For example, the Kenyan government should support the training programs and widen the use of ICT among SMEs and they will see positive impacts on the country's development [45].

C. Limitations Faced by SME Owners in Course Selection

SME owners need to intensify their knowledge regularly [37]. They cannot afford to be left behind as the world is moving very fast, and they would incur losses if the way they do business is traditional [57]. There are various government initiatives to encourage SMEs to participate in courses and trainings to increase knowledge and skills. A study conducted by [68] shows that the support given by the government has resulted in a tremendous improvement in ICT skills amongst the SMEs. However, there are still many SMEs who have not yet had the opportunity to follow the courses provided. There are several factors that contribute to the limitations of SMEs in participating in the courses and training provided.

Among the factors are commitments such as family, children, and others. Thus, they are unable to find a suitable time to enhance their skills and knowledge of the business [70]. Additionally, time is also the main deterrent for them, preventing them to pursue the skills and knowledge that they need for their business. Therefore, for those who want to enforce or carry out short courses, they need to understand these constraints and find ways so that the courses that were developed can overcome these problems, and also are able to facilitate the affairs for both parties. Other than that, one of the major concerns shouldered upon the SME owners are the financial constraints [7] [70]. This is because with the present commitments, they are not ready for additional hardships. This is not the case for younger generations, as they have more opportunities and financial assistance for their learning level. Most SME operators are unaware of the existence of free courses provided by the government. This is due to the inadequate disclosure about the courses offered for them. They assume that the courses offered for them cover supplementary knowledge related to ICT, and that they are incapable to participate because they think it will incur a high cost.

In order to enhance knowledge, SMEs are urged to continuously provide mental and physical support to their employees. This is because a consistent support system from family, friends and employers will poise their spirit and diligence to seek knowledge. There is a positive opportunity for employees to complete course sessions if they get support from their employers [21]. In ensuring them to stay focused and motivated while engaging in a course, it is important for them to know the purpose and outcome of them attending the course [20]. This is needed to make them feel confident with their learning process.

D. Thematic Analysis (TA)

Thematic Analysis developed by [17] which have later been applied in various fields of study. According to them, TA is a process of identifying patterns or themes in qualitative data. The goal of TA is to identify themes and patterns in important data and use these themes to articulate an issue in research. This method will not only summarize the data but will also interpret and understand the data better. The interview method has been identified as the main method in finding the theme.

TA is a qualitative method that can be learned easily because it has complete methods and procedures [17]. TA is also described as flexible theory and meaningful knowledge [18], therefore, researchers should perform a structured procedure in controlling data thus producing reliable reports. This clearly proves that this method provides unique flexibility in accordance with the research questions and data forms developed by the researchers. This analysis takes into account the experiences, perspectives, practices and behaviors of the respondents through direct interaction. Qualitative research should show that the analysis that has been conducted is reliable, accurate, consistent, and covers the whole subject through recording, systematic in nature, with complete details to prove that a method such as TA is reliable to use [53]. An article wrote by [62] implemented TA as a method to identify factors affecting SME owners in adopting ICT in business and found reliable reasons that can contribute to other research. For the number of respondents, four is the minimum number of respondents coming from various backgrounds to get a clear picture of the problems encountered [42]. There are six phases developed by [17]. Table I shows the phases and descriptions of TA method.

For the first phase, the first step in qualitative analysis by [17] is for the researchers to familiarize themselves with the data by reading, repeating and understanding the entire data and transcripts. As for the second phase which is to generate the initial code, [17] suggested that researchers should manage the data in a more systematic way by encoding each relevant data segment on research study questions and used open code when it does not have pre-code-set, and should also grow and modify the code as it does through the coding process. After that, the researcher needs to discuss and develop some introductory ideas about the code. Then the transcript code assignment is done separately. Transcript encoding is performed on each relevant text segment to address the research question. After that, the code is compared and discussed before moving on to another transcript. Researchers have the option to create manually through printed transcripts

or by using qualitative data analytics software such as ATLAS.ti and Nvivo [34] which are suitable for large data sets. In addition, Microsoft Excel can also be used as an option in identifying themes.

For the third phase of finding the theme, [17] identified that the theme is a pattern that captures something important about the data from the research questions. If the data set is small, there may be similar data between the encoding level and the initial theme identification level. In this case the researcher studies the codes that eventually become the theme. At the end of this step, the code has been organized into a broader and more specific theme of the research question. This theme is descriptive and illustrates patterns in data related to research questions. This fourth phase as the theme review phase, where researchers need to study, modify, and develop the introductory theme already known in Phase 3 [17]. At this point, all data related to each theme is collected and classified according to colors. The next step is to relate with the function of the theme in an interview. The themes should be clear, and they should be different from the others. This is the final stage of the theme which is developed by [17], where the fifth phase is the determination phase, used to identify the essence of the theme. Themes need to be discussed, to know if there are sub-themes, how they interact and relate to the main theme and how the themes relate to one another. In this analysis, feedback is a holistic theme rooted in other themes. Lastly, the results of this study from TA should be recorded in the form of a report.

E. Expert Panel Judgement

In the development of a model or the use of modified research instruments, a panel of experts can be involved as individuals who can provide feedback on the importance, appropriateness, and accuracy of the content, as well as clarity of meaning of each item in the research instrument [6]. There are three steps in the instrument validation staged by the panel of experts, namely, pilot study, reliability, validity analysis and item improvement [8]. The panel of experts is selected from a population of experienced researchers and industry players from various fields in entrepreneurship and language. Definition of panel of experts in the context of this study as people who have published journals in their respective fields or have experience working in the field for several years and hold important positions in an organization [14]. According to the study, the appropriate number of expert panels for each study depends on the research's needs. There are studies that use the services of three expert panels [29], and five expert panels [3] [9] [80].

TABLE I. PHASES OF THEMATIC ANALYSIS BY [17]

Phase	Description
1	Familiarizing yourself with the data
2	Generating initial codes
3	Searching for themes
4	Reviewing potential themes
5	Defining themes
6	Producing the report

III. METHODOLOGY

A. Objectives and Aims

The methodology used in this study is divided into two main phases, namely, the data collection phase and the conceptual development phase of the qualitative study. These two main phases have their own methods, but the output from the first phase is the input for the second phase. The first phase is the data collection phase using the Thematic Analysis method. The second phase is the development phase of the smart entrepreneurship training framework for the use of SMEs in Malaysia. The purpose of this study is to identify courses that are relevant for SMEs in Malaysia. The target groups in this study are SME owners who own businesses in Kuala Lumpur and Selangor.

B. Thematic Analysis Method

The first phase conducted in this study is the data collection process. This study was conducted through the method of interviewing a total of eight SME owners and analyzing them using the TA method developed by [17]. This is the method to identify and analyze the meaning patterns in data sets, and to describe the themes that are the most important in a conducted study. TA method is suitable to identify the objectives for this study. Furthermore, this process begins when researchers are able to identify potential issues while collecting the data [17].

C. Instruments

The method used for this study was interview. Interviews are used to gather information ranging from how to select the appropriate courses and the type of courses, to the performance changes shown after attending the course. The first part of the interview begins with the consent notice requesting that each respondent express their consent to record the interview using a Sony ICD-UX543F digital voice recorder. Next, demographic questions are provided in the second part of the interview form to get more information about the participants and businesses. The third part of the form presents objective of the study along with the questions used by the researcher, to each respondent. The questions are used as the basis of the interview and the researcher reserves the right to add to the other questions depending on the answers given by the respondents. This is to get the actual reactions and reality faced by the respondents in sharing information about the course and the impact after attending the course.

After recording the data from the respondents, the researchers transcribed the data into word form and used Microsoft Excel to store the data. Survey data were then analyzed using the Thematic Analysis method. The list of interview questions is as follows:

- 1) Do you know about the courses?
- 2) Is there a need for employees to attend courses?
- 3) List the courses that you or the staff has attended?
- 4) How often do you send employees to attend courses?
- 5) Which section needs the most courses?
- 6) How do companies / employees find suitable courses?
- 7) Is the employer or employee deciding to attend the course?

- 8) Does the company provide financial assistance for the use of the course?
- 9) What is your opinion on the duration of the course?
- 10) How to identify the appropriate time for employees to be sent for courses?
- 11) How are the changes shown by the employees after attending the course?
- 12) What changes are seen in the employees after attending the course?
- 13) Is there a connection between the course and the performance of employees?
- 14) What is your response to the course you have attended?
- 15) Can you share what you want in a course?

D. Respondents

A total of eight SME owners had agreed to be the respondents for this study. All respondents were selected randomly after identifying their business background. The respondents consisted of 2 males (33.3%) and 6 females (66.7%). The age of the respondents selected was between 32 and 57 (Average: 42.625; Standard Deviation: 7.652). The type of businesses of the respondents was based on products (33.33%) and on services (66.67%). All respondents had been in business for 4 to 21 years. 75% of the respondents had been in business from 1 to 10 years, 12.5% from 11 to 20 years, and 12.5% from 21 to 30 years. Table II shows the details of the respondents' background. The eight respondents came from various business backgrounds. Their age ranged from 32 to 57 years, and the year of establishment was from 1999 to 2016. All eight respondents had attended business courses throughout their business.

TABLE II. DETAILS OF EIGHT (8) RESPONDENTS FOR INTERVIEW SESSIONS

R	Gender	Establishment Year	Product / Service	Type of Business
R1	Female	2016	Product	Cosmetic
R2	Female	2013	Service	Tailor
R3	Male	1999	Product	Halal Meats
R4	Female	2016	Product	Bakery
R5	Female	2006	Service	Tailor
R6	Male	2010	Service	Medical Supplies
R7	Female	2014	Product	Printing and Souvenirs
R8	Female	2016	Service	Tailor

E. Procedure

As mentioned before, this study was conducted through interviews. The respondents required to complete this study are the SME owners who have been conducting business activities for more than four years by selling products or offering services. The selection of respondents is random, and business is conducted around Kuala Lumpur and Selangor, Malaysia. Researchers had contacted respondents and set a time to meet. After meeting at the promised place and time, the researcher explained about the study to be conducted and followed by an agreement form to record the interview. Respondents were

asked to sign a consent form for the interview to be recorded and asked to fill in personal and business information in the second part of the form. In the third part of the consent form, the researcher had listed three (3) study objectives and questions used for the study. During the interview, the researcher used the question and occasionally added questions depending on the answers and information shared by the respondents. As a token of appreciation, the researchers presented souvenirs from the Famous Amos shop at the end of the interview session. The same process was repeated until the eighth respondent.

Phase 1: Thematic Analysis:

After the interview session, the researchers collected all the audio data and uploaded the file to a laptop and created a set of storage files for security purposes. All twelve audio data files were then transcribed into words using Excel software. The Excel data was then printed and repeated reading was done to understand the details of the interview.

Based on the set of questions asked during the interview session, the researchers tried to get the starting code based on the answers given by the respondents. Researchers reduced the data to be more focused on important data. After creating almost nine types of code and classifying using pencil colors, the process then moved on to the next transcript.

After identifying the starting code, the researcher performed the third phase which is to find the theme by collecting the codes which then formed the same set of data sets for each transcript.

After finding the theme, the researcher reviewed, modified, and developed the initial theme and reviewed the theme to ensure it coincides with the objectives of this study. The data was then reviewed repeatedly and only data that felt like it was important was stored. The theme should be clear with the objectives of the study and should be found in all twelve of the transcripts.

The researcher then analyzed and determined the theme and found sub-themes where this data interacts and relates between the main themes in the data set. After obtaining the code, themes and sub-themes, the researcher then formed a model framework based on the findings from the study. A report is produced which includes the objectives of the study, literature review, research methods, analysis, results, and a thorough discussion of this study.

Phase 2: Expert Panel for Course Selection Concept Framework

The involvement of a panel of experts is aimed at obtaining their confirmation of the accuracy of the concept framework of suitable course selection for SMEs presented by the researchers. It refers to the basic knowledge that SME owners need to know in managing an organization. All instruments and items identified were brought to a group of experts to check their accuracy according to the purpose of the study. To this end, researchers have met with four experts in the areas of organizational management, marketing and digital marketing, accounting, operations, and Language. In the meeting, the researchers demonstrated the smart entrepreneurship training

framework developed using the TA method. The concept was then accompanied by 55 items based on the theme namely business management, sales and marketing, accounting and finance, information technology and operations and production. The combination of these four experts consisting of academics and business consultants aims to get a more comprehensive view of the SME course selection concept framework. These experts not only look at the theoretical aspects, but also look at the practical aspects that can be applied and will have an impact on SME entrepreneurs in Malaysia. Table III shows the background of the U-PPM system expert panels according to expertise [1].

TABLE III. BACKGROUND OF THE EXPERT PANEL

Panel Experts	Work Experiences
Panel Expert 1 Educational Leadership and Language Expert	1. Associate Professor Universiti Teknologi Malaysia (1989 – Current) 2. Director General Dept. of Community College (2009-2011)
Panel Expert 2 Marketing and Entrepreneur Development Expert	1. Lecturer Politeknik Sultan Abdul Halim Mu'adzam Shah (2017 – current) 2. Lecturer Politeknik Mu'adzam Shah Pahang (2016-2017) 3. Lecturer Politeknik Tuanku Syed Sirajudin Perlis (2010-2013) 4. Lecturer Politeknik Sultan Abdul Halim Mu'adzam Shah (2008-2009)
Panel Expert 3 Business Management, Logistic and Operation Expert	1. Director Technoputra Division Universiti Kuala Lumpur (2019 – Current) 2. General Manager UniKL Resources Sdn.Bhd (2014 – 2018) 3. Senior Lecturer UniKL Malaysian Institute of Industrial Technology (2009 – 2014)
Panel Expert 4 Accounting and Digital Marketing Consultant Expert	1. Founder & Director UpRiser Success Terminal 2. Director of Training Management JNJ Management Consultants

IV. RESULTS

A. Phase 1: Thematic Analysis

As indicated earlier, the TA method was used in this study to identify the list of relevant courses for the SMEs. The results of this method will be used as an input for the development of a smart entrepreneurship training framework for SMEs. This study has found a list of courses attended by entrepreneurs, and it has been divided into five basic divisions in an organization. The divisions are in the areas of business management, sales and marketing, accounting and finance, ICT and technology and production and operations. Table IV is an analysis using the TA Method to identify the theme. It is arranged with the prefix code obtained by the researcher according to the same set of answers given by the respondents.

TABLE IV. THEMES, SUB-THEMES AND CODES FOR THE LIST OF COURSES ATTENDED

Theme	Sub-Theme	Code	Respondents
Management	Management Strategy	Find ways to plan strategies for business fundamental, management and communication	R2, R4, R5, R8
	Leadership Skills	Find the internal skills you need to have as an owner	R7
	Human Resource Management	Courses on managing Human Resource	R1
Sales & Marketing	4P's (Place, Price, Promotion, Product)	Courses on digital marketing, Marketplace	R1, R2, R4, R5, R7, R8
Account & Finance	Cash Flow	Courses on accounts and finance	R1, R2, R4, R7, R8
	Tax	Courses on tax	R3
	Asset Management	Courses on managing company assets	R5
ICT & Technology	Infrastructure	Courses on technical and computer	R2, R6
	Digital Platform	Courses on apps and internet	R2, R7
Production & Operation	Standard Operating Procedure	Courses based on organization's need	R3, R5

Researchers identify that through the code of business planning strategy, business basics, and communication; it coincides with the sub-theme of management strategy and is directly within the business management theme. Four respondents have said that they have attended courses related to this topic to gain knowledge and understanding for the purpose of improving their business. Two of the narratives are as follows:

"Motivation classes, apps and knowledge about accounts been thought in the second week. Other than that, communication course and the techniques depend on each field." - R2.

"When I started my business, I took a basic business course from MARA. Then, I joined Baitulmal Entrepreneur Program, and they gave more entrepreneur courses." - R8.

For the code under internal skills, the sub-themes are leadership skills and management. The narratives from the respondents are as follows:

"Dr Azizan's course - Marketing and internal strength as an owner. He taught us how we have to handle a company, how we want to drive. What is the future holds? Mastering Marketing, Zero Marketing, Account. We went to all courses." - R7.

Still under the theme of Management with the sub-theme of Human Resource Management, the initial code achieved is to find a course on human resources and its narrative, which is as follows:

"I have attended the Hasbul Brothers course for Management. We have to take care of everything including cash flow, handle customer rejection and HR to take care of staff." - R1.

Under the theme of Sales and Marketing, the sub-themes listed include four 'Ps', namely, Place, Price, Promotion, and Product. The pre-code leads to the need of finding the courses related to marketing, digital marketing, and online platforms. Examples of narratives are as follows:

"Entrepreneurship courses, pre-business courses, Marketing courses." - R5.

"We went to Dr. Azizan's courses because it is more to Marketing." - R7.

"It is an online marketing now. Hence, we can't wait for customers at the store. So, we have to do something for example to search for suitable platform to go online." -R8.

For the accounting and finance theme, the first sub-theme is Cash Flow, and the initial code is looking for financial and account management courses. The narratives are as follows:

"Now, after attending the course, we knew that the cash flow has to be recorded and saved. Before this, the money was spent just like that. Hence, we managed to get the Management awareness which is an impact after attending the course." - R8.

"Motivation classes, apps and knowledge about accounts been thought in the second week. - R2.

"I went to Entrepreneurship course, ways to manage a business, managing financial data and our business channel and digital marketing." - R4.

Next, in the theme of accounting and finance, the sub-theme is taxation, where the code is a course on taxation. The narrative is as follows:

"It's all like Tax. You see, I'm good at talking, I can approach, but in terms of management, I am not good at that. The staff who went for the course had to handle everything (tax)." - R3.

Still under the theme of accounting and finance, the important course is to start with the initial code of managing the company's assets and lead to the sub-theme of asset management, the respondents' narrative is as follows:

"There are also financial and asset management courses for companies. That is important." - R5.

Under the theme of ICT and technology, the courses are considered important by the respondents. Under the infrastructure sub-theme, the initial codes are technical, and computer related. The narrative is as follows:

"Sometimes, we have to know the ICT and computers." - R2.

Furthermore, still under the theme of ICT and technology, the initial codes found are of applications and the internet. Both are under the sub-themes of digital platforms. The narrative is as follows:

"I went through a course at Taylor's University to learn about promoting our own products. I learnt to download the apps to beautify those pictures and the importance of WhatsApp Business." - R2.

"Irfan Khairi is more to Technical. He taught on the technical part on how to upload something through the Internet. It was in 2010 and the internet did not go very well back then." - R7.

Under the initial code of requirements of a particular organizational field, it leads to the sub-theme of Standard Operating Procedures (SOP), which is under the theme of production and operation. The respondents' narrative is as follows:

"It is a course for Halal to cut chicken and beef so this course is a must for hygiene." - R3.

"I am in the field of sewing. In 1994, I studied the course for full-time for a year." - R5.

Researchers have developed a smart entrepreneurship training conceptual framework using the results of the study using the TA method. Fig. 1 shows a complete chart on the selection of courses that are important for the needs of SME organizations doing businesses in Malaysia.

B. Phase 2: Findings from Expert Panel's Feedback

Following the development of a smart entrepreneurship training conceptual framework, a website was developed to help SME owners to make the right course selection for their organization. This is important because, there are various selections of courses available in the market; hence the SME owners have to make the right decision in choosing the courses that offer knowledge and experience which will then be beneficial for their business.

Researchers have developed the items based on the result of the analysis and the course selection concept framework development. Next, the development of the items and instruments needs to get an approval from the experts in the various fields. Hence, a group of experts in the related fields have conducted a research and validated the items and instrument used in the system known as the Malaysian SMEs Psychometric Test or U-PPM. This activity is done to examine the accuracy of items and instrument developed by researchers, to determine the courses chosen by SMEs to meet the needs of their organizations. Prior to the development of this U-PPM, a study was conducted based on in-depth interviews with eight SME owners. Researchers have analysed the results of the interviews using TA method and have produced an instrument that has five themes along with fifty-five items.

For that purpose, each expert has been asked to validate each theme and item. After reviewing all the items, the experts are required to fill out a form as an evidence of their agreement and feedback. Feedback obtained from all experts leads to the removal, addition, modification, and positioning of items. The development of this holistic and practical U-PPM system is expected to contribute to the field of theory and practice that benefits SMEs, especially to the training providers in Malaysia. Table V shows the results by the four experts. Items have been reduced to 50 from 55.

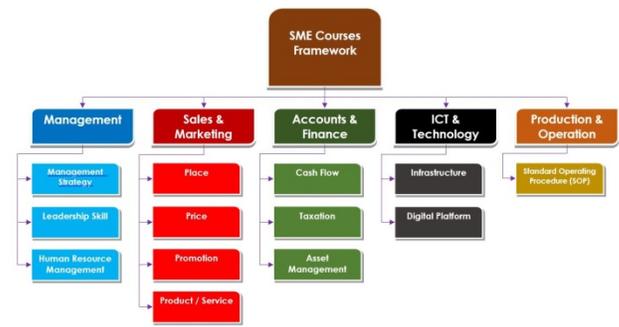


Fig. 1. SME's Smart Entrepreneurship Training Framework Components.

TABLE V. EXPERT PANEL (EP) RESULTS FOR U-PPM ITEMS

No	Details	EP 1	EP 2	EP 3	EP 4	%
1.	The format of the study instrument is appropriate and interesting	Agree	Agree	Agree	Agree	100
2.	The meaning of each item is clear	Agree	Agree	Agree	Agree	100
3.	The language used is easy to understand	Agree	Agree	Agree	Agree	100
4.	The instructions given are clear	Agree	Agree	Agree	Agree	100
5.	The instructions for the measurement scale are clear	Agree	Agree	Agree	Agree	100
6.	The objectives of the instrument stated are clear	Agree	Agree	Agree	Agree	100
7.	The number of items used is appropriate	Dis agree	Agree	Agree	Dis agree	50
8.	The whole idea presented in the instrument is interesting	Agree	Agree	Agree	Agree	100

V. DISCUSSION

Based on the result from the researcher's interview with eight SME owners, there are five main components in the selection of business training courses for SMEs in Malaysia. The main themes are business management, sales and marketing, accounting and finance, ICT and technology, and production and operations. These themes show that most SMEs chose to deepen their knowledge in the related field towards the organization they are leading. Table VI lists the themes and sub-themes derived from the TA method.

As for Business Management theme, the sub-themes resulting from the findings of this study are business management strategies, leadership skills, and human resource management. According to a study conducted by [52], SMEs can upgrade the entrepreneurial skills among them by acquiring lofty knowledge and management capabilities. This includes competency in distributing the tasks in the business, followed by assessing skills, and proficiency in human resource management. This finding is supported by [55] who says that human resource management is one branch of economics that is closely related to the success of an organization and business. For that, SME owners are encouraged to attend management and leadership skills courses

to boost their internal and peripheral capabilities in order to maintain sustainability and success in developing the enterprises [79].

As for the sales and marketing theme, an important sub-theme in marketing development is the application of mixed marketing elements, which are also known as 4P's, namely, place, price, promotion, and product. This mixed marketing framework plays an important role in strengthening the brand of SME products nationwide [66] other than influencing consumer purchasing decisions [64]. SMEs are found to be accomplished in conducting market research and defining the criteria and advantages of the products to customers and are capable in determining the rivalry strengths and weaknesses as well as to boost organisational monthly profits [52]. Apart from that, SMEs should engage in courses that apply knowledge related to ICT in order to enhance the way they do marketing and promotion of the products [84].

On the other hand, for accounting and finance codes in TA, the theme arising was involved with the knowledge of cash flow, taxation, and asset management. A study on financial management practices on MADA entrepreneurs conducted by [81] shows that those with a moderate level of financial experience were unable to manage their own finances without the need of help from others. Scores on financial knowledge, skills, attitudes, and financial planning are in the rank because they followed programs organized by their top management in acquiring the knowledge on financial management. Some of the entrepreneurs did not practice the system in documenting their transactions in the ledger, thus making the course on cash flow and taxation an important element in accounting and financial management [81].) Investigation on KADA entrepreneurs shows that their financial management is at a bearable level [67]. This is because, they have less experience. Therefore, KADA entrepreneurs are encouraged to participate in courses to improve financial knowledge and smoothen their business progress.

TABLE VI. THEMES AND SUB-THEMES FOR TA METHOD

Theme	Sub-theme
Management	Management Strategy
	Leadership Skills
	Human Resource Management
Sales & Marketing	Place
	Price
	Promotion
	Product
Accounts & Finance	Cash Flow
	Taxation
	Asset Management
ICT & Technology	Infrastructure
	Digital Platform
Production and Operation	Standard Operating Procedure

For the ICT and technology theme, the sub-themes are related to the infrastructure provided by SME owners and the importance of digital platforms. ICT has a tremendous economic impact on entrepreneurs if they have benefited to the fullest on the effective use of technology in advancing their businesses. However, their monthly income also affects the status of computer ownership among SMEs [61]. For SME participation in the digital field, a study by [54] said that digital participation is important for the advancement of their business and the source of information related to the latest products in the market. Entrepreneurs no longer need to face customers or have a one-way communication, because they have switched to streamline sales. Therefore, the application of knowledge on ICT needs to be further strengthened to bring these SMEs to a greater height. A study conducted by [60] showed that the factors of digital literacy, customer relationship building, and digital ecosystem building play an important role in influencing the use of digital platforms by SMEs. He added that these findings are a useful guide to stakeholders in designing educational programs related to the use of digital platforms for SME entrepreneurs.

For the production and operational theme, the sub-themes involved are related to the Standard Operating Procedures (SOPs) in an organization. According to [82], informal learning and previous experience have given lots of input to SMEs in managing their businesses. Therefore, SMEs need to acquire knowledge related to the management and products operation and services, because this would provide a positive impact in their business, will help in boosting their income, and hence help them remain competitive.

VI. CONCLUSION

This survey was conducted to develop a comprehensive course selection system for SMEs. The reason being, selecting an appropriate course could influence the performance of the SMEs other than to enhance profits for the organization. The development of this model uses two qualitative methods, which are, Thematic Analysis and Expert Panels. Through this study, the researcher found that there are five important business education courses needed by the organizations. The themes are business management, sales and marketing, accounting and finance, ICT and technology, and production and operations. For the management theme, there are three sub-themes, namely, management strategy, leadership skills, and human resource management. For the sales and marketing theme, sub-themes are mixed with marketing, which follows the 4Ps, namely, place, price, promotion, and product. The three sub-themes for the accounting and finance theme are cash flow, taxation, and asset management. The two sub-themes to the ICT and technology theme are infrastructure and digital platform. Finally, the sub-theme for production and operations is the Standard Operating Procedures in accordance with the needs of the respective organizational area of expertise.

These findings are useful for SME owners and the top management, who would be able to ensure that efforts and actions are taken in order to fully utilize the usage of an ICT in their business operation, and that the sustainability of their

organization would be maintained. As a result of the Thematic Analysis method, researchers have developed a smart entrepreneurship training framework of courses selection system for SMEs that has led to the development of items and instrument for a website known as the Malaysian SMEs Psychometric Test or U-PPM. Four expert panels from the areas of organizational management, marketing and digital marketing, accounting, and operations and language have agreed to carry out their responsibilities to validate the items used in the development of this U-PPM system for the next agenda of the study. It is hoped that with the existence of this U-PPM, SME owners can make the right decision in identifying the needs for acquiring business knowledge in entrepreneurship courses.

REFERENCES

- [1] Abd. Ghani, Basri and Adnan, Ahmad Azrin (2017) Penilaian pakar dalam model konseptual penggunaan berhierarki Islam / Basri Abd. Ghani and Ahmad Azrin Adnan. In: 2nd International Islamic Heritage Conference (ISHEC 2017).
- [2] Aca, L. A. O. (2018). Owner Financial Literacy Characteristic: Implication for Access to Finance Among Smes in Kwara State Nigeria. Fountain University Osogbo Journal of Management, 3(1).
- [3] Adnan, A. A. (2017). Penilaian pakar dalam model konseptual penggunaan berhierarki Islam/Basri Abd. Ghani and Ahmad Azrin Adnan.
- [4] Al-Awlaqi, M. A., Aamer, A. M., & Habtoor, N. (2018). The effect of entrepreneurship training on entrepreneurial orientation: Evidence from a regression discontinuity design on micro-sized businesses. The International Journal of Management Education, 100267.
- [5] Albors-Garrigós, J., Hervás-Oliver, J. L., & Márquez, P. (2009). Internet and mature industries. Its role in the creation of value in the supply chain. The case of tile ceramic manufacturers and distributors in Spain. International Journal of Information Management, 29(6), 476–482.
- [6] Ali, R., & Buang, N. A. (2018). Kompetensi Keusahawanan Sebagai Mediator Kesiapan Penerapan Elemen Keusahawanan dalam Kalangan Pensyarah Institut Pendidikan Guru (IPG) (Entrepreneurship Competency as a Readiness Mediator in the Implementation of Entrepreneurship Elements Among Lecturers of Teacher Institute). Jurnal Pendidikan Malaysia (Malaysian Journal of Education), 43(1SI), 123-130.
- [7] Antonioli, D., & Della Torre, E. (2016). Innovation adoption and training activities in SMEs. The International Journal of Human Resource Management, 27(3), 311-337.
- [8] Ariffin, S. R., Ahmad, J., & Najmuddin, N. A. (2010). The Developing Students Generic Skills Instrument Through Lecture Assessment Based on Many-Facets Rasch Measurement Model. Jurnal Pendidikan Malaysia (Malaysian Journal of Education), 35(2), 43-50.
- [9] Asbulah, L. H., Lubis, M. A., & Aladdin, A. (2018). Kesahan dan Kebolehppercayaan Instrumen Strategi Pembelajaran Kolokasi Bahasa Arab: Analisis Menggunakan Model Rasch (Validity and Reliability of Arabic Collocation Learning Strategies Instrument: Analysis Using Rasch Model). Jurnal Pendidikan Malaysia (Malaysian Journal of Education), 43(1SI), 131-140.
- [10] Assaad, R., Krafft, C., & Yassin, S. (2020). Job creation or labor absorption? An analysis of private sector job growth in Egypt. Middle East Development Journal, 1-31.
- [11] Assante, D., Castro, M., Hamburg, I., & Martin, S. (2016). The use of cloud computing in SMEs. Procedia computer science, 83, 1207-1212.
- [12] Azmi, A., Sapiei, N. S., Mustapha, M. Z., & Abdullah, M. (2016). SMEs' tax compliance costs and IT adoption: the case of a value-added tax. International Journal of Accounting Information Systems, 23, 1-13.
- [13] Bayrakdaroglu, A., & Şan, F. B. (2014). Financial literacy training as a strategic management tool among small–medium sized businesses operating in Turkey. Procedia-Social and Behavioral Sciences, 150, 148-155.
- [14] Beecham, S., Hall, T., Britton, C., Cottee, M., & Rainer, A. (2005). Using an expert panel to validate a requirements process improvement model. Journal of Systems and Software, 76(3), 251–275.
- [15] Bongomin, G. O. C., Ntayi, J. M., Munene, J. C., & Malinga, C. A. (2017). The relationship between access to finance and growth of SMEs in developing economies. Review of International Business and Strategy.
- [16] Bradford, S. K., Rutherford, B. N., & Friend, S. B. (2017). The impact of training, mentoring and coaching on personal learning in the sales environment. International Journal of Evidence Based Coaching and Mentoring, 15(1), 133.
- [17] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.
- [18] Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589-597.
- [19] Brien, E. O., & Hamburg, I. (2014). Supporting Sustainable Strategies for SMEs through Training, Cooperation and Mentoring. Higher education studies, 4(2), 61-69.
- [20] Byrne, J., Delmar, F., Fayolle, A., & Lamine, W. (2016). Training corporate entrepreneurs: an action learning approach. Small Business Economics, 47(2), 479-506.
- [21] Castaño Muñoz, J., Kalz, M., Kreijns, K., & Punie, Y. (2016). Influence of employer support for professional development on MOOCs enrolment and completion: Results from a cross-course survey. Research Track, 251.
- [22] Daud, S., Ahmad, S. N., & Zohor, R. M. (2018). Kajian Tinjauan Keberkesanan Promosi Kursus Pendek di Kolej Komuniti Segamat 2. Politeknik & Kolej Komuniti Journal of Life Long Learning, 2(1), 50-64.
- [23] Dubois, D., & Rothwell, W. J. (2004). Competency-based or a traditional approach to training. T and D, 58(4).
- [24] Dumitriu, D., Militaru, G., Deselnicu, D. C., Niculescu, A., & Popescu, M. A. M. (2019). A perspective over modern SMES: managing brand equity, growth and sustainability through digital marketing tools and techniques. Sustainability, 11(7), 2111.
- [25] Foroudi, P., Gupta, S., Nazarian, A., & Duda, M. (2017). Digital technology and marketing management capability: achieving growth in SMEs. Qualitative Market Research: An International Journal.
- [26] Frost, S. (2016). The importance of training & development in the workplace. Small Business, <http://smallbusiness.chron.com/importance-trainingdevelopment-workplace-10321.html>.
- [27] Garavan, T., Watson, S., Carbery, R., & O'Brien, F. (2016). The antecedents of leadership development practices in SMEs: The influence of HRM strategy and practice. International Small Business Journal, 34(6), 870-890.
- [28] Gielnik, M. M., Uy, M. A., Funken, R., & Bischoff, K. M. (2017). Boosting and sustaining passion: A long-term perspective on the effects of entrepreneurship training. Journal of Business Venturing, 32(3), 334-353.
- [29] Harun, N., & Ghani, F. A. (2016). Kesahan dan Kebolehppercayaan Soal Selidik Amalan Belajar Pelajar Berpencapaian Rendah Sekolah Berasrama Penuh. Jurnal Kemanusiaan, 14(3).
- [30] Imran, M., & Tanveer, A. (2015). Impact of training & development on employees' performance in banks of pakistan. European journal of training and development studies, 3(1), 22-44.
- [31] Jackson, G., & Ahuja, V. (2016). Dawn of the digital age and the evolution of the marketing mix. Journal of Direct, Data and Digital Marketing Practice, 17(3), 170-186.
- [32] Jain, T. K., & Sharma, A. (2019). Impact of Training and development on Employee Performance in Retail Sector: A Review paper. Available at SSRN 3316856.
- [33] James, J., Deacon, J., & Huxtable-Thomas, L. (2016). Relationship marketing in high technology based SMEs: A customer perspective. In Let's Get Engaged! Crossing the Threshold of Marketing's Engagement Era (pp. 413-425). Springer, Cham.

- [34] Joffe, H. (2012). Thematic analysis. *Qualitative research methods in mental health and psychotherapy*, 1.
- [35] Karim, M. M., Choudhury, M. M., & Latif, W. B. (2019). The impact of training and development on employees Performance: An analysis of quantitative data. *Noble International Journal of Business and Management Research*, 3(2), 25-33.
- [36] Katz, J. A., Hanke, R., Maidment, F., Weaver, K. M., & Alpi, S. (2016). Proposal for two model undergraduate curricula in entrepreneurship. *International Entrepreneurship and Management Journal*, 12(2), 487-506.
- [37] Kim, N., & Shim, C. (2018). Social capital, knowledge sharing and innovation of small-and medium-sized enterprises in a tourism cluster. *International Journal of Contemporary Hospitality Management*.
- [38] Kirby, D. A. (2004). Entrepreneurship education: can business schools meet the challenge? *Education+ training*.
- [39] Krishnan, T. N., & Scullion, H. (2017). Talent management and dynamic view of talent in small and medium enterprises. *Human Resource Management Review*, 27(3), 431-441.
- [40] Mabhungu, I., & Van Der Poll, B. (2017). A review of critical success factors which drives the performance of micro, small and medium enterprises.
- [41] Manresa, A., Bikfalvi, A., & Simon, A. (2019). The impact of training and development practices on innovation and financial performance. *Industrial and Commercial Training*.
- [42] Maree, J. G. (2015). Career construction counseling: A thematic analysis of outcomes for four clients. *Journal of Vocational Behavior*, 86, 1-9.
- [43] Milner, J., McCarthy, G., & Milner, T. (2018). Training for the coaching leader: how organizations can support managers. *Journal of Management Development*.
- [44] Mohd Aris, N., Abas, S. A., Mohd Adnan, S. D., Md Nasir, M. F., & Jalani, H. (2018). Modul usahawan tani Islam. *e-Journal of Islamic Thought and Understanding (e-JITU)*, 1(1), 16-32.
- [45] Mokaya, S. O., & Njuguna, E. W. (2017). Adoption and use of information and communication technology (ICT) by small enterprises in Thika Town, Kenya.
- [46] Mullins, R., Duan, Y., Hamblin, D., Burrell, P., Jin, H., Jerzy, G., ... & Aleksander, B. (2007). A Web Based Intelligent Training System for SMEs. *Electronic Journal of E-learning*, 5(1), 39-48.
- [47] Mungai, B. (2012). The relationship between business management training and small and medium-sized enterprises' growth in Kenya. Unpublished PhD Thesis, Kenyatta University.
- [48] Musa, H., & Chinniah, M. (2016). Malaysian SMEs development: future and challenges on going green. *Procedia-Social and Behavioral Sciences*, 224(2016), 254-62.
- [49] Mustapha, Z., Ahmad, A. H., Kob, C. G. C., & Zairon, I. Y. (2017). Peranan Kolej Komuniti Sebagai Pembimbing Dalam Memantapkan Niat Keusahawanan Pelajar. *Politeknik & Kolej Komuniti Journal of Social Sciences and Humanities*, 2(1), 209-216.
- [50] Nguyen, T. V., & Bryant, S. E. (2004). A study of the formality of human resource management practices in small and medium-size enterprises in Vietnam. *International small business journal*, 22(6), 595-618.
- [51] Noor, N. M. (2015). Pengukuran personaliti keusahawanan dan prestasi perniagaan bahan binaan di Negeri Kelantan (Doctoral dissertation, Universiti Malaysia Kelantan).
- [52] Noorzeli, N. M., & Wahab, M. N. A. (2017). Keupayaan Usahawan Kecil Melaksanakan Kemahiran Keusahawanan Menggunakan Kaedah HRV Biofeedback. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 2(1), 65-76. (Semua Perbincangan).
- [53] Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1609406917733847.
- [54] Omar, F. I., Rahim, S. A., & Dimiyati, H. A. (2019). Analisis Pola Penyertaan Digital ICT dan Transformasi Keusahawanan. *Jurnal Komunikasi: Malaysian Journal of Communication*, 35(2).
- [55] P Rameli, M. F., Sharif, D., & Che Man, N. (2017). Etika pengurusan sumber manusia bagi Muslimpreneurs dalam perniagaan berskala kecil/Mohd Faizal P. Rameli.
- [56] Pavlovic, S., Olukuru, J., & Coelho, J. (2019). Impact of Training on Small and Growing Businesses. Available at SSRN 3472940.
- [57] Penglin, L., & Honghao, L. (2017). The New Thinking of the SMEs Transformation and Upgrading in Shaanxi Province: from the Internet Finance Perspective.
- [58] Premand, P., Brodmann, S., Almeida, R., Grun, R., & Barouni, M. (2016). Entrepreneurship education and entry into self-employment among university graduates. *World Development*, 77, 311-327.
- [59] Ptak-Chmielewska, A., & Matuszyk, A. (2018). The importance of financial and non-financial ratios in SMEs bankruptcy prediction. *Bank i Kredyt*, 49, 45-62.
- [60] Ramdan, M. R., Abdullah, N. L., Isa, R. M., & Hanafiah, M. H. (2020). Meneroka Faktor-faktor yang Mempengaruhi Penggunaan Platform Digital oleh Perusahaan Mikro dan Kecil. *Jurnal Pengurusan (UKM Journal of Management)*, 59.
- [61] Rashid, S. M. R. A. (2016). Keupayaan ICT dalam meningkatkan pencapaian usahawan wanita: Satu kajian kes usahawan luar bandar di Malaysia. *e-BANGI*, 11(2), 78-103.
- [62] Rozmi A.N.A, Nohuddin P.N.E, Hadi A.R.A, Bakar M.I.A and Nordin A.I, "Factors Affecting SME Owners in Adopting ICT in Business using Thematic Analysis" *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(7), 2020.
- [63] Saad, M., Kumar, V., & Bradford, J. (2017). An investigation into the development of the absorptive capacity of manufacturing SMEs. *International Journal of Production Research*, 55(23), 6916-6931.
- [64] Sahir, S. H., & Rosmawati, R. (2020). Improve Marketing Mix for Marketing Plan Strategic in Coffeeshop Business. *Management Analysis Journal*, 9(4), 459-466.
- [65] Samašonok, K., Išoraitė, M., & Leškienė-Hussey, B. (2016). The internet entrepreneurship: opportunities and problems. *Entrepreneurship and Sustainability Issues*, 3, 329-349.
- [66] Sarker, T. R., & Al Saeed, L. S. L. A. (2020). Benchmarking Marketing and Business Strategy of UNIQLO to Start-up a Retail Shop in Bangladesh. *Benchmarking*, 12(2).
- [67] Sepeai, N., & Ramli, Z. (2019). Pengurusan Kewangan Usahawan: Kajian Kes Usahawan KADA di Kota Bahru, Kelantan. *Jurnal Wacana Sarjana*, 3(3), 1-13. (Kewangan).
- [68] Singh, R. K., Luthra, S., Mangla, S. K., & Uniyal, S. (2019). Applications of information and communication technology for sustainable growth of SMEs in India food industry. *Resources, Conservation and Recycling*, 147, 10-18.
- [69] Smith, A., & Hayton, G. (1999). What drives enterprise training? Evidence from Australia. *International journal of human resource management*, 10(2), 251-272.
- [70] Suseno, Y., Bao, C., Baimbridge, M., & Su, C. (2019). Informal training in Chinese small-and medium-sized enterprises. *International Journal of Entrepreneurship and Small Business*, 37(1), 1-24.
- [71] Teklehaimanot, M. L., Ingenbleek, P. T., Tessema, W. K., & van Trijpp, H. C. (2017). Moving Toward New Horizons for Marketing Education: Designing a Marketing Training for the Poor in Developing and Emerging Markets. *Journal of Marketing Education*, 39(1), 47-60.
- [72] Ubfal, D., Arraiz, I., Beuermann, D. W., Frese, M., Maffioli, A., & Verch, D. (2019). The Impact of Soft-Skills Training for Entrepreneurs in Jamaica. Available at SSRN 3374406.
- [73] Vancell, J. (2018). e-Learning for older workers in SMEs?: the perceptions of owners and workers in Maltese microenterprises.
- [74] Vandenberg, P., Yoshino, N., Goto, A., Intarakumnerd, P., & Miyamoto60, J. (2016). Policies to enhance SME internationalization. *Integrating SMEs into Global Value Chains*, 117.
- [75] Vazirani, N. (2010). Review paper: Competencies and competency model—A brief overview of its development and application. *SIES Journal of management*, 7(1), 121-131.

- [76] Yasmin, A., Tasneem, S., & Fatema, K. (2015). Effectiveness of digital marketing in the challenging age: An empirical study. *International Journal of Management Science and Business Administration*, 1(5), 69-80.
- [77] Youssef, A. B., Boubaker, S., & Omri, A. (2018). Entrepreneurship and sustainability: The need for innovative and institutional solutions. *Technological Forecasting and Social Change*, 129, 232-241.
- [78] Zaidan, E. (2017). Analysis of ICT usage patterns, benefits and barriers in tourism SMEs in the Middle Eastern countries: The case of Dubai in UAE. *Journal of Vacation Marketing*, 23(3), 248-263.
- [79] Zaimah, R., & Abdullah, S. (2017). Tahap keupayaan usahawan dalam Perusahaan Kecil dan Sederhana di Kuala Terengganu (The entrepreneurs' capability level in the Small and Medium Enterprises in Kuala Terengganu). *Geografia-Malaysian Journal of Society and Space*, 13(4).
- [80] Zain, R. M., & Ramli, A. (2020). Model Permasalahan Kesesakan Trak di Depoh Kontena Kosong: Kesahan Melalui Teknik Delphi. *Journal of Information System and Technology Management*, 5 (16), 35-49.
- [81] Zainol, A. S., & Ramli, Z. (2019). Pengurusan Kewangan Usahawan MADA: Kajian Kes di Kubang Pasu, Kedah. *Jurnal Wacana Sarjana*, 3(4), 1-10.
- [82] Zakaria, N. A. M., Hamid, M. A. A., Norazman, I., & Thukiman, K. (2020). Pengalaman Pembelajaran Informal Dalam Kalangan Usahawan Bumiputera Perusahaan Kecil Dan Sederhana (Pks) Di Besut, Terengganu. *Jurnal Kemanusiaan*, 18(2).
- [83] Zakaria, N., Zainal, S. R. M., & Nasurdin, A. M. (2011). Investigating the role of human resource management practices on the performance of SME: A conceptual framework. *Journal of global management*, 3(1), 74-92.
- [84] Zaman, M. M. K., & Othman, N. (2019). Amalan Pengurusan Perniagaan Usahawan Wanita Felcra Berhad. *Journal of Business Innovation*, 3(1), 1.
- [85] Zulkarnain, E. F., Abdullah, A. N. F. B., & Ridzuan, A. A. (2019). Tahap Pendigitalan Perniagaan Dalam Kalangan Usahawan PKS MARA di Melaka. *Politeknik & Kolej Komuniti Journal of Life Long Learning*, 3(1), 130-140.

A Framework for Protecting Teenagers from Cyber Crimes and Cyberbullying

Sultan Saud Alanazi¹, Adwan Alownie Alanazi²
College of Computer Science and Engineering
University of Ha'il, Ha'il
Saudi Arabia

Abstract—Social applications consist of powerful tools that allow people to connect and interact with each other. However, its negative use cannot be ignored. Cyberbullying is a new and serious Internet problem. Cyberbullying is one of the most common risks for teenagers to go online. More than half of young people report that they do not tell their parents when this will occur, this can have significant physiological consequences. Cyberbullying involves the deliberate use of digital media on the Internet to convey false or embarrassing information about others. Therefore, this article provides a way to detect cyberbullying in social media applications for parents. The purpose of our work is to develop an architectural model for identifying and measuring the state of Cyberbullying faced by children on social media applications. For parents, this will be a good tool for monitoring their children without invading their privacy. Finally, some interesting open-ended questions were raised, suggesting promising ideas for starting new research in this new field.

Keywords—Cyberbullying; cyber bullying; internet crimes; social media security; e-crimes

I. INTRODUCTION

Due to increased use of social media and trending social applications, children are involving in different internet-based activities. The ratio of increase in the use of social media is shown in Fig. 1. On one side it is proving to be an informative and interactive medium for indolence's but on the other side it has a lot of emerging issues as well. One of the biggest concerns is towards the increasing cyberbullying cases day by day. As the technologies are increasing day by day and many of the parents are not well aware about it, more children are curious towards using internet-based facilities which led them to cyberbullying victimization [1][2]. Cyberbullying effects individual's life mentally and physically. Some cases of cyberbullying lead to death as well. So, the motivation of this article is to educate parents to monitor their children online activates, make them aware of such increasing cyberbullying crime and to provide their children an E- safe environment [1].

In recent years, the continuous development of technology has made it possible to keep in touch with friends and family around the world anytime, anywhere, making many conveniences and ease of life. With its increasing use we must focus on its uses to help us in our daily lives. As human is a social animal, we need to build a social network to survive [3][4]. In this modern era, social networking is most important. As a result, social networking sites are now widely used for a variety of purposes, including entertainment and networking. All social media platforms require the consent of

all participants forming a network to communicate [4].

Benefits of this include the ability to acquire, deploy, maintain, and enhance the software used by consumers quickly and effectively. As a result, the number of new applications for platforms such as Mac, Android and iPhone have increased, and this model has been adopted to provide consumers with a variety of modern, low-cost applications. However, this paradigm shift has created a series of new security challenges [6].

As the market for new applications grows rapidly, so do the security threats with this software provisioning model. In the Android market, there are many reports of applications infected with malware and spyware [6].

As the use of cell phones increases, so does the misuse of information. Organized criminal groups may use services such as the WhatsApp mobile messaging application to engage in illegal activity, and the encryption technology used in these applications [7] is helping to cover the traces of this behaviour and make it undetectable to trace [8]. The growth of digital technology is mainly focused on the virtual life of individuals (mainly adolescents (11-19 years)) [9]. Internet users have a great opportunity to exchange ideas, interact and engage with people by developing a virtual community [1][2][10]. Often these interactions lead to passive use rather than active use. Cyberbullying is one of the most dangerous incidents in social media applications like WhatsApp and Facebook, and individuals (especially teens) are victims of cyberbullying [11]. Cyber security is a major concern of all security companies around the globe.

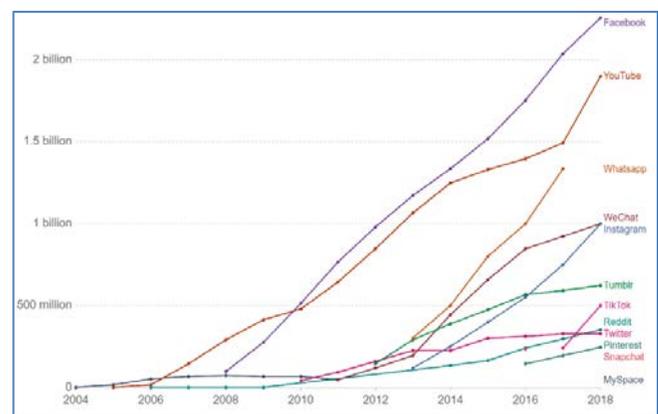


Fig. 1. Number of People using Social Media Platforms, 2004 to 2018 [5].

II. CYBERBULLYING

In recent years, many studies have been proposed to make the Internet a safer place to work. Security threats are on the rise, most notably cyber bullying and online grooming. Cyberbullying is a way to bully someone online. This can be done by maliciously posting or posting bad content online. This situation is disastrous for social media users, especially teenagers [11]. There have been many reported cases of cyberbullying as a cause of adolescent suicide. Online platforms like Facebook and WhatsApp are most commonly used by teens for all types of relationships as shown in Fig. 2. This social networking portal attracts more cybercriminals and engages in more illegal activities. As you know, billions of people around the world use it for many purposes [9][11].

On the other hand, online grooming presents you as dislike and uses victims to achieve their goals, usually through sexual activity [11]. As an open access platform, cyber-groomers also have access to the same internet platforms that are used by young people. Cyber bullying can take many forms, including posting harmful or threatening content on the Internet, spreading rumours, taking, and posting disgusting photos of someone without permission, posting pornographic images and information, and much more as shown in Fig. 3. Many teens suffer from cyber bullying and many of them engage in cyber bullying. More than a third of teens fell victim to cyberbullying, and they didn't even tell their parents about cyberbullying [11][12].

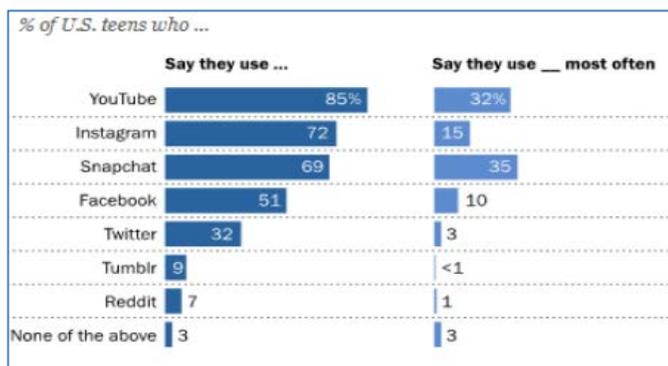


Fig. 2. Use of Social Media by Teens [9].

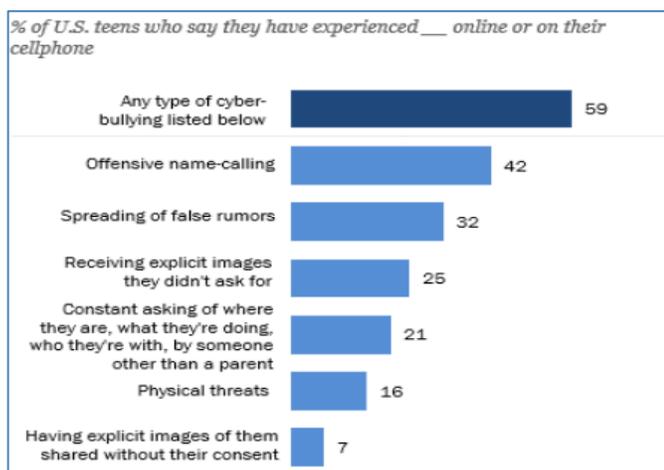


Fig. 3. Some Form of Cyberbullying [11].

The consequences of cyberbullying include anxiety, depression, and even suicide. Cyberbullying is primarily aimed at young people [11][13]. Cyberbullying begins when the victim begins to react to the harmful messages [13][14]. Internet use has been found to be a reliable predictor of victims experiencing cyberbullying. For men, Internet use is also an indicator of online cyberbullying [13][14].

III. DEADLY CHALLENGES

Now a day's cyberbullying is of great concern especially for teenagers. Parents need to check their children for such incidents which could reportedly took lives. Increasing social media use and growing apps have led to high level of curiosity among teenagers [13]. Due to this, teenagers try different games or interactive applications which led them to cyberbully victimization. One of these applications is the famous Blue Whale game. It is also known as Blue Whale Online Challenge, is a life-threatening online game that is believed to have killed hundreds of teenagers [15][16]. This online game is a great challenge for teenagers on various internet portals. Our dependence on technology is increasing. When curiosity leads teenage boys to online games, it can explain the Cyberbullying victimization very well. The boy faces a challenge online. He called the blue whale and agreed [15]. He gets a job, and he must do it. But is this task really important or problematic? These missions forced him to leak personal information used by criminals and began to threaten. The accident had a psychological impact on the victims, and sometimes even led to suicide. The game administrator has strict control over who can play or download. After research and analysis of the victims, they decide whom to grasp and who can be a very suitable and easy victim [15][16].

Social media cyberbullying can lead to suicides [17]. Another new WhatsApp group, MOMO CHALLENGE, was discovered by the popular YouTuber ReignBot in July 2018, and it has caught the attention of the public. After a few days, someone might accept the Peach Challenge. Momo information has been sent to countries such as Mexico, India, Argentina, the United States, France and Germany. In early August, WhatsApp Momo game reportedly killed two people, Manish Sarki (18) on August 20 and Aditi Goyal (age 26) in Kursong in Darjeeling the next day [18]. Police suspect they may have taken action because they are playing this suicide game.

After messaging a man named Momo on WhatsApp, a 12-year-old girl was found dead in a backyard near Buenos Aires, Argentina. "After searching for videos and chats on WhatsApp via a phone hack, they are now looking for a teenager who allegedly exchanged these messages with," police told the news portal [19]. Police believe the teenager's intention was to "upload the video to social media as part of a challenge, crediting the Momo game" for the suicide [19]. One of the challenges in this game is that people have to communicate with unknown numbers.

IV. MONITORING THE CHILD

With the growing technology parents are falling behind to cope up and understand the security threats of using internet. This is due to a lack of time and inadequate knowledge of

computers usage, social networking and rapid growing mobile apps they are unable to inspect what their children are doing when they are using internet. Parents are unaware or do not fully understand the evolution of bullying. Thus, it is necessary for parents of modern era to monitor their child.

Parents cannot control all of their children's behaviours, but there are steps they can take to prevent cyberbullying and protect their children from harm [20]:

- Track and monitor teenagers' social media, apps, and browsing history. Cyberbullying is usually occurring when parents notice negative changes in adolescent behaviour.
- View or reset the privacy settings and phone location (GPS) of your child.
- Follow your child on social media, and try to be in his friends list, or ask a trusted adult to do the same.
- Parents should pay attention and to be aware of the latest applications, social media platforms and digital languages used by teenagers.
- Know your child's social media email address, username, and password.
- Establish rules for good behaviour, digital content and applications.
- Limit internet use accordingly.

Following (Fig. 4) is a depiction of different stages to cyberbullying victimization.

Parents can use different available parental control tools and techniques that can help to build a non-invasive approach towards their teenage children and thus prevents them from cyberbullying, malicious behaviour on the internet, and inappropriate content. Free monitoring software and apps are available that can be used to help parents restrict inappropriate and unwanted content, block specific domains, and view children's online activities (including social media activities) without having to physically check their children's devices daily. In fact, it is also less invasive and sometimes invisible to children [21][22][23].

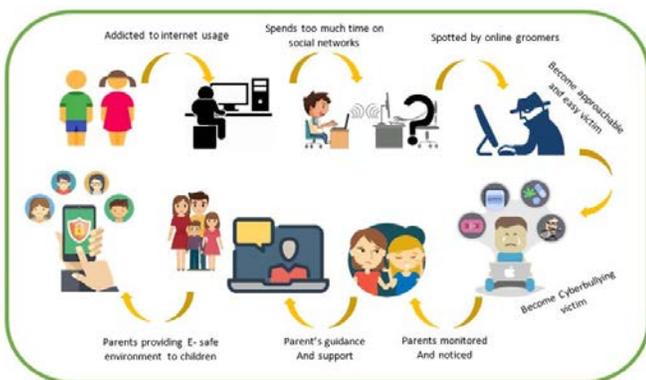


Fig. 4. Stages to Cyberbullying Victimization [13].

When choosing the right internet application for a child, parents should consider their children's ages, device usage, and digital network behaviour, which may help. Parents should build trust relationships with their children by engaging in open and honest discussions. These conversations provide an opportunity to exchange values and expectations regarding digital family-friendly behaviours, such as viewing and sharing content, and available and unavailable applications. Ask your child about their online experiences to avoid potential risks of cyberbullying and abuse. Make it clear to them that your goal is to protect them and that you want to have an open conversation with them. Discuss their concerns, opinions and also manifest your opinion [21][22][23].

V. CYBERBULLYING DETECTION FRAMEWORK

In this section, the proposed cyberbullying detection architecture is presented. The proposed architecture includes five different modules, which not only detects the bullying content on the social application but also helps in the detection of mobile security level used by the child. As the use of mobile phones is growing tremendously, various mobile phone applications are also emerging. Different Social media applications are emerging through which youth can communicate in many different ways. Popular example of these type of social applications includes, WhatsApp, Facebook, Line and many more. The proposed architecture will help parents to monitor the child and detect the bullying content before time. This helps to save child from becoming victims to such crimes online. The details of the architecture are shown in Fig. 5 and also discuss in the following sections.

A. Registration Module

This module is used to ask the user (parent) to enter the Application login information of his or her minor. This login information is used to retrieve the relevant data from the desired logged-in App. It also provides information of the application used by his or her minor.

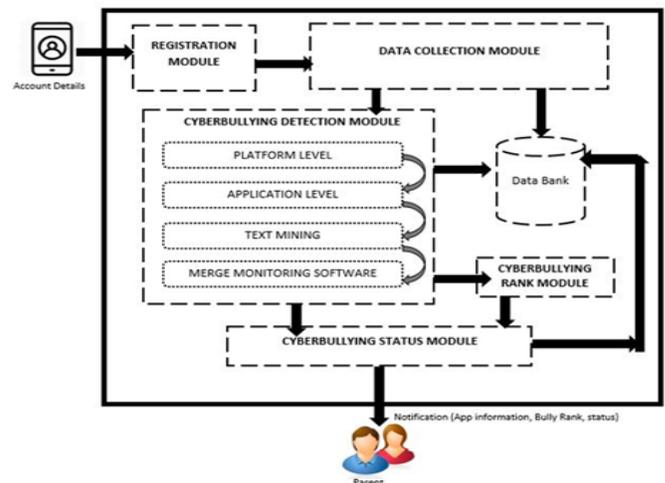


Fig. 5. Cyberbullying Detection Framework.

B. Data Collection Module

The Data Collection Module is the component that collects all the relevant information from the logged-in application for example:

- Contacts.
- Chat history.
- Backup of chat database.
- Avatars of contact.
- Received files.
- Sent files.
- Locations.

C. Data Bank

Stores all the relevant information regarding the particular topic and area of cyberbullying that is used by the application and helps to detect the bullying rank status further. It works as a repository of persistence information, keywords, chats, history, and any other relevant information based on the user requirements that can be used for reference for detection module.

D. Cyberbullying Detection Module

The Cyberbullying Detection Module uses the retrieved data to determine if a child or teenager is being bullied online. The main result of this component is the Bullying Rank (BR) calculated according to the state of art and specific vulnerability. Since we are focusing on Android apps, the following security settings will be checked accordingly.

E. Application Platform Level Security Check

A mobile application platform is a set of software tools used in the development, creation and maintenance of mobile applications. In part, it is an enterprise mobile application platform that provides organizations with mobile application tools. It is also vulnerable to many security threats. The continued emergence of mobile malware not only provides knowledge and copyright for developers, but also leads to information leakage and even financial loss. The security requirements of mobile applications have become a serious issue in the development of the entire application market. A tabularized Platform Level Security Comparison is shown in the Table I below.

TABLE I. PLATFORM LEVEL SECURITY COMPARISON

Feature	Android	iOS
App Download	Crowdsourcing	iOS App store only
Signing Technology	Self-Signing	Code signing
Inter-process Communication	Yes	No
Open/ Close Source	Open source	Close
Memory Randomization	Yes	Yes
Storage	Internal & External	Internal only
Shared User ID	Yes	No

F. Application-Level Security Check

Application-level security refers to the security services that are invoked on the interface between the application and the queue manager connected to it. A tabularized Comparison is shown in Table II.

TABLE II. APPLICATION-LEVEL SECURITY CHECK

Feature	Found (✓)/ Not Found (x)
Shared User ID	✓
No. of permissions	✓
Application Verification	✓
Encrypted Messages	✓
Decrypted Messages	✓
Contacts	✓
Media	✓
Location	✓

G. Text Mining Check

Here, we're going to use two sets of data. It uses the Bad Words Dataset and Sensitive Word Dataset to categorize the user's message and record the bad history in the database. If a user exceeds the limit for inappropriate posts or adult photos, they will be automatically blocked. Some abusive words are given in the Table III.

TABLE III. ABUSIVE WORDS

ABUSIVE WORDS
BAD WORDS (Retarded, Dumb)
SWEAR WORDS (Bit**)
SECOND PERSON (You, Yours)

H. Merge Monitoring Software

Implementing monitoring software will help to understand the big picture. There are different types of monitoring software available (paid or free), including Spyzie and net nanny. Spyzie is a professional monitoring solutions provider that provides tracking and surveillance tools for smartphone users [24]. It is used to protect children from online threats and to ensure that they are always in a safe place. Parents will be able to track and store call logs, messages, social activities, etc.

I. Cyberbullying Rank Module

Rank the cyberbullying according to the following scale:

No.	RANK	RANGE
1	Low	0-3
2	Medium	4-6
3	High	7-10

The ranking will be based on the severity of the words which affects the mentality and behaviour of the teenagers and attract other teenagers and public opinion of the circle. Keywords will low affect will have less effect and thus will be

ranked low accordingly. For example, sexting and related keywords will be considered as high severity, posting embarrassing photos, stealing identity of a person will be considered high ranking but keywords related to beauty, body, may be considered as low rank or medium rank.

VI. ADVICE FOR PARENTS

To reduce the harms of cyberbullying and online behaviour, parents can:

- Set clear expectations about their child online reputation and behaviour.
- Make their children aware about the negative consequences of cyberbullying, make annoying and inappropriate comments and messages, share nude photos of you and other people, sexting etc. and the legal issues associated with all this.
- Be clear and strict about what can be viewed and what can be shared.
- Determine which apps are right and allowed to use for your child and which are not.
- Set rules for how your child spends time online or on the device.
- Simulate positive and respectful online behaviour to him through your devices and online and social media accounts.
- Talk to your child about the bad effects of being a bystander of cyberbullying.

Talking to your children about cyberbullying and digital behaviour is not a one-off activity, but a continuous conversation. Before your child switches to text messaging, social media, online games, and chatting, discuss these topics with them. It can help them better understand the reality and potential bad effects of cyberbullying and provide them with ongoing opportunities to practice the coping skills. In this way, you can support the transition from a bystander to ally. SAMHSA provides a free 'Learn about Bullying' app for parents, teachers and educators with dialogue options, tips and other tools to help prevent bullying [25].

If you think your child has witnessed cyberbullying, you can encourage them to do something they should and shouldn't. For example, encourage children to dislike or not like, share, or comment on someone's racial, or hurtful text or posted information, or send hurtful text to others. Without sharing and participating on the malicious or cyberbullying content, you can limit the harm the message can do to others or to yourself.

When the children feel they should react, encourage them to respond calmly, clearly, and constructively. Anger and violent reactions to a situation can make the situation worse. Encourage children (and adults!) To stay away from the device and not accuse, insult, or retaliate against others. It helps you stay calm and focused and shows that the digital behaviour of others is harmful and unacceptable.

Please reply directly to the person who sent the offending message or comment. If you feel safe, you can contact each other online, by phone, or directly with people creating or transmitting malicious messages. This way you can clearly show that they can't stand the negative behaviour. This is also an opportunity to honestly share concerns about the behaviour and its causes.

VII. CONCLUSION

This section concludes this article by identifying a complex set of open-ended research questions and compelling solutions. Cyber bullying is an emerging concern with growing technology. Main victims of cyberbullying are teenagers who are bullied online. Teenagers are emotionally weak and fall an easy prey. In this era to provide an e safe environment to children we proposed an architecture to detect cyberbully, rank it and provide information to the concerned parent.

In future, we will develop a prototype of this framework for general purpose use for both Android and iOS.

VIII. NOTES ON CONTRIBUTORS

Sultan Saud Alanazi completed his bachelor's in Information Technology from Deakin University in 2019. He is currently working for Al-Shamli General Hospital, Saudi Arabia as an IT admin. He is also doing his Master's in Cybersecurity at University of Ha'il. His area of interests includes Application security and Network security.

Dr. Adwan Alownie Alanazi received the PhD degree in Computer Science and engineering from the University of Bridgeport, USA. He is currently holding an assistant professor position with the University of Hail, Saudi Arabia. His research interests include Internet of things (IoT) and security in wireless Sensor networks.

REFERENCES

- [1] Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(4), 376–385. <https://doi.org/10.1111/j.1469-7610.2007.01846.x>.
- [2] Langos, C. (2012). Cyberbullying: The Challenge to Define. *Cyberpsychology, Behavior, and Social Networking*, 15(6), 285–289. <https://doi.org/10.1089/cyber.2011.0588>.
- [3] O'Keeffe, G. S., & Clarke-Pearson, K. (2011). The Impact of social media on Children, Adolescents, and Families. *Pediatrics*, 127(4), 800–804. <https://doi.org/10.1542/peds.2011-0054>.
- [4] Khan, G. F., Swar, B., & Lee, S. K. (2014). Social Media Risks and Benefits. *Social Science Computer Review*, 32(5), 606–627. <https://doi.org/10.1177/08944393145247011>.
- [5] Ortiz-Ospina, E. (2019, September 18). The rise of social media. Retrieved from Our World in Data website: <https://ourworldindata.org/rise-of-social-media>.
- [6] Jang-Jaccard, J., & Nepal, S. (2014). A survey of emerging threats in cybersecurity. *Journal of Computer and System Sciences*, 80(5), 973–993. <https://doi.org/10.1016/j.jcss.2014.02.005>.
- [7] Madden, D., & Deb, A. (2019, December 16). WhatsApp is a Threat to Society. Here's How to Fix It. Retrieved August 16, 2020, from Techonomy website: <https://techonomy.com/2019/12/whatsapp-is-a-threat-to-society-heres-how-to-fix-it/>.
- [8] Bernal, N. (2018, July 4). WhatsApp admits encryption has helped criminals as it plans crackdown on illegal behaviour. The Telegraph. Retrieved from <https://www.telegraph.co.uk/technology/2018/07/04/whatsapp-admits-encryption-has-helped-criminals-plans-crackdown/>.

- [9] Anderson, M., & Jiang, J. (2018, May 31). Teens, Social Media & Technology 2018. Retrieved from Pew Research Center: Internet, Science & Tech website: <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>.
- [10] Rivera-Vargas, P., & Miño-Puigcercós, R. (2018). Young people and virtual communities. New ways of learning and of social participation in the digital society. *Páginas de Educación*, 11(1), 67. <https://doi.org/10.22235/pe.v11i1.1554>.
- [11] Anderson, M. (2018, September 27). A Majority of Teens Have Experienced Some Form of Cyberbullying. Retrieved from Pew Research Center: Internet, Science & Tech website: <https://www.pewresearch.org/internet/2018/09/27/a-majority-of-teens-have-experienced-some-form-of-cyberbullying/>.
- [12] Whittaker, E., & Kowalski, R. M. (2014). Cyberbullying Via Social Media. *Journal of School Violence*, 14(1), 11–29. <https://doi.org/10.1080/15388220.2014.949377>.
- [13] Chan, T. K. H., Cheung, C. M. K., & Lee, Z. W. Y. (2021). Cyberbullying on social networking sites: A literature review and future research directions. *Information & Management*, 58(2), 103411. <https://doi.org/10.1016/j.im.2020.103411>.
- [14] Patchin, J. W. (2015, February). Our latest research on cyberbullying among school students. Retrieved April 8, 2021, from Cyberbullying Research Center website: <http://cyberbullying.org/2015-data/>.
- [15] Mahadevaiah, M., & Nayak, R. B. (2018). Blue Whale Challenge: Perceptions of First Responders in Medical Profession. *Indian Journal of Psychological Medicine*, 40(2), 178–182. https://doi.org/10.4103/ijpsym.ijpsym_399_177.
- [16] Narayan, R., Das, B., Das, S., & Bhandari, S. S. (2019). The depressed boy who accepted “Blue Whale Challenge.” *Indian Journal of Psychiatry*, 61(1), 105–106. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_234_18.
- [17] Luxton, D. D., June, J. D., & Fairall, J. M. (2012). Social Media and Suicide: A Public Health Perspective. *American Journal of Public Health*, 102(S2), S195–S200. <https://doi.org/10.2105/ajph.2011.300608>.
- [18] NDTV. (2018, August 26). After Two Deaths, Bengal Grapples with Suicide Game “Momo Challenge.” Retrieved April 8, 2021, from NDTV.com website: <https://www.ndtv.com/india-news/after-two-deaths-bengal-grapples-with-suicide-game-momo-challenge-1906357>.
- [19] Chiu, A. (2018, September 5). The “Momo Challenge”: A sinister threat to young people or an urban myth?. *Washington Post*. Retrieved from <https://www.washingtonpost.com/news/morning-mix/wp/2018/09/05/the-momo-challenge-a-sinister-threat-to-young-people-or-an-urban-myth/>.
- [20] Ybarra, M. L., & Mitchell, K. J. (2004). Online aggressor/targets, aggressors, and targets: a comparison of associated youth characteristics. *Journal of Child Psychology and Psychiatry*, 45(7), 1308–1316. <https://doi.org/10.1111/j.1469-7610.2004.00328.x>.
- [21] Elsaesser, C., Russell, B., Ohannessian, C. M., & Patton, D. (2017). Parenting in a digital age: A review of parents’ role in preventing adolescent cyberbullying. *Aggression and Violent Behavior*, 35, 62–72. <https://doi.org/10.1016/j.avb.2017.06.004>.
- [22] Mehari, K. R., Moore, W., Waasdorp, T. E., Varney, O., Berg, K., & Leff, S. S. (2018). Cyberbullying prevention: Insight and recommendations from youths, parents, and paediatricians. *Child: Care, Health and Development*, 44(4), 616–622. <https://doi.org/10.1111/cch.12569>.
- [23] UNICEF. (2020). Cyberbullying: What is it and how to stop it. Retrieved from www.unicef.org website: <https://www.unicef.org/end-violence/how-to-stop-cyberbullying>.
- [24] Spyzie. (2020). Spyzie. Retrieved from [Spyzie.com](http://www.spyzie.com) website: <https://www.spyzie.com/>.
- [25] SAMHSA. (n.d.). KnowBullying Mobile App | Publications and Digital Products. Retrieved from store.samhsa.gov website: <https://store.samhsa.gov/product/knowbullying>.

Anonymity Feature in Android Mobile Apps for Interest Groups

Lee Sin Yi¹, Nurul A. Emran², Norharyati Harum³

Fakulti Teknologi Maklumat dan Komunikasi,
Universiti Teknikal Malaysia Melaka,
Malaysia

Abstract—Mobile apps that provide platforms for interest-oriented communities (or interest groups) allow people with common interests to gather virtually in sharing their shared passion and ideas. While the participation of such apps is voluntary, some people feel uncomfortable revealing their real identities due to privacy and safety concerns. Thus, some mobile apps provide an anonymity feature that allows people to join the group anonymously. Nevertheless, little is known on how the anonymity feature relates to the people who prefer to join interest groups. Thus in this paper, we hypothesize that mobile apps users that highly value interest groups will also highly appreciate the anonymity feature provided in the interest groups. In particular, we explored the market segment of mobile Android apps with anonymity features within selected interest groups. A pilot study was conducted where 34 Android apps users, primarily Malaysian, filled up the questionnaires designed to investigate the anonymity feature in the apps. The results of the pilot study show that most Android apps nowadays offer their users to remain anonymous. The findings show that most users who give a high score on the importance of interest-based groups also provide a high score on the importance of the anonymity feature offered by the mobile apps providers.

Keywords—Anonymity feature; mobile social apps; quantitative observation; android apps; anonymous social media; interest groups component

I. INTRODUCTION

Online communities are often connected to common interest communities voluntarily. The advantage of online communities is that there will be no face-to-face or physical access requirements. With the interest-based groups that can be accessed through mobile apps, more people can easily find new friends with similar interests. The so-called social apps are not limited to the young generation and have become more popular among the elders [1]. For example, people who are fond of hiking can find other people with similar interests in the hiking-interest community from mobile apps.

Furthermore, with the proliferation of interest-oriented mobile apps such as Meetup, Smactive, WeGoDo, the mobile apps users community can connect with people with a similar passion more than ever before [2]. Major social media apps platforms like Facebook, Whatsapp, and Telegram also provide accessible platforms for interest group creations. Martin, in 2019 reported that marketers could use interest-based groups in Facebook to boost their product sales against the targeted audience [3], which is an indication of the potential of interest-based groups from a marketing perspective.

Facebook keeps profiles of its users' interests based on users' behavior, such as the pages a user likes, the previous ads they clicked, the websites and posts they engaged. For example, we can see various interest-based categories and subcategories if we utilize Facebook's ads manager platform. An 'interest' category like Hobbies and activities has several subcategories such as Arts and music, Home and garden (Furniture, Gardening, Do it yourself (DIY)).

Within the interest groups, the members usually seek to exchange information or ideas on specific topics or get solutions for their problems. Participation in an interest group can be entertaining, compelling and people often return and stay active for an extended period. Most of the time, these people cannot be defined by a specific geographical area. Without knowing each other's real identity, everyone will have equal rights and freedom to express their minds. Self-disclosure can help in releasing stress and pressure, satisfying needs, and adapting communications behaviors [4].

As the identity of the members is usually being disclosed, there is a rising concern about the privacy and confidentiality of the members. Thus, the anonymity topic has become a focus in many application domains [3]–[8], and a new type of social media apps called anonymous social media has emerged [9]. Anonymous apps (such as Secret and Yik Yak) gain popularity because many people want to express themselves freely [10]. Such apps also become a 'tree hole' (confidant) where people can share their problems anonymously. The actual name and information of the app users will not be made compulsory during the registration process. In this way, people will share the problems that they do not want their family and close friends to know. Help or guidance from the other members can be received without the pressure of being judged by the people who do not know them personally. People with low self-esteem especially suffer the most when it comes to identity disclosure [11]. Besides, in research conducted by Ma, Hancock, and M. Naaman (2016), they found that people tend to minimize self-disclose as content intimacy increases [4].

In online platforms, people usually tend to express themselves more openly. They can share the things that they will not share in the face-to-face world as they feel more uninhibited in cyberspace. From the psychological perspective, this is being called the disinhibition effect by the researchers [12]. People tend to align their virtual identity with their self-guide when reconstructing a new identity in an anonymous application. The self-discrepancy between their self-guided and

This work is sponsored by the Universiti Teknikal Malaysia Melaka under research grant PJP/2020/FTMK/PP/S01771

virtual identity will be lesser as they will show more true selves with the latest online identity [12].

They may be more satisfied and motivated to show their true self freely without much constraint in an anonymous community online. According to Suler (2004), when people tend to show more generosity and kindness without their real identity being known by people, this is called benign disinhibition. However, there may be toxic disinhibition in the form of rude language or harsh criticisms [13], which lead to cyberbully. Denzil et al. (2015) divide anonymity based on the sensitivity level of the social media contents where anonymity is beyond the binary concepts of anonymous or non-anonymous types [9].

As there is a call for mobile apps that subscribe for anonymity feature, in this paper, we are motivated to answer the following research questions:

a) What is the proportion of the market segment for mobile apps with anonymity features in interest groups understudy?

b) How perception of the importance of interest group relates to the perception of the anonymity feature?

In the next section, related work on the anonymity feature will be presented. Section III consists of the methods used to answer the research questions; Section IV covers the findings and discussion. Finally, Section V concludes the research.

II. RELATED WORK

There is a 13.4% increase in the number of social app downloads from the Google Play Store from the year 2019 to 2020, which makes 1.18 billion downloads in 2020 [14]. On average, people spent 131 minutes every week on the Social app, which is the most extended amount of time spent within the mobile apps category. Gaming, Communication, and Music are also among the top apps category in time spent statistics. Facebook is the most popular social media platform worldwide, with 2.6 billion users, as of July 2020, followed by YouTube and Whatsapp [14].

A survey conducted from September 2016 to January 2021 revealed that 86 percent of Malaysians were actively engaged with social media. Facebook was the most popular social media platform of choice among Malaysian, followed by Instagram, Facebook Messenger, and LinkedIn [15].

The popularity of Facebook might be driven by its flexibility in creating and joining interest-based groups. In Facebook, a Facebook Group feature allows the creation of interest-based groups for its members to communicate, express their minds, and share their mutual interests or passion [16][17]. Facebook allows Facebook users to join up to 6000 groups at the same time. One requirement set by Facebook is the real-name policy, where Facebook users must use their real name to register a Facebook account and configure the user's profile. This is a compulsory requirement for all Facebook users. Through this method, the real-name policy will ensure the safety of the group members as one will know about the identity of the persons they are connecting with.

A "real name" refers to the actual name in this policy as it would be shown on the identity card, student ID, driving license, or credit card of Facebook users. This naming policy of Facebook has prohibited names that are prone to judgemental elements, such as names that contain too many words, excessive capital letters, or first names that have initials. The Facebook accounts that have such names will be detected and thus suspended by the monitor software of Facebook [18]. However, the real-name policy maintained by Facebook has raised safety concerns on the privacy and data protection of Facebook users. This situation is due to the reason that one's real name often reflects some identities and cultures. The real name of some people may also contain many potentially sensitive details that might be obtained by people with bad intentions while browsing their Facebook profiles [19]. With the rising privacy demands in the surveillance society, some Facebook users started to protest against this real-name policy which causes Facebook to relax the real names requirement to "provide the name they use in real life" [20]. Despite the real name issue, some argue that real name is crucial to establish the trustworthiness of the apps [21].

Nevertheless, as privacy and safety issues have received more weight in interest-based groups, the anonymity feature becomes more attractive than trustworthiness. Socializing within the interest-based groups that do not require their members to provide their real identities is a desired characteristic of a social app for some people. Besides, due to the disinhibition effect of being anonymous, people tend to loosen up and share personal things about themselves. They may reveal their wishes, fears, or secret emotions when they can separate what they express from their real identity and their natural world. Anonymity gives them more courage to say and speak more openly. Everyone has an equal opportunity to express their mind regardless of their race, gender, or status in the real world [22]. Privacy and secrecy issues also have been highlighted in mobile apps/systems in education and business domains [23], [24]. In 2017 a survey on awareness of privacy has been conducted among Android users, which revealed that the level of understanding is still low [25] even though there is a raising need for privacy.

Caution must be taken as the true negative self that is usually hidden in the physical world may no longer be hidden in anonymous social media platforms [12]. Should there be any user with negative minds and intentions, they can quickly spread inappropriate or unhealthy information. Thus, a feedback or reporting feature should be added to this apps to lodge a report about any inappropriate behavior of its members. As little is known on the need of reporting feature and on how perception of the importance of interest group relates to the perception of the anonymity feature these are the research gaps addressed by this paper.

III. METHODOLOGY

This section will describe how we answer the research questions as stated earlier in Section I. An explanatory approach has been adopted for market segment analysis to observe the anonymity feature in Android apps worldwide. In particular, we adopted the quantitative observation method

which is usually adopted in studying market segments for marketing purposes such as in [26], [27].

A quantitative approach has been adopted in a pilot study to determine the relationship between the interest-based and anonymity features, especially among Malaysian users. Descriptive statistical analysis is adopted to analyze the result of questionnaire responses. IBM SPSS software, version 2, was used to perform the statistical analysis and for data screening (error and missing values checking). A pilot study is a part of our full-scale project that aims to determine the potential of anonymity features in mobile apps development for Malaysian interest-based groups.

A. Research on Market Segment

To answer the first research question, we use Google Play Store to find inter-interest-based group apps on the Android platform. As we set English as the apps' language, apps in other languages such as Chinese, Japanese, and Korean are excluded. The apps with invitation links and software bugs are also excluded from this study as they cannot be accessed quickly during the research phase.

After filtering all of the apps, we further analyze them by conducting a member registration process. Unlike the mobile apps reported in [28], not all app providers reveal the anonymity feature in their apps. Thus, registration of a new member account is performed. This process is crucial to check the anonymity features within the apps by experiencing an actual registration process. In this study, we perform binary classification. Apps requiring a real name or actual information (such as the mobile phone number) are classified as apps without anonymity. In contrast, apps allow the member to be anonymous as apps with anonymity.

We also observed the criteria of whether the app offers a single-interest group or multiple-interest groups. Apps that serve one main topic but with multiple groups are classified as multiple-interest groups. For example, an app focuses on drawing art as the main topic with multiple groups such as the water-color community, oil painting community, and sketch community. A market segment analysis of this quantitative observation study will be reported in Section IV.

B. Questionnaire Development

A questionnaire has been designed based on samples of best practices of online questionnaires related to mobile apps usability and UX [29]–[31]. The questionnaire is divided into three categories: demographic, mobile user experience, and preference. The demographic category consists of three items, namely gender, age, and nationality; mobile usage experience has four items where respondents will tell whether they have experience with interest group apps and the anonymity feature. The preference category consists of nine items. The first two categories consist of close-ended questions, while the third category has five closed-ended questions (Yes/No and five-scale Likert-chart questions) and four open-ended questions. The questionnaire is developed using an online Google form, as shown in Fig. 1.

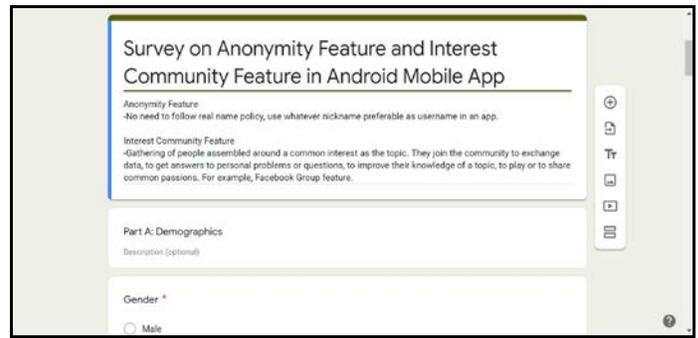


Fig. 1. The Online Questionnaire.

We utilize Facebook's interest-based group to distribute the questionnaire as it is the most popular social media platform to date. In particular, we use a public Facebook group called "Thesis / Survey Questionnaire Filling Group" with 14.0K members and a private group called "Student Survey Exchange" with 13.8K members. Both groups facilitate their member worldwide in getting the respondents for research surveys that are based on mutual collaboration and voluntary basis. The members mainly gather in this group to collect responses for their surveys. To exchange for the responses from other members, they are required to help other group members to fill up their questionnaires or do their surveys. A call for the survey is made in the group posting, as shown in Fig. 2. Even though the focus of the pilot study is Malaysian users, we also record responses from other countries that took part in the survey. We set to observe if non-Malaysians show common or different characteristics from Malaysians. In this pilot study, we allow the respondents to answer the questionnaire within seven days before collecting the results.

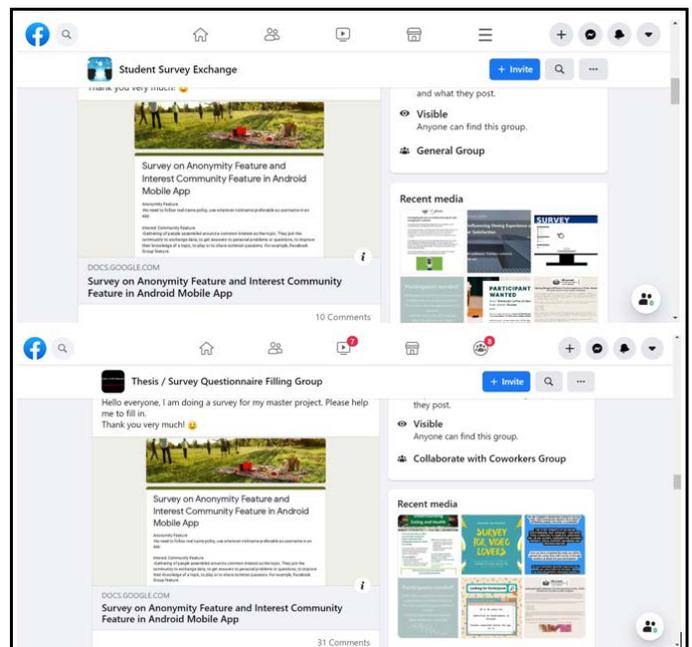


Fig. 2. Posting on the Call for the Survey.

IV. RESULTS AND DISCUSSION

A. Market Segment Analysis

In this section, the results of market segment analysis using descriptive statistics will be presented. A set of interest-based Android apps (n=65) have been selected from the Google Play Store, as shown in the Appendix. These apps belong to interest group categories such as social, gaming, communication, and art.

The top pie chart in Fig. 3(a) shows that apps that offer anonymous features dominate the Android apps market segment by 69%. In contrast, Fig. 3(b) shows that most of the apps under study offers multiple interest groups, with 68% of the market segment.

To analyze further, we seek to examine the number of apps with (and without) anonymity features by interest group types. As shown in Fig. 4, it has been found that the apps with the anonymity feature offer more multiple interest group creation than the apps without the feature, with 43%. Only 23 % of the apps without anonymity feature offer multiple interest group creation.

Based on the results presented so far, we can see that most Android apps provider are ready to give options for their users to be anonymous or not. As a result, these kinds of apps dominate the Android apps market segment. The results also indicate the popularity of multiple interest group creation within a single app regardless of whether it comes along the anonymity feature or not.

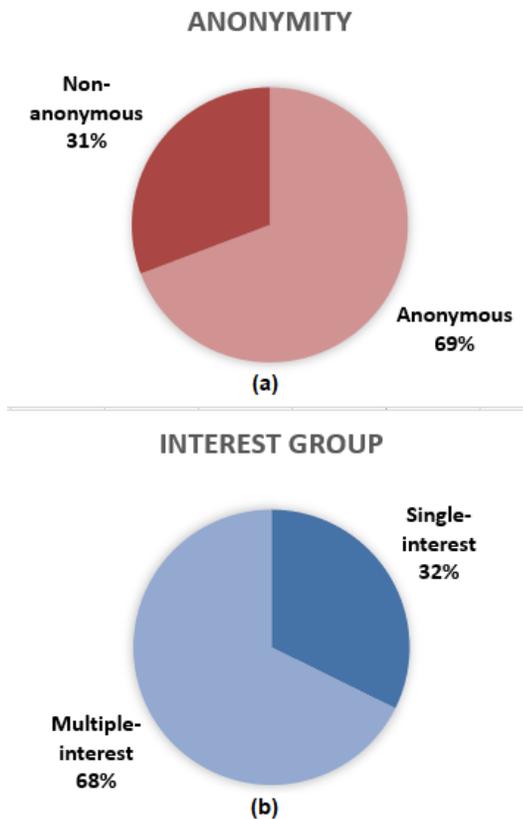


Fig. 3. (a)The Proportion of Android Apps based on Anonymity, (b) the Proportion of Android Apps based on the Type of Interest Group.

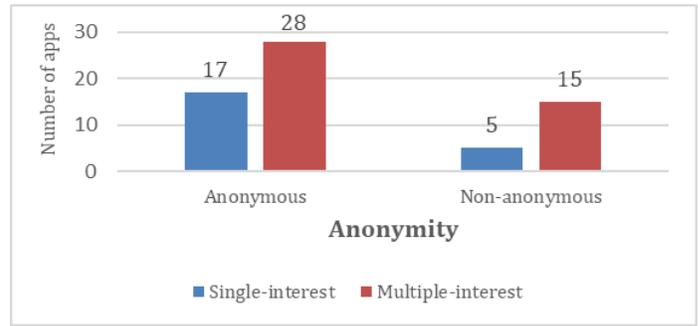


Fig. 4. The Proportion of the Interest Group Type by the Anonymity Type.

B. Close-Ended Question Responses Analysis

Data collection was facilitated via Facebook interest groups, as mentioned earlier in Section 3.0, where 34 responses (n=34) were received from Android app users. Table I shows the demographic information of the participants.

Table II consists of the results retrieved from the Experience category items that cover questions regarding the experience of users with interest groups and anonymity features. The results show that, in terms of age group, those within 18-24 years old are more familiar with interest groups app and apps with anonymity features as compared to those from 25-34 age category. We can say that respondents who use these apps are among the early adults based on their age range. It is also found that female dominates both, apps with interest groups and with anonymity features. 44% of Malaysian have experience with group interest apps, and 41.2 % of them also used apps with anonymity features before. If we accumulate the scores for non-Malaysian, we can also see most of the respondents already exposed to the group interest apps and apps with anonymity features.

TABLE I. DEMOGRAPHIC INFORMATION OF THE PARTICIPANTS

Characteristic	Value
Age (years), n (%)	
18-24	19 (55.9)
25-34	15 (44.1)
Gender, n (%)	
Female	21 (61.8)
Male	13 (38.2)
Country, n (%)	
Austria	1 (2.9)
British	2 (5.9)
Brunei	1 (2.9)
Egyptian	1 (2.9)
Indian	3 (8.8)
Kenyan	1 (2.9)
Malaysian	16 (47.1)
Maltese	1 (2.9)
Pakistani	5 (14.7)
Singaporean	2 (5.9)
Swiss	1 (2.9)

TABLE II. THE RESULTS OF CLOSED-ENDED ITEMS IN THE EXPERIENCE CATEGORY

Characteristic	Have you used apps with interest groups before? (%)		Have you used apps with anonymity features before? (%)	
	Yes	No	Yes	No
Age (years)				
18–24	44.1	11.8	47.1	8.8
25–34	41.2	2.9	41.2	2.9
Gender				
Female	50.0	11.8	52.9	8.8
Male	35.3	2.9	35.3	2.9
Country				
Austria	2.9	0.0	2.9	0.0
British	5.9	0.0	2.9	2.9
Brunei	2.9	0.0	2.9	0.0
Egyptian	2.9	0.0	2.9	0.0
Indian	5.9	2.9	8.8	0.0
Kenyan	2.9	0.0	2.9	0.0
Malaysian	44.1	2.9	41.2	5.9
Maltese	0.0	2.9	0.0	2.9
Pakistani	11.8	2.9	14.7	0.0
Singaporean	2.9	2.9	5.9	0.0
Swiss	2.9	0.0	2.9	0.0

Table III consists of the results retrieved from the closed-ended question items in the Preference category. In this category, respondents state their preference regarding apps with an interest group and anonymity feature and the type of interest group. The results show that most respondents prefer apps with interest groups and anonymity features, in which female users and those from 18-24 years old dominate the score. Respondents from the same categories also show the highest preference on multiple interest groups than the single-interest group. The findings also reveal that most Malaysian prefer apps with interest groups and anonymity features. The same behavior also can be seen among non-Malaysian. Nevertheless, in terms of groups, Malaysians equally prefer the single and multiple interest groups, while non-Malaysians show more interest in the multiple groups' apps.

To examine the relationship between one's perception of the importance of apps with interest groups and anonymity features, we extract the responses from the five-scale Likert-chart questions in the Preference category. The questions require the respondents to rate the importance of apps with interest groups and anonymity features from 1 (the least important) to 5 (the most important).

Fig. 5 illustrates the Scatterplot graph that shows the distribution of users based on the rating on the importance of interest groups and anonymity features. If we divide the chart equally into four quadrants, we can see that most users are grouped at the top right quadrant, where a high rating was given to the importance of both interest and anonymity features. However, four respondents who value the importance of interest groups do not have the same opinion on the

importance of the anonymity feature (11.76%). Only one respondent who gives a high rating on the anonymity feature gives a low rating on the importance of the interest group. As the left bottom quadrant is empty, we can say none of the respondents think less about the importance of the interest group and the anonymity features.

TABLE III. THE RESULTS OF CLOSED-ENDED ITEMS IN THE PREFERENCE CATEGORY

Characteristic	Do you prefer apps with an interest group?		Do you prefer apps with an anonymity feature?		Type of preferred interest group apps	
	Yes	No	Yes	No	Single	Multiple
Age (years)						
18–24	50.0	5.9	44.1	17.6	14.7	41.2
25–34	41.2	2.9	38.2	0.0	20.6	23.5
Gender						
Female	55.9	5.9	50.0	11.8	14.7	47.1
Male	35.3	2.9	32.4	0.0	20.6	17.6
Country						
Austria	0.0	2.9	0.0	2.9	0.0	2.9
British	5.9	0.0	2.9	2.9	2.9	2.9
Brunei	2.9	0.0	2.9	0.0	0.0	2.9
Egyptian	2.9	0.0	0.0	2.9	0.0	2.9
Indian	8.8	0.0	8.8	0.0	0.0	8.8
Kenyan	2.9	0.0	2.9	0.0	0.0	2.9
Malaysian	44.1	2.9	44.1	2.9	23.5	23.5
Maltese	0.0	2.9	2.9	0.0	2.9	0.0
Pakistani	14.7	0.0	8.8	5.9	8.8	5.9
Singaporean	5.9	0.0	5.9	0.0	0.0	5.9
Swiss	2.9	0.0	2.9	0.0	0.0	2.9

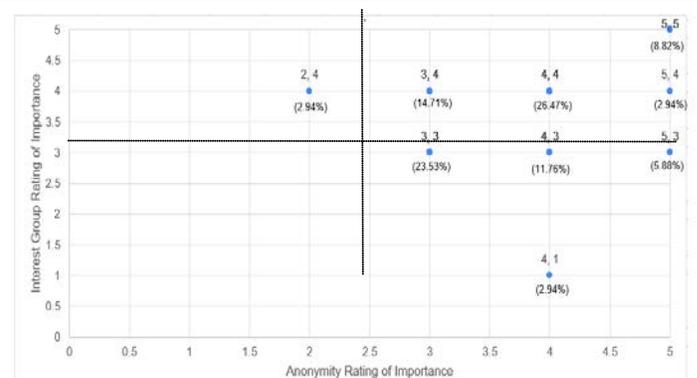


Fig. 5. The Distribution of Users on the Rating on the Importance of Interest Groups and Anonymity Features.

C. Open-Ended Question Responses Analysis

The responses to the open-ended questions from the Preference category allow the participant to express their opinion regarding the rating they give on the importance of interest group and anonymity feature and the type of interest group they prefer. According to some of the responses collected, most participants think that the interest group is important because this feature increases the possibility of meeting people with similar interests and helps introverts discover people with the same interest. Some think that it is nice to post or communicate with others with similar topics and interests and get valuable and specific topic information. Some of the respondents who prefer the multiple-interest group over the single-interest group mentioned that the multiple-interest group allows more options, more choices, and more possibilities to find groups that match their interests. They think that the group will enable people to exchange ideas and thoughts about the shared passion and receive more information. One of the respondents mentioned that the group helps make the mobile phone more organized as the number of applications installed can be minimized.

However, some respondents prefer the single-interest group over the multiple-interest type. A female respondent mentioned that "I don't like to mix up all the interest community in a single app. I tend to categorize the social apps so that I can know that these certain apps are aimed for which community". A respondent who does not favor the interest groups mentioned that "I have less willingness to come together to address issues and have no commitment to deliver ideas."

Those who favor the anonymity feature stated that they think the feature helps in privacy and security aspects and makes them feel safer. They mentioned that the anonymity feature is essential as it can improve internal communications. The feature also enables people to communicate with strangers anonymously through their smartphones when they are unwilling to disclose their real names to the public or people not so close to them. They also suggested that users be given more options to hide their identity and that there should be more options to be anonymous on different applications. A female respondent says that "The interest groups feature in mobile apps allow users to be anonymous for more honesty, openness, and diversity of opinion, meanwhile encourage expressiveness and interaction among users."

Some respondents also suggested that a real name is required for identity verification but only shows the username in the application. A male respondent who likes the anonymity feature mentioned that "Even though anonymity feature is preferable, the authority should have some form of access to a small portion of users' data like age and name to avoid any illegal behavior and the application users have to know about it."

The results yielded from the questionnaires in the pilot study imply the on-demand characteristics of the mobile apps market, especially for Android users. The market segment analysis shows the readiness of the apps providers in terms of providing interest groups (single/multi groups) with anonymity features. As for Malaysian users, the anonymity feature is highly desirable for interest group apps, but there is no

difference in their preference in the type of interest group. Nevertheless, as Malaysia is a multi-racial country, we discovered that a better understanding of the categories of the interest-based apps with anonymity features favored by Malaysians (races/ethnicity) could be gained if the questionnaires include questions regarding this topic. A similar improvement can also be made on the market segment analysis of the mobile apps.

V. CONCLUSIONS

In conclusion, we have approached anonymity features in mobile apps interest groups using an explanatory, quantitative approach in a pilot study. Within the scope of the research presented in this paper, the results support our hypothesis that mobile apps users' who highly value interest groups will also highly appreciate the anonymity feature provided in the interest groups. The study has allowed us to provide three implications for mobile apps design: including the anonymity feature and interest groups, considering the multi-interest groups in a single app, and the need to embed fair security aspects when the anonymity feature is activated. Even though they are preliminary, the results of this pilot study are adequate for us to proceed with wider audiences (i.e., more age groups, nationality). They need to be validated using representative research or similar case studies in different contexts. Also, the scope of this research is limited to the Android platform. Thus, the results of the study cannot be used to represent other platforms. However, the results could open new research ideas and the design of focused mobile apps. Research that looks at the security feature of mobile apps with anonymity will contribute to a broader understanding of the subject. Such a study could consider such questions as "How can we embed the acceptable security aspect of mobile apps that allow anonymous users?"

ACKNOWLEDGMENT

We would like to thank the *Fakulti Teknologi Maklumat Dan Komunikasi, Universiti Teknikal Malaysia Melaka (UTeM)* for supporting this research. This work is under research grant *PJP/2020/FTMK/PP/S01771*.

REFERENCES

- [1] A. Rosales and M. Fernández-Ardèvol, "Smartphones, apps and older people's interests: From a generational perspective," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2016, 2016*.
- [2] K. Matthews, "6 Tools to Find Awesome People with Similar Interests," *makeuseof.com*, 2015. <https://www.makeuseof.com/tag/6-tools-find-awesome-people-similar-interests/>.
- [3] L. Martin, "Facebook Ad Targeting: How to Use Interest Groups," *Lightning AI*, 2019. <https://blog.lightningai.com/facebook-ad-targeting-how-to-use-interest-groups-1ec7dcc9d04a>.
- [4] X. Ma, J. Hancock, and M. Naaman, "Anonymity, intimacy and self-disclosure in social media," in *Conference on Human Factors in Computing Systems - Proceedings, 2016*, doi: 10.1145/2858036.2858414.
- [5] A. Vasilateanu and C. Casaru, "Anonymity in Health-oriented Social Networks," in *Procedia Computer Science, 2015*.
- [6] T. Wright, "Security, privacy, and anonymity," *XRDS Crossroads, ACM Mag. Students, 2004*.
- [7] S. K. T. Febriana and Fajrianthi, "Cyber Incivility Perpetrator: The Influenced of Dissociative Anonymity, Invisibility, Asynchronicity, and Dissociative Imagination," in *Journal of Physics: Conference Series, 2019*.

- [8] C. Piao and X. Li, "Privacy Preserving-Based Recommendation Service Model of Mobile Commerce and Anonymity Algorithm," in Proceedings - 12th IEEE International Conference on E-Business Engineering, ICEBE 2015, 2015.
- [9] C. Denzil, A. S. Leandro, M. Mainack, B. Fabrício, and P. G. Krishna, "The many shades of anonymity: Characterizing anonymous social media content," in Proceedings of the 9th International Conference on Web and Social Media, ICWSM 2015, 2015, pp. 71–80.
- [10] D. Bulatovych, "How to Build a Successful Anonymous Social Networking App," Yalantis.com, 2019. <https://yalantis.com/blog/anonymous-social-network-development/>.
- [11] A. L. Forest and J. V. Wood, "When social networking is not working: Individuals with low self-esteem recognize but do not reap the benefits of self-disclosure on facebook," *Psychol. Sci.*, 2012.
- [12] C. Hu, S. Kumar, J. Huang, and K. Ratnavelu, "Disinhibition of negative true self for identity reconstructions in cyberspace: Advancing selfdiscrepancy theory for virtual setting," *PLoS One*, 2017.
- [13] J. Suler, "The online disinhibition effect," *Cyberpsychology Behav.*, vol. 7, no. 3, pp. 321–326, Jun. 2004.
- [14] M. Iqbal, "App Download and Usage Statistics (2020)," *BusinessofApps*, 2020. <https://www.businessofapps.com/data/app-statistics/>
- [15] H. Nurhayati-Wolff, "Social media users as a percentage of the total population Malaysia 2021," *statistica.com*, 2021. <https://www.statista.com/statistics/883712/malaysia-social-media-penetration/#:~:text=Facebook's popularity in Malaysia, social media users in Malaysia.>
- [16] A. Gotter, "Everything You Need to Know About Facebook Groups Marketing," *Adespresso.com*, 2020. <https://adespresso.com/blog/facebook-groups-everything-you-need-know/>.
- [17] E. Moreau, "Everything You Need to Know About Facebook Groups," *lifewire.com*, 2020. <https://www.lifewire.com/facebook-groups-4103720>.
- [18] E. Grinberg, "Facebook' real name' policy stirs questions around identity," *CNN*, 2014. <https://edition.cnn.com/2014/09/16/living/facebook-name-policy/index.html>.
- [19] S.-L. Chen, "What's In A Name? – Facebook's Real Name Policy And User Privacy," *Kansas J. Law Public Policy's*, vol. 28, no. 1, pp. 146–172, 2019, Accessed: Mar. 11, 2021. [Online]. Available: https://lawjournal.ku.edu/wp-content/uploads/2019/08/V28_I1_05_Chen_Web.pdf.
- [20] CBC News, "Facebook makes changes to 'real names' policy after complaints," *cbc.ca*, 2018. [https://www.cbc.ca/news/technology/facebook-real-names-1.3367403#:~:text=Technology %26 Science-,Facebook makes changes to "real names" policy after complaints,to have their accounts suspended.](https://www.cbc.ca/news/technology/facebook-real-names-1.3367403#:~:text=Technology%26Science-,Facebook makes changes to 'real names' policy after complaints,to have their accounts suspended.)
- [21] A. M. Cirucci, "Facebook's affordances, visible culture, and anti-anonymity," in *ACM International Conference Proceeding Series*, 2015, doi: 10.1145/2789187.2789202.
- [22] J. Suler, "The online disinhibition effect," *Cyberpsychology and Behavior*. 2004.
- [23] C.-H. Lai, B.-S. Jong, Y.-T. Hsia, and C. Yuan, "Paper-Use of a Mobile Anonymous Question-Raising System to Assist Flipped-Classroom Learning Use of a Mobile Anonymous Question-Raising System to Assist Flipped-Classroom Learning," *Int. J. Interact. Mob. Technol.*, vol. 14, no. 03, pp. 66–81, Feb. 2020.
- [24] B. Jagdale and J. Bakal, "Privacy aware monitoring of mobile users in sensor networks environment," *Int. J. Interact. Mob. Technol.*, vol. 13, no. 2, pp. 127–140, 2019.
- [25] M. M. Alani, "Android Users Privacy Awareness Survey," *Int. J. Interact. Mob. Technol.*, vol. 11, no. 3, pp. 130–144, Apr. 2017.
- [26] I. Bose and X. Chen, "Quantitative models for direct marketing: A review from systems perspective," *Eur. J. Oper. Res.*, vol. 195, no. 1, pp. 1–16, May 2009.
- [27] D. A. Siegel, "The mystique of numbers: Belief in quantitative approaches to segmentation and persona development," in *Conference on Human Factors in Computing Systems - Proceedings*, 2010, pp. 4721–4731.
- [28] J. Bolluyt, "20 Android and iOS Apps to Give You Privacy and Anonymity," *cheatsheet.com*, 2014. <https://www.cheatsheet.com/technology/20-android-and-ios-apps-to-give-you-privacy-and-anonymity.html/>.
- [29] Lulu, "14 Important UX Questions to Include in Your Mobile Application Survey," *UX design.cc*, 2019. <https://uxdesign.cc/14-important-ux-questions-to-include-in-your-mobile-application-survey-b1e094c2620c>.
- [30] Delighted, "Mobile app surveys: Best practices and sample questions," *delighted.com*, 2020. <https://delighted.com/blog/mobile-app-survey-best-practices>.
- [31] Michael Litwin, "Mobile App Survey: Complete Guide with Question Examples," <https://survicate.com/>, 2018. <https://survicate.com/mobile-app-survey/mobile-app-survey-questions/>.

APPENDIX

No	App Name	Logo	No	App Name	Logo	No	App Name	Logo
1	Amino		23	Geek Dating		45	Qooiver	
2	Anime		24	GLIDE		46	SHEROES	
3	AnimeFansBase		25	Glstars		47	SoloLearn	
4	Capture		26	Goodwall		48	Still Active	
5	Clubify		27	Grasscity Forum		49	Still Active	
6	ComeOut		28	hi-hive		50	Tandem	
7	Comma		29	Hobbinity		51	Tapatalk	
8	Commune Community		30	iConek Community Hub		52	The Mighty	
9	Community		31	Instagram		53	Topiks	
10	Community		32	InterNations		54	Topiks	
11	Community		33	IQ		55	TUBBR	
12	Community		34	KStarLive		56	Tumblr	
13	Community Café		35	LMK		57	Vent	
14	DailyAct		36	Lysn		58	Vingle	
15	Defense News		37	Meetup		59	We Heart It	
16	DeviantArt		38	Meipai		60	WEBTOON	
17	DogChal		39	Orbis		61	We Chat	
18	Doongle		40	Panion		62	Writco	
19	Edapt		41	Photory		63	Yapper	
20	Facebook		42	PicsArt		64	Youtube	
21	FAN GURU		43	Pinterest		65	Zest	
22	Gamer Community Forum		44	PS Community				

A Method to Prevent SQL Injection Attack using an Improved Parameterized Stored Procedure

Kamsuriah Ahmad¹, Mayshara Karim²

Center for Software Technology and Management
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia, 43600, Bangi, Selangor, Malaysia

Abstract—Structured Query Language (SQL) injection is one of the critical threats to database security. The effects of SQL injection attacks cause the data contained in the database to be at risk of being exploited by irresponsible parties, compromising data integrity, disrupting server operations and in return affecting the organization's image. Although SQL injection is an attack performed at the application level, SQL injection prevention requires security controls at all levels, namely application level, database level and network level. The absence of SQL injection prevention measures at the application level makes the database vulnerable to attack. Reviews indicate that the current approaches still not sufficient in addressing these three issues, which are i) improper use of dynamic SQL, ii) lack of input validation process and iii) inconsistent error handling. Currently, program and database code security is based solely on basic security measures that are focused at the network level such as network firewalls, database access control and web server request filtering. Unfortunately, these measures are still inadequate and not sufficient to safe guard the program code and databases from the attack. To overcome this shortcoming as addressed by these three issues, a new comprehensive method is proposed using an improved parameterized stored procedure to enhance database security. Experimental results prove that the proposed method is able to prevent SQL injection from occurring and able to shorten the processing time when compared with existing methods, hence able to improve database security.

Keywords—SQL injection prevention; database security; parameterized stored procedure; network firewall

I. INTRODUCTION

The sophistication of information technology makes life easier but at the same time poses a threat if the security aspect is not given special attention [1]. The internet that connects people around the world carries the threat of hackers. SQL injection is used as a method to steal and exploit information on the database [2]. SQL injection can affect organizational data security, and now they start to realize the importance of preventing this attack before it happens [3]. The prevention from this injection has become the focus of database administrators, network administrators, application code programmers, database system providers, software developers and the top management of the organization. The problems brought by SQL injection have been around for a long time and are still a hot topic in information security issues [4] [5]. Web application security and database security are inter-

dependent. If a web application is manipulated by hackers, databases that have been equipped with security features can still be exploited. There are a number of threats to database security: among them are i) excessive privilege which refers to users who are given permission to access or carry out various transactions in the database but abuse the permission instead, and ii) SQL injection which is the entry of unauthorized input into the database to carry out any instructions that are not valid [6]. In addition, weak audit trails or automatic recording of database transactions that are not performed properly and media used as storage for backups such as hard disks and tapes are also prone to theft [7]. Therefore to improve database security, this paper aims to propose a new method in preventing SQL injection. To discuss the existing prevention methods on SQL injection and the motivation of this work, this paper is organized as follows: section II discusses the issues of SQL injection attack. Section III describes various ways that cause web applications to be attacked by SQL injection. Section IV highlights the impact of SQL injection on database security. Section V reviews the existing SQL injection prevention method. Based on the limitations faced by the existing methods, a new method is proposed in Section VI. Section VII explains the experiment conducted to prove the efficiency of the proposed method and Section VIII concludes the study and future work suggestion.

II. SQL INJECTION ATTACK

SQL is a standard programming language used to access databases. SQL injection is the act of putting malicious code as an input to a web application and sending it to a database to execute various commands [8]. An attack occurs when a hacker exploits an SQL injection to execute an unauthorized command [9]. Examples of common exploits are: stealing confidential information that can bring profits such as credit card numbers, bank savings account numbers, passwords, business transaction records and medical records. In addition, exploitation can also involve data modification [10]. For example; students trying to change grades or exam marks. There are also cases which are not personal attacks; such as sabotage the database by deliberately deleting certain tables, shutting down database operations and disrupting network traffic [4]. The term 'injection' is used because malicious code which considered as a non-valid input is injected into a valid SQL statement. The database will execute this command because it is valid in terms of the syntax and rules [11]. Fig.1 explains the scenario of SQL injection attack.

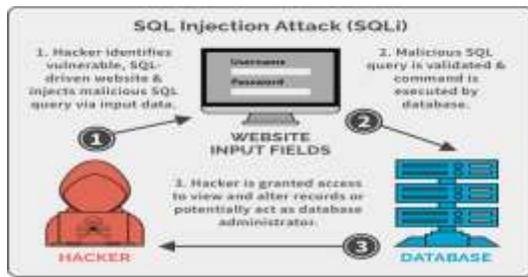


Fig. 1. SQL Injection Attack [2].

The term attack in the context of SQL injection is defined as an unauthorized access to applications or systems which normally done by hackers [12]. This access is obtained through SQL injection mechanisms derived from SQL query modifications [9]. Web applications that have user input can be attacked by SQL injection because the application code has certain vulnerabilities that expose to the attacks.

This will lead to the occurrence of bugs, loopholes, vulnerabilities or defects. These weaknesses shall be manipulated by unauthorized users to gain unrestricted access to stored data [10]. Therefore, a proper mechanism is needed to prevent the attack, such as validate the user input both at the client and server side [13] [14]. Since lack of proper mechanism in preventing the attack at the application and database level exists in the literature, therefore it has become the motivation of this study. In order to understand in details, next section will discuss the types of SQL injection that usually attack databases.

III. SQL INJECTION MECHANISM

Web applications have many loopholes and make them vulnerable to SQL injection attacks. There are various ways that cause web applications to be attacked by SQL injection. Among them are discussed as follows.

- SQL Injection through error messages.

An error-based injection is a type of SQL injection derived from an error message to figure out the structure and the type of database [15]. The attacker intentionally enters wrong input logically in order to allow database generates an error message. This error message contains information that allows the attacker to identify the parameters vulnerable to the injection.

- Boolean-based blind SQL injection

The term blind means that the SQL injection is performed when the programmer has set a generic custom error message in case web application encounters an error [16]. Without displaying error messages, database vulnerabilities can be protected. The hacker has to deal with a database system that does not display error messages and as an alternative, hackers submit a series of "TRUE" and "FALSE" queries via SQL queries [17]. Information about the database will be revealed through the results of these queries.

- Time-based blind SQL injection

Time-based SQL injection means hackers obtain information on database based on response times. SQL

command is sent to database with code to force the database to wait for a specified amount of time during the execution of the queries. The response time indicates whether the result of the query is true or false. While waiting for the query to be processed, the attacker able to execute another query and might inject malicious code in the query [18].

- UNION-based SQL injection

The hacker connects the SQL injection to the original SQL query by using the UNION clause to retrieve data from another table. The result of this injection is that the database provides a data set that contains a combination of the results of the original query and the results of the SQL injection query [19].

To avoid SQL injection through error messages will be the focus of this study. Hackers have a variety of reasons and motivations when it comes to hacking. Hackers gain access and control to databases illegally through SQL injection. With such access hackers can perform various actions to manipulate the data stored in the database. The main impact when SQL injection occurs is the disruption of the confidentiality and integrity of the data in the database [20]. Table I describes the actions that hackers can take and the consequences of the actions taken.

TABLE I. THE ACTION AND THE EFFECTS OF SQL INJECTION ATTACKS

	Action	Effects
1	Identify parameters that can be injected	Attackers able to find input parameters that are vulnerable to SQL injection
2	Perform a fingerprint on the database	An attacker can find out the type and version of the database being used. Easier to devise a more specific attack strategy.
3	Identify database schema	Attackers able to extract the information on database schema accurately
4	Extraction of data	Attackers able to extract complete data from database table
5	Add, remove and edit data	Attackers can alter certain data for their own benefit

IV. THE IMPACT OF SQL INJECTION ATTACK ON DATABASE SECURITY

Most organizations are unaware that their web applications are vulnerable to SQL injection attack or have been attacked by SQL injection [21]. Awareness and knowledge of SQL injection is still lack in the organization, any prevention steps are mostly at the basic level such as network firewall installation, the use of Secure Socket Layer (SSL) and network access control [2]. However, these measures only provide protection at the network level, and not at the application level [11]. Reviews state that the weakness in preventing SQL injection lies in the programming code of the web application itself [15]. Although various types of techniques have been introduced, most of the researchers insist that the application program code needs to be ensured its security first [22]. Firewalls and protocol information unable to protect web applications from hackers due to their position are behind the web application infrastructure [12]. The risks of attacks are increased due to the exposure of web application, firewall-friendly HTTP protocols and lack of database security [23].

Even though, there are approaches to improve database security through SQL injection prevention both at application and database level, however they are not capable to protect database against these three attacks which are:

- 1) vulnerabilities triggered when using dynamic SQL [13][18].
- 2) malicious code entered through input data [5][6].
- 3) inappropriate error handling message [10][14].

The cause of these attacks is explained as follows:

- Dynamic SQL usage

Dynamic SQL is a query language to build SQL statements and performs functional logic such as data search, user login and others, where this query is created dynamically at run-time. Although, these language are beneficial and user-friendly; but at the same time they expose the web applications to SQL injection. Often web applications are vulnerable to the attacks when dynamic SQL statements are being used [7]. Because, this language is dynamic in nature, therefore, an unauthorized user can modify the original query by injecting a malicious code to the query at run-time [24]. The following shows the impact of using dynamic SQL on database security. An example is used to show how SQL injection attacks might occur when using dynamic SQL. During runtime, the following dynamic SQL statement is sent by the web server to the database server:

```
SELECT * FROM program WHERE ID = 9
```

The intention of this query is to list the data from all the columns in a table named program; where column ID is 9. Since the query is dynamic, an hacker with a good knowledge of programming able to manipulate the query by injecting malicious code and modify the query dynamically at runtime, for example adding the symbol (') after the number 9. SQL statement is changed to

```
SELECT * FROM program WHERE id = 9'
```

This example shows that the use of dynamic SQL gives an opportunity to the unauthorized user to perform attack on database. Therefore the used of dynamic SQL should be avoided [18]. Since most organization used stored procedure with dynamic SQL [10], therefore, a preventive step at the database level is needed to avoid SQL injection attempts.

- Input validation

Web applications allow users to log in into the system. However, the data entered by the user may contain some malicious code and can easily exploit through SQL injection [25]. The inputs need to be checked and validated beforehand. The absence of an input validation process will contribute to the vulnerability of SQL injection. When a web application ensures that the received input is within the expected range, the malicious code could be prevented from entering the database server thereby preventing SQL injection. Therefore, a proper mechanism is needed to validate the input in order to prevent malicious code from entering the database server.

- Error handling process

During data processing, system will prompt an error message if they encounter problems. However, the default error messages reveal sensitive information and weaknesses of the database. The intruder of the system will learn from these errors, and this will give an opportunity for them to breach the system [14].

The limitations of the existing prevention methods at the application and database levels in the organization have become the motivation of the study. An improved method that able to overcome these three issues discussed above is needed in enhancing organization security level.

V. AN OVERVIEW OF THE EXISTING SQL INJECTION PREVENTION METHOD

Measures to improve the security of web applications and databases can be done at the server, network, database and application levels. However, most organizations take the basic security measures at the server and network level only [26]. The server and ICT network level is the physical level which is also an asset subject to the security procedures set by the existing policies. For example, organizations in the public sector need to comply with the Security Policy which outlined rules for internet access control, use of firewalls, server configuration settings, user access control and others.

Since this study focuses on the improvement of SQL injection prevention at the application and database level, the reviews of the existing methods are based on these two levels, which are summarized in Table II. Existing methods are reviewed based on their ability to address the threats causes by the three issues discussed namely i) dynamic SQL usage, ii) data validation at both client and server side, iii) error handling process. These three issues need to properly address to prevent SQL injection.

As Table II indicates, the existing methods did not address the three issues highlighted in one single method; therefore they are not sufficient in securing the database from SQL injection attacks. The method proposed by [4][12][21][24] validates the data at the client side and ignore the validation process at the server side. Therefore malicious code might enter the database. Even though methods proposed by [6][23][27][28] used parameterized query however dynamic SQL is still used. Therefore, this method is unable to prevent unauthorized user from modifying the original query by injecting malicious code. The used of dynamic SQL should be avoided to increase database security [2]. Dynamic analysis as used in [25] increases the query processing response time. As an organization that constantly deals with users through web applications, fast response times are essential. There are previous researchers who embedded handled error messages in their approaches, but the usage is inconsistent since the data is not validated at the application and database level [2][14]. Therefore these approaches are not sufficient in dealing with SQL injection attacks.

TABLE II. EXISTING APPROACHES IN SQL INJECTION

Author	Method used	Remarks
[2]	Error handling	Users are able to access database information through error messages. However, the error handling messages are not done consistently
[4]	Input Validation	The validation is only at the application side
[6]	Parameterized query	The data given by the users is encoded. However since this approach used dynamic SQL to encode the data, therefore the method did not prevent the system from the attack.
[12]	Input Validation	Input validation is only at the application or client side and ignores the validation at the server side.
[13]	Dynamic SQL	The data is not properly validated and error messages are not handled consistently.
[14]	Error handling	Proposed signature-based detection to handle error messages. However, this approach did not embed the input validation at both client and server side.
[18]	Dynamic SQL	Validate the input at both client and server. However, SQL injection is still being used and error handling is not properly done.
[21]	Input Validation	Input validation is only at the application or client side and ignores the validation at the server side.
[23]	Parameterized query	This approach did not validate the input data; therefore invalid data might enter the database.
[24]	Input validation	Use dynamic SQL to separate between normal data and malicious data. However, this process still vulnerable to the attack.
[25]	Dynamic SQL	This method uses combined static and dynamic analysis to prevent malicious code produced by dynamic SQL. However, the process will increase the query processing time. The data is not properly validated and error messages are not handled consistently.
[26]	dynamic SQL	Used dynamic SQL in the methodology, therefore not efficient in preventing SQL injection
[27]	Parameterized query	Use Knuth-Morris-Pratt matching algorithm and parameterized query to detect and prevent SQL injection. However, the error messages are not handled properly.
[28]	Parameterized query	Use signature-based approach to prevent the attack. However, the error messages are not handled properly.

Based on the comparative analysis and to overcome the limitations of the existing methods this study proposed a comprehensive method to address the highlighted issues. The existing stored procedure methods is referred and improved by incorporating parameterized queries as a SQL injection prevention mechanism at the database level and developing input validation and error handling as a preventive mechanism at the application level. The proposed method intends to address the issues highlighted in Section IV. The details on the proposed method are discussed in the next section.

VI. AN IMPROVED STORED PROCEDURE WITH PARAMETERIZED QUERY

Currently, most organization used stored procedure as a prevention method at the application level [23][27][28].

Stored procedures are subroutines which contain a set of predefined SQL code that can be saved and reused over and over again when needed. The code is written, compiled and executed before it is being used by the web application to communicate with databases. Frequently used queries can be saved in a stored procedure and can be recalled when requested. Since the location of stored procedure is within the database server, therefore the interaction between the program code and the table containing the data can be avoided [29]. This will improve the security of application program code because it can add another layer of abstraction before interacting with database [19]. With these features, stored procedure is able to provide high security feature to the database.

A parameterized query (or a prepared query statement) is a pre-compiling SQL statement. One or more parameters (or variables) need to embed into the statement for it to be executed. Parameters are sent to the query as soon as the user presses the 'Submit' button or presses the link in the web application. Normally, the existing code for the web application contains dynamic SQL statements when performing a query in the search functions. Due to limitation in dynamic SQL, parameter queries will be used instead, in conjunction with stored procedures. Parameter queries have program code security features that can be used instead of dynamic SQL. Parameters containing embedded user input will not be interpreted as commands to execute, thus will not allow code to be injected by SQL injection. The use of these parameterized queries is intended to replace dynamic SQL queries that can cause vulnerabilities to SQL injection. Application code will be more secured because functional logic is defined first and input parameters are entered after functional logic is processed [2].

Therefore this study proposed a comprehensive method to overcome the three issues highlighted earlier as discussed in Section 4. The proposed method makes an improvement based on these three steps:

- 1) use stored procedure with parameterized query instead of SQL dynamic,
- 2) validate the input at both client and server to reduce vulnerabilities to SQL injection. This is to ensure the input parameters sent to the database server do not contain malicious code and to prevent the code from accessing and harm the stored procedure,
- 3) improve the error handling process by not displaying confidential database information. Since error messages may reveal the limitation of the system to the intruder.

Fig. 2 shows the workflow of the proposed method. The workflow consists of four components which are:

- 1) the search page where the user enters the data to search,
- 2) the data validation at the client and the server site,
- 3) the error message if problems encountered during the process, and
- 4) the output of the query.

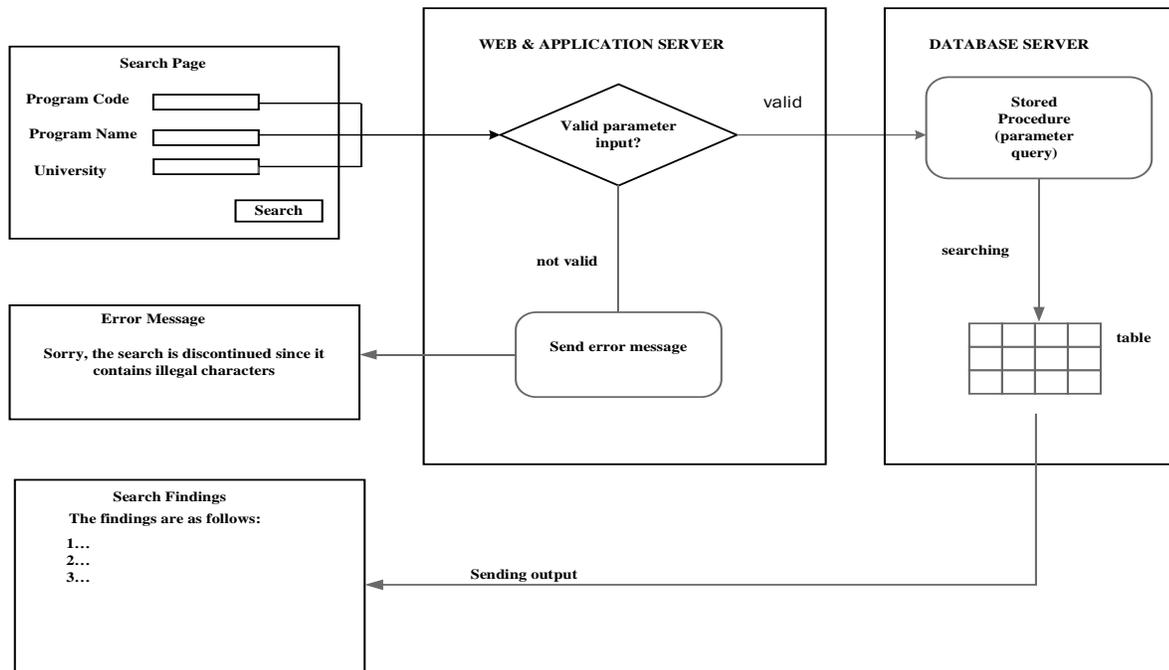


Fig. 2. The Workflow of the Parameterized Stored Procedure Improvement.

```

USE [ProgDB]
GO
/*****Object: StoredProcedure [dbo].[search_programmes] *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo] [search_programme]
    @arg_programmcode NVARCHAR(10),
    @arg_programmename NVARCHAR(25),
    @arg_instname NVARCHAR(25),
AS
BEGIN
    DECLARE @sql NVARCHAR(max),
            @a NVARCHAR(100),
            @b NVARCHAR(100),
            @c NVARCHAR(100),
            SET @sql 'SELECT programme_code,
                    autoID,
                    inst_name_E,
                    inst_name_ML,
                    inst_name_EL,
                    prog_name_EN,
                    programme_name_E,
                    Category,
                    Status,
                    Valid_date, AA_yes,
                    FROM Q_search where 1=1'
    SET @a = '%' + @arg_programmcode + '%'
    SET @b = '%' + @arg_programmename + '%'
    SET @c = '%' + @arg_instname + '%'
    IF (@arg_programmcode is not NULL)
        SET @sql @sql + ' AND programme_code LIKE ' + @a
    IF (@arg_programmename is not NULL)
        SET @sql @sql + ' AND programme_name LIKE ' + @b
    IF (@arg_instname is not NULL)
        SET @sql @sql + ' AND inst_name_E LIKE ' + @c
    EXECUTE sp_executesql @sql,
        N '@a NVARCHAR(100), @b NVARCHAR(100), @c NVARCHAR(100)
    ',
    @a = @a, @
    -
    
```

Fig. 3. The Proposed Method using Parameterized Stored Procedure.

A stored procedure with parameterized query will be constructed that consists of the search function and the four components described above. SQL dynamic is not used in the query. For instance, the user will enter data regarding the program code, program name and university in the search page and the search button is clicked when finished. The input parameter will be validated at both the application and the database site. If the input data is not valid at the application site then an error message will be displayed. If the input data is valid then parameters of the search function will be sent to the stored procedure. The query will search from the table for the answers and displays the result.

The stored procedure that has the search function and the four components used in this study is shown in Fig. 3. This stored procedure consists of validating features and a proper error handling process.

VII. EXPERIMENTAL DESIGN

Since most approaches used dynamic SQL in preventing SQL injection attacks [13][18][25]; therefore this study intends to prove that the proposed method using parameterized query stored procedure is better than the existing approaches. To evaluate the effectiveness of the proposed approach, two experiments are conducted. The purposes for evaluation are to prove that:

- 1) The stored procedures that have parameterized queries are better than dynamic SQL in terms of preventing databases from SQL injection attacks.
- 2) The proposed approach has a shorter query processing time when compared the approaches that used dynamic SQL.

To achieve the aims, the experiment is conducted in two versions, which are:

- Experiment 1: the simulation attacks on web applications with dynamic SQL queries, and
- Experiment 2: the simulation attacks on web applications with stored procedures.

In this study, the web application that used as a case study will be represented by a pseudo-version web application. This application is developed to simulate and to prove the existence of web application vulnerabilities to SQL injection and subsequently became a medium for the experiment. A pseudo web application is considered a basic replica of the original web because it is developed with the same program code and functional logic as the original web application. Fig. 4 explains the used of pseudo-version web application in identifying the SQL injection attacks.

Pseudo-version web application is developed and experiment is conduct to simulate the SQL injection attacks. The experiment is run in a computer using the Microsoft Windows 7 Enterprise Edition operating system, 4GB of RAM, Intel (R) Core (TM) i5-2320 CPU @ 3.00GHz processing chip. Pseudo web applications are developed using

Microsoft SQL Server 2008 database. The web server operating system used is Windows 2008 R2 with ASP.Net, Coldfusion and Microsoft IIS 7.5 web application technologies. The stored procedures are written in the Transact-SQL language in Microsoft SQL Server.

SQLMap is used in the experiment which works in conjunction with the Python library and Netsparker Trial Edition software. SQLMap is chosen as the SQL injection vulnerability scanner instrument in this study due to its availability as open source software. This software is the most effective and popular among hackers and web application penetration testers [25]. Both experiments used the same test plan and software namely SQLMap, Netsparker and web browser. The query used for these experiments are based on the search query as in the stored procedure in Fig. 3. The setting for these experiments is explained in Fig. 5.

A. Experiment 1

The purpose of experiment 1 is to investigate the capabilities of dynamic SQL and the proposed parameterized stored procedure in preventing SQL injection attacks. To accomplish this experiment, two testing are required.

- Test 1: to evaluate SQL Injection attack on web application with the existence of dynamic SQL queries, and.
- Test 2: to evaluate SQL Injection attack on web application with parameterized stored procedures.

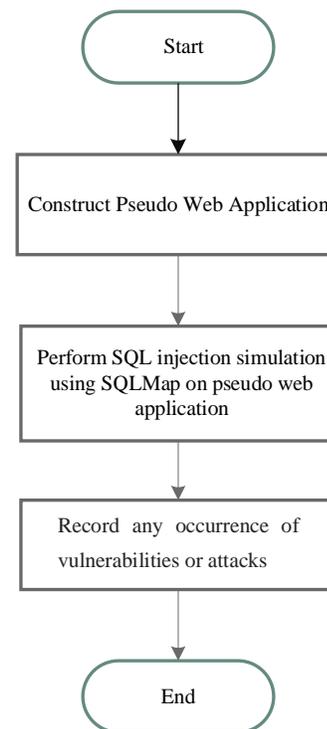


Fig. 4. Process Flow to Identify Web Application Vulnerabilities against SQL Injection.

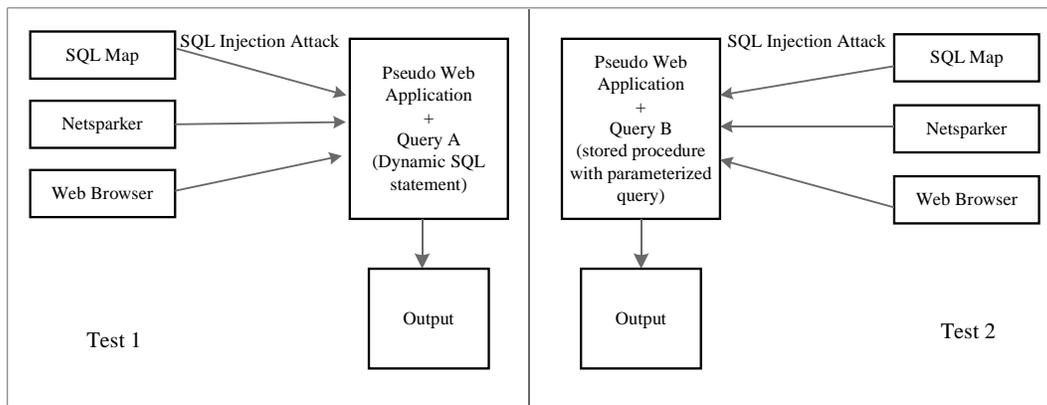


Fig. 5. The Setting for Experiment 1.

SQL which contains a malicious code is injected to SQLMap, Netsparker and web browser separately. SQL injection will penetrates web applications with queries containing dynamic SQL. Web applications that have strong defenses will be able to prevent attacks caused by SQL injection. Upon completion of the experiment, the outputs from these two tests were compared to investigate whether web application with dynamic SQL and stored procedures successful prevent SQL injection attacks. The results of this experiment are displayed in Table III.

TABLE III. THE RESULTS FOR EXPERIMENT 1

Attack	Experimental Results	
	Test 1	Test 2
SQLMap	Prevention Fail	Successful prevention
Netsparker	Prevention Fail	Successful prevention
Web Browser	Prevention Fail	Successful prevention

It is found that web applications with dynamic SQL is unable to prevent the attack cause by SQL injection when using SQLMap, Netsparker or web browser as stated in Test 1. The same instruments are used to test the web applications equipped with parameterized query, error handling and input validation. However when using the proposed approach, the web application able to prevent the attack. This proves that the proposed parameterized stored procedure is able to prevent web applications from SQL injection attacks.

B. Experiment 2

The purpose of Experiment 2 is to evaluate the performance of the proposed method and dynamic SQL in terms of time processing. This study used two evaluation instruments, namely

- Microsoft SQL Server Client Statistics and
- SQLQueryStress.

Client Statistics is a facility available in Microsoft SQL Server while SQLQueryStress is a standalone tool available as open source.

1) Evaluation using microsoft SQL server client statistics: This evaluation use MS SQL Server Client Statistics as an instrument to investigate the amount of data transmitted from server to client side. The intention is to evaluate whether SQL statement contributes to high traffic workload. The client execution time and processing time will be recorded during this evaluation.

a) Measuring client execution time: Client execution time is the cumulative amount of time used at the client side to execute query. To obtain the average execution time, ten trials were performed. In this evaluation, Test 1 represents the pseudo-web application which uses dynamic SQL during the trial, while Test 2 represents web-application which uses the proposed parameterized stored procedure. The output from this evaluation is displayed in Fig. 6.

Fig. 6 states that for each execution, Test 1 generates high execution time when compared with Test 2. This is due to a stored procedure which located at the database server. The SQL query is executed first before the web application submits the request. As such, the client is able to shorten the execution time by calling SQL queries that are already executed at the database server. In contrast, dynamic SQL are executed at runtime when the requests are submitted by the user. However, the query execution can be influenced by various factors. Such as, network traffic conditions as well as web and database server workloads. This will increase client execution time during query processing when using dynamic SQL. As the results show, the proposed parameterized stored procedure is better than dynamic SQL in terms of processing time.

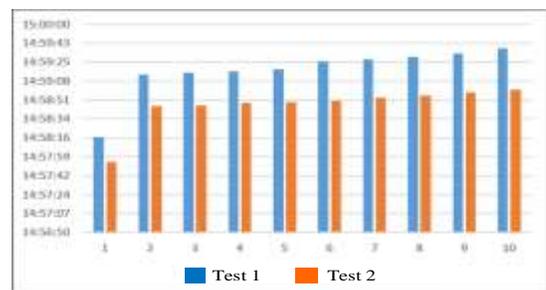


Fig. 6. Query Processing Time at the Client Side.

b) *Measuring Time Statistic:* Time statistics provide information on the amount of time used during query processing at the client side versus the time used to wait for the server. Three metrics that can be used to measure time statistics namely client processing time, server response time and total execution time.

- Client processing time- to measure query processing time at the client side.
- Wait-time on server replies - is the time spent by the client waiting to receive the response from a database server after submitting the request.
- Total execution time - the time spent by the system to execute the query, including the time spent waiting for the server to respond. Using the same instrument, the results from this experiment are shown in Table 4. Comparison of time statistics between Test 1 and Test 2.

TABLE IV. THE OUTPUT FOR TEST 1 AND TEST 2

Evaluation Criteria	Test 1 (ms)	Test 2 (ms)
i. client processing time	3.5	1.9
ii. server response time	40.2	39.6
iii. total execution time (i+ii)	43.5	41.5

As can be seen in Table IV, web applications that use stored procedures are able to process queries at the client side in a shorter time. Also in the second and third evaluation criteria, the applications with stored procedure are able to improve query processing time at the server and total execution time both at the client and server when compared with the applications with dynamic SQL. This is due to the query in the stored procedure is executed first before the web application submits the request. Using this improvement method, the time taken to process the queries will be reduced.

2) *Evaluation using SQL query stress:* This evaluation uses SQLQueryStress software and used as a benchmark to investigate the performance of the query when given heavy workload. The intention is to investigate the impact of SQL queries (whether dynamic SQL queries, parametric queries or stored procedures) towards system performance. The differences in the structure of dynamic SQL queries and stored procedures make the query executed in different ways. Therefore, it is important to know the impact of both SQL queries on the system performance. Based on the result in Table V, it is found that the time taken by Test 2 is less than the time taken by Test 1.

The results state that Test 2 has a logic reading of 1547 compared to 1620 in Test 1. SQL queries that have low logic readings were more efficient than SQL queries that produced high logic readings. This is because high logic readings have negative implications to system memory. In other words, high logic readings place heavy workload to computer systems. The results show that query embedded in stored procedures produce less workload than dynamic SQL when using the same data hence improve query performance. This is due to

the query is processed earlier at the database server rather than at the runtime. Thus, it will reduce the elapsed time, the CPU time, the actual processing time and the client time as well.

TABLE V. COMPARISON OF SQL QUERY STRESS CRITERIA ON WEB APPLICATIONS USING DYNAMIC SQL QUERIES (TEST 1) AND STORED PROCEDURES (TEST 2)

Criteria (average)	Test 1	Test 2
elapsed time (ms)	1.3600	0.7350
CPU seconds (ms)	0.0996	0.0167
actual seconds (ms)	0.2834	0.0401
client seconds (ms)	0.1009	0.0475
logical reads	1620.0	1547.0

VIII. CONCLUSIONS

This study successfully proposed a new comprehensive method using an improved parameterized stored procedure to overcome the three issues highlighted in preventing web application from SQL injection attack. These three issues are i) improper use of dynamic SQL, ii) lack of input validation process and iii) inconsistent error handling. Unfortunately, the existing approaches did not properly address these issues. Hence they are not able to prevent SQL injection attacks at both the application and database side. In this study, a stored procedure is constructed that consists of a comprehensive method to address these three issues which are built in the code. To prove the effectiveness, the proposed method was evaluated through three approach of attack simulation, namely using SQLMap software, Netsparker and web browser. The experiments conclude that SQL injection does not successfully penetrate web applications and databases when the proposed method is implemented hence able to overcome the limitations faced by the existing methods. This indicates that the SQL injection prevention method developed has successfully prevents SQL injection from occurring. The proposed method is also evaluated from the perspective of time used and its impact on the overall system. Evaluations using Microsoft SQL Server Client Statistics and SQLQueryStress software have concluded that although there are slight differences in time processing, the proposed method uses a shorter query processing time when compared with dynamic SQL. It can be concluded that the SQL injection prevention method used does not generate high overhead that may lead to high response time. This study can be further improved by focusing the prevention of SQL injection in network systems. There are various factors to be considered such as the types of server used, network traffic and the attacks from various sources. More efforts are needed and further enhancements are required to improve database security.

ACKNOWLEDGMENT

This research was sponsored by the Research Incentive Grants (Grant No. GGP-2019-024), Centre for Software Technology and Management (SOFTAM) of Faculty of Information Science and Technology, National University of Malaysia (UKM).

REFERENCES

- [1] A.Jamil and Zawiyah Mohammad Yusof, "Information Security Governance Framework of Malaysia Public Sector", *Asia-Pacific Journal of Information Technology and Multimedia*, vol. 7 no. 2, 2018, pp.85 – 98.
- [2] L.Ma, D.Zhao, Y.Gao and C. Zhao, "Research on SQL Injection Attack and Prevention Technology Based on Web", *International Conference on Computer Network, Electronic and Automation (ICCNEA)*, Xi'an, China, 2019, pp. 176-179.
- [3] N. Singh, M. Dayal, R. S. Raw and S. Kumar, "SQL injection: Types, methodology, attack queries and prevention," *3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2016, pp. 2872-2876.
- [4] Z. C. S. S. Hlaing and M. Khaing, "A Detection and Prevention Technique on SQL Injection Attacks", *IEEE Conference on Computer Applications (ICCA)*, Yangon, Myanmar, 2020, pp. 1-6.
- [5] H. Gupta, S. Mondal, S. Ray, B. Giri, R. Majumdar and V. P. Mishra, "Impact of SQL Injection in Database Security", *International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, Dubai, United Arab Emirates, 2019, pp. 296-299.
- [6] M.Malik, and Patel, T., "Database Security - Attacks and Control Methods", *International Journal of Information Sciences and Techniques*, 6(1/2), 2016, pp.175–183.
- [7] H. Gaikwad, Bhavesh B. Shah and Priyanka Chatte, "SQLi and XSS Attack Introduction and Prevention Technique", *International Journal of Computer Applications (0975 – 8887)* May 2017 volume 165 – No.2, 23-27.
- [8] J. Clarke, *SQL Injection Attacks and Defense*, Second Edition, 2012. Syngress, pp.1–473.
- [9] P. S. P. Pullagura and Gokilavani, "Defeating SQL Injection on Preventing Run Time Attacks", *International Journal Of Science & Technoledge*, 2(5), 2014, pp. 93–96.
- [10] Kindy, D. A. & Pathan, A. K.. A Detailed Survey on various aspects of SQL Injection in Web Applications : Vulnerabilities , Innovative Attacks and Remedies. *International Journal of Communication Networks and Information Security*, 5(2), 2013, pp.80–92.
- [11] Q. Li, W. Li, J. Wang and M. Cheng, "A SQL Injection Detection Method Based on Adaptive Deep Forest," in *IEEE Access*, vol. 7, 2019: pp. 145385-145394,
- [12] V. Dwivedi , Himanshu Yadav and Anurag Jain, "SQLas: Tool to Detect And Prevent Attacks in PHP Web Applications", *International Journal of Security, Privacy and Trust Management (IJSPTM)* Vol 4, No 1, February 2015, pp.21-30.
- [13] G. Deepa and Thilagam, P. S.. "Securing Web Applications from Injection and Logic Vulnerabilities: Approaches and Challenges", *Information and Software Technology*, 74, 2016, pp.160–180.
- [14] A.Jumaa and Omar, A.. "Online Database Intrusion Detection System Based on Query Signatures", *Journal of University of Human Development*, 3(1), 2017, pp.282–287.
- [15] M. Rami and F. Jnena, "Modern Approach for WEB Applications Vulnerability Analysis", *Master of Science in Computer Engineering. The Islamic University of Gaza Deanery of Graduate Studies Faculty of Engineering Computer Engineering Department*, 2013.
- [16] M.Amirulluqman Azman, Mohd Fadzli Marhusin and Rossilawati Sulaiman, "Machine Learning-Based Technique to Detect SQL Injection Attack", *Journal of Computer Science*, 17 (3), 2021, pp.296-303.
- [17] J.Minhas and Kumar Raman, "Blocking of SQL Injection Attacks by Comparing Static and Dynamic Queries", *International Journal of Computer Network and Information Security*; 5(2), Feb 2013, pp.1-9.
- [18] M.Amin Mohd Yunus, Muhammad Zainulariff Brohan and Nazri Mohd Nawi. "Review of SQL Injection: Problems and Prevention". *International Journal On Informatics Visualization*, vol 2, 2018, No 3 – 2.
- [19] S. Choudhary and Nanhay Singh.. Safety "Measures and Auto Detection against SQL Injection Attacks", *International Journal of Engineering and Advanced Technology (IJEAT)*, Volume-9 Issue-2. 2019, pp. 2827 – 2833.
- [20] P. Sadotra and Chandrakant Sharma. SQL "Injection Impact On Web Server And Their Risk Mitigation Policy Implementation Techniques: An Ultimate Solution To Prevent Computer Network From Illegal Intrusion", *International Journal of Advanced Research in Computer Science*, Volume 8, No. 3, March – April 2017, pp. 678-686.
- [21] K.G. Vamshi, V. Trinadh, S. Soundabaya, and A. Omar, "Advanced Automated SQL Injection Attacks and Defensive Mechanisms", in *Annual Connecticut Conference on Industrial Electronics, Technology & Automation (CT-IETA)*, 2016, p. 1-6.
- [22] A.Alazab and Ansam Khresiat, "New Strategy for Mitigating of SQL Injection Attack", *International Journal of Computer Applications* 154(11), 2016, pp.1-10.
- [23] M.Horner and Hyslip, T.. "SQL Injection: The Longest Running Sequel in Programming History", *Journal of Digital Forensics*, 12(2), 2017, article 10.
- [24] R. Mohamed Thiyab, Musab A. M. Ali, Farooq Basil and Abdulkader, "The impact of SQL injection attacks on the security of databases", *Proceedings of the 6th International Conference of Computing & Informatics*, 2017, pp. 323-331.
- [25] I.Lee, Soonki Jeong, Sangsoo Yeo and Jongsub Moon, "A novel method for SQL injection attack detection based on removing SQL query attribute values", *Mathematical and Computer Modelling*, volume 55, issues 1–2, 2012, Pages 58-68.
- [26] S. Nanhay, D. Mohit, R.S. Raw, and K. Suresh, "SQL Injection: Types, Methodology, Attack Queries and Prevention", in *3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, p. 2872 – 2876.
- [27] O.C. Abikoye, Abubakar, A., Dokoro, A.H. et al. "A novel technique to prevent SQL injection and cross-site scripting attacks using Knuth-Morris-Pratt string match algorithm", *EURASIP J. on Info. Security*, 2020, 14.
- [28] S. Choudhary, Arvind Kumar and Anil Kumar, "A Detail Survey on Various Aspects of SQLi", *International Journal of Computer Applications*, 161(12), 2017, pp.34-39.
- [29] A. Zhu and Wei Qi Yan, "Exploring Defense of SQL Injection Attack in Penetration Testing", *International Journal of Digital Crime and Forensics*, 9(4), 2016, pp. 62-71.

From User Stories to UML Diagrams Driven by Ontological and Production Model

Samia Nasiri¹, Yassine Rhazali², Mohammed Lahmer³, Amina Adadi⁴
LMMI Laboratory of ENSAM, Moulay Ismail University
ISIC Research Team of ESTM
Meknes, Morocco

Abstract—The User Story format has become the most popular way of expressing requirements in Agile methods. However, a requirement does not state how a solution will be physically achieved. The purpose of this paper is to present a new approach that automatically transforms user stories into UML diagrams. Our approach aims to automatically generate UML diagrams, namely class, use cases, and package diagrams. User stories are written in natural language (English), so the use of a natural language processing tool was necessary for their processing. In our case, we have used Stanford core NLP. The automation approach consists of the combination of rules formulated as a predicate and an ontological model. Prolog rules are used to extract relationships between classes and eliminate those that are at risk of error. To extract the design elements, the prolog rules used dependencies offered by Stanford core NLP. An ontology representing the components of the user stories was created to identify equivalent relationships and inclusion use cases. The tool developed was implemented in the Python programming language and has been validated by several case studies.

Keywords—Ontology; prolog rules; natural language processing; UML diagrams; user stories

I. INTRODUCTION

Requirements engineering (RE) represents an important role in all types of software development processes. They aim to define the scope of development together with customers [1]. In agile software development, requirements are presented in documents named user stories. These documents are an efficient way to express requirements from the user. User stories are written in natural language that renders them easily understandable to stakeholders, indeed they are short text that depicts a semi-structured specification. A user story often uses the following format type: As <role>, I want <feature> to <reason> [2, 3].

Recently, agile software development has become more and more widely used. However, unlike the extensive automation research on RE in traditional software development, the automation of RE in agile development has not yet been investigated sufficiently, especially in the area of requirements modeling [4]. Requirements modeling is a critical process in the software engineering life cycle. It is a multi-faceted and time-consuming process. However, it provides a solid guide for the final product. The success of software projects depends mainly on careful and timely analysis and modeling of system requirements. In [5], the authors propose an approach to generate a conceptual model using heuristic

rules and the NLP tool, but this model is not complete as it lacks the attributes of each entity. In [6], the authors also analyzed user stories in order to generate a UML use case diagram, but their approach is limited as they did not extract the relationships between the use cases. Furthermore, the authors of both approaches have not refined the relationships in the conceptual model or in the use cases in the use case diagram.

Our contribution aims to automate RE in agile development in order to generate automatically three UML diagrams which are class diagram, use case diagram, and package diagram with the refinement of results. To achieve the refinement task, at first, ontology engineering is created for defining synonyms and relationships between actions, given that action is a relationship or part of the use case, secondly, Prolog rules are used to eliminate the relationships that are at risk of error in the class diagram. Our purpose is to minimize the errors of the relationships extraction and to avoid the redundancy of the associations not taken into account by Wordnet in the class diagram. Prolog rules are applied at first to determine the relationship between engineering requirements. All statements are converted to rules and facts in the SWI-Prolog language. A user story is made up of three elements: the role which represents the actor who acts, then the action represented by a verb, and finally the object which has undergone the action. Ontology is created to describe the components of user stories as the role, the action and the object. This ontology represents the field of agile methods by focusing on the part of user stories.

After the creation of the classes in ontology editor, we proceeded to the stage of filling the ontology by enriching it with vocabulary and equivalent of class instances; we concentrated on the class "action" which represents the relation between the classes in a diagram of UML. The object properties reflect the relations that can be established between instances of the ontology classes.

The structure of this paper is as follows: This section introduces agile methods and our proposed approach. In Section 2, we present the related work of our proposal. However, we detail our proposal approach, and we present the main of the platform in Section 3. Then, we present a generated UML diagrams in Section 4. In Section 5, we present the discussion and analysis. At finally, conclusion is presented in Section 6.

II. RELATED WORK

Several research projects have been carried out to automate the requirement engineering, but few researchers have developed a tool to automate the agile requirement presented in user stories. Since the agile method is the most used in software engineering, it was necessary to think about developing a solution to automate the design phase in the software development Life cycle. A user story is a very effective means of communication between future users of the software and developers and designers.

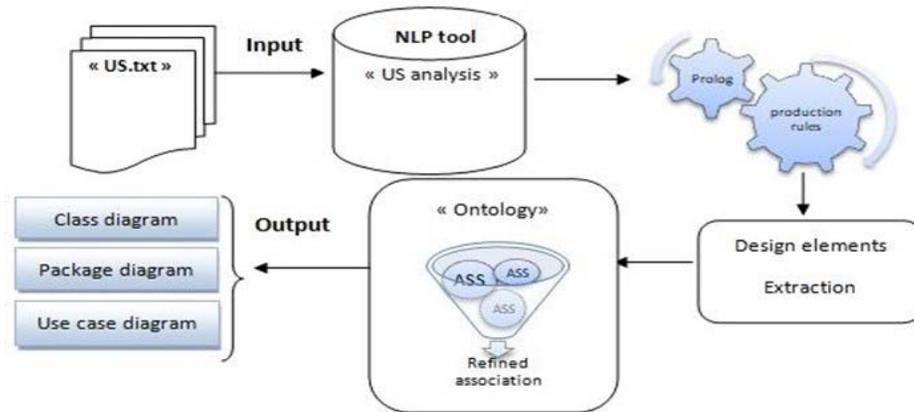
Our approach automates the design phase, i.e. from several user stories; our tool generates several diagrams as output: the class diagram, package diagram, and use cases. The tool is carried out by developing prolog language rules allowing the extraction of design elements such as actors and associations between classes, and subsequently refining the associations obtained by using ontology. The created ontology represents the user stories and based on looking for the synonyms of the associations that are found in the ontology. The use of ontology was primordial, firstly, in order to avoid the redundancy of the associations in a generated class diagram, secondly, in order to detect the inclusion between use cases in the use case diagram.

To analyze requirement engineering most of the researchers used the ontology domain to achieve their goal. Our approach combines prolog rules and the domain of ontology. The majority of the researchers have tended to analyze the requirement engineering [1] but in our approach, we start with the extraction of the design elements constituting the UML diagrams, and then we analyze the requirements using the ontology that represents the strong point of our approach. In [7], the authors propose the business process ontology design scheme. The built ontology is considered as a knowledge base by collecting the user stories to reuse them from previous projects. Classes are created in ontology according to Role-Action-Object relations. In [8], the authors used an NLP tool named "OpenNLP Parser" and Wordnet in order to analyze the requirements. Their aim is to extract concepts to constitute a class diagram. The authors developed a desktop tool named "RAPID", the limitation of this tool is that each sentence in the requirements must match a specific structure. In [9], the authors develop a formal Web Ontology Language ontology for the standard representation of engineering requirements. The proposed ontology uses explicit semantics that makes the ontology amenable to automated reasoning. The approach allows the evaluation and classification of engineering requirements. In [10], the author's Approach allows to Test Case Generation Based on Inference Rules. In [11], the authors are built an automated tool named "ABCD" for class diagram generation from user requirements. They used NLP techniques to extract class diagram concepts and generate an XMI file representing a class diagram. The limitation of their tool is that the system does not focus on the problem of concept redundancy. In [12], the authors have developed a framework

to automate the documentation by elaborating the ontology. 60 percent is a percentage of their automation. In [5], a conceptual model is generated from a set of user stories, their tool named visual narrator, this tool does not extract attributes of entities in the conceptual model, and they focus on detecting entities and relationships. In [13], the conceptual model is generated from an unrestricted format such as general requirements, user stories, or use cases; but the attribute extraction rule is based on a set of previously designed verbs. In [14], the searchers suggest an approach that generates a class diagram from use case specifications, parts of speech tags (POS tags), and typed dependencies (TD), were used to reach their objective, however, the developed tool analyses simple sentences. The rules used to extract attributes are not valid in most sentence structures, due to the failure of consecutive names processing. In [6], the authors used the NLP tool named TreeTagger analyzer and developed a JAVA plugin to generate the use case diagram from the user stories; their tool does not handle sentences containing compound nouns. Also; it does not support inclusion or exclusion relationships between use cases. In [15], the authors analyze the requirements by combining the ontological model with prolog rules. This analysis relies on tracing the requirements, eliminating duplicate requirements, and identifying conflicting requirements. The authors used agile requirements. In [16], an NLP-based tool is implemented to generate an Entity Relationship diagram from requirement specification. The machine-learning module is implemented by using a supervised learning mechanism. In [17], through a linguistic analysis of sentence structures and action verbs in user stories, the authors discover patterns of labeling refinements. The refinement goal is a transformation of User Stories into Backlog Items. In [18] the authors propose a technique to automatically transform textual user stories into visual use case scenarios.

III. AN APPROACH TO EXTRACT DESIGN ELEMENTS AND ANALYSE RELATION BETWEEN THE CLASSES AND USE CASES

In our previous approach [19], our objective was to define the extraction rules of the object-oriented design elements, such as actors, classes, attributes, operations of classes, and associations. These components were essential to generate a class diagram, presented in an XMI file and also in a PNG image. We used a natural language processing (NLP) tool named "Stanford CoreNLP" and python language to achieve our goal. After extraction of associations, we have used Wordnet to delete the redundancy associations between the two classes. The use of Wordnet was not sufficient to avoid Redundancy that's why we have thought of another approach that integrates artificial intelligence materialized firstly in the use of prolog language for the definition of production rules to generate associations. Secondly in the use of requirement ontology in which we have defined synonyms of verbs presenting associations. The ontology is created in Protégé editor.



US: user stories; Ass: Association

Fig. 1. Architecture of the Proposed Approach.

All treating was done in python language, even access to ontology to search for synonyms. Prolog language is used firstly to define the Production rules for extraction of the design elements. Secondly to define rules to detect errors in association extraction. The output of our framework is three diagrams: class diagram presented in XMI file, use case, and package diagrams presented in a PNG image. This image is carried out by using Plant UML.

The processing of a text in user stories goes through several steps: Splitting, Tokenizing, POS, Lemmatization, and typed dependencies. The user stories analysis was done using the NLP tool named Stanford core NLP. Fig. 1 shows the architecture of the proposed approach.

A. Prolong Rules for Extracting Relations

To extract the design elements which constitute the class diagram, from a set of user stories, we followed the steps described in the algorithm presented below:

Algorithm: Design elements extraction

- 1: Procedure (Stories S, Actors A, Classes C, Rels R, Attributes ATT, Operations Op)
- 2: for each s in S
- 3: p=POS(s)
- 4: Word_tokenize(s)
- 5: Dependency_parse(s)
- 6: Extract_Nouns(s)
- 7: Extract_Verbs(s)
- 8: A=Extract_Actor(s)
- 9: C=extract_class(s)
- 10: R= Extract_all_relationships(s) [prolog rules: Association(X,Y,Z); X is association name, Y and Z are classes]
- 11: Comp= filtrate_composite_rel (R)
- 12: for each c in C

13: if c in comp then ATT=extract_attribute (Comp)

14: for each r in R

15: If ATT in r

16: Op=r

Based on rules previously defined in [19], we have defined a production rules written in prolog language to extract actors, classes, and associations which connect classes (lines 8-10). The facts are provided from NLP tool that provide the nouns, the verbs and typed dependencies (lines 3-7). To extract attributes of classes we have followed the same approach of [19] i.e. from the resulting classes we do a refinement; some classes become attributes and thereafter some associations become operations of a class (lines 10-16).

The rules are presented in this form: Association(X, Y, Z); The objective of these prolog rules is to extract X which represents the association name, and Y and Z which are classes in the UML class diagram.

B. Prolog Rules for Detecting Errors in Relation Extraction

After extraction of association between classes, we proceed to the refining step by applying some prolog rules which detect errors in the list of association.

Rule1: if two or more actors have the same action to execute.

Rule2: if there is an association between A class and B class and the same association between B class and C class then there isn't an association between A class and C class. This rule avoids transitivity between associations which can clutter the class diagram with several unsuccessful relations.

C. Ontology for Analysis of Relations between the Classes in Class Diagram and use Cases in use case Diagram

Our ontology is important to represent knowledge of the application domain and to identify the relations between requirements such as composition or synonyms. USOn is an ontology that describes an agile requirement; we have created ontology classes and instances through Protégé ontology editor.

Fig. 2 shows the hierarchy of the ontology USon. Table I presents the description of some classes.

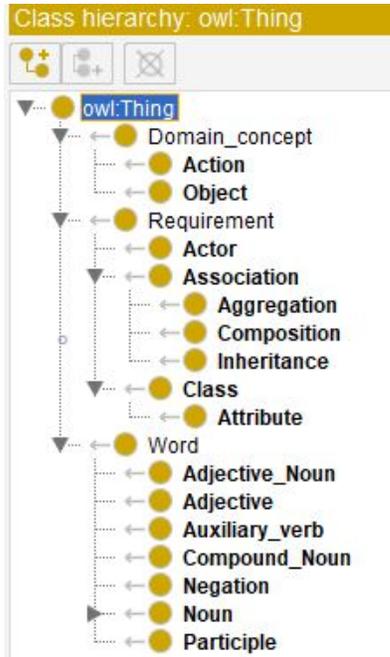


Fig. 2. Hierarchy of the Ontology USon.

TABLE I. DESCRIPTION OF SOME CLASSES

Class name	Description
Action	class whose elements are verbs which represent the associations in class diagram
Object	class whose elements are nouns which represent the classes in class diagram
Actor	Class whose elements are nouns which represent the role in user story (As role,...). The actor is who perform the action. Actors are present in the use case diagram.
Association	Symantec link between classes in the class diagram
Class	class whose elements are nouns which represent the classes in class diagram
Attribute	class whose elements are nouns which represent the attributes of classes in class diagram
Word	class whose elements are tokens which represent part of user story

We have defined a set of Synonyms to *Action* class in order to refine association name. The same process is for the *Actor* class. Fig. 3 shows an example of defined synonyms and inclusion action.

Our tool accesses the *USon* ontology in order to compare the names of the associations obtained using the prolog rules with those defined as synonym of the *Action* class. Subsequently, the associations will be refined. The refinement of actors in use case diagram is done by browsing synonyms of Actor instances.

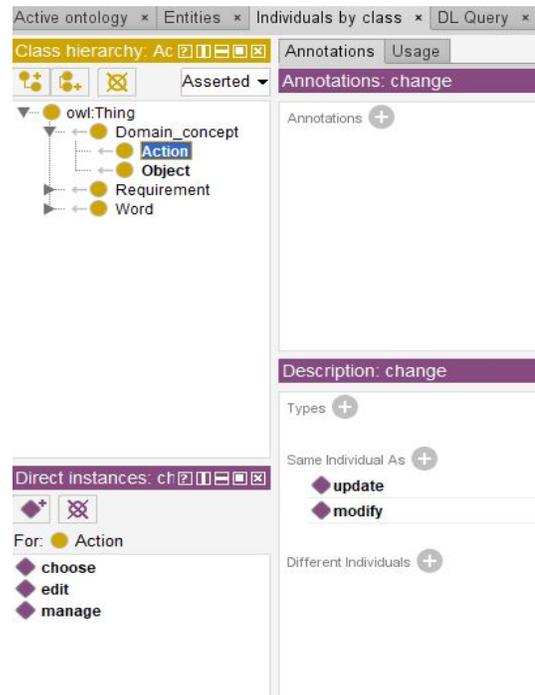


Fig. 3. Synonyms of Change Action.

Consider an example of user stories:

- US1: As a manager, I want to manage account of users.
- US2: As a manager, I can create a new account.
- US3: As a user, I can modify login account.
- US4: As a user, I can update my login.

In US1 the action is the verb “manage”, according to our approach the following association is extracted: Manage (manager, account).

In US2 the action is the verb “create”, according to our approach the following association is extracted: Create (manager, account).

Our tool, at first looks for the relationships between the same classes as in US1 and US2, and US3 and US4, after Wordnet is used in order to avoid redundancies, then browsing of *USon* ontology is mandatory to detect synonym and inclusion relations between use cases; in the example, the actions modify and update are synonyms as shown in Fig. 3 so an association will be removed from the list of associations in order to avoid duplicate associations.

In *USon* Create is part of Manage as shown in Fig. 4, then there is an inclusion between two use cases extracted from US1 and US2: “manage account” and “create account”.

Consider these user stories:

- As a user, I can change the account information.
- As a user, I am able to edit account information.

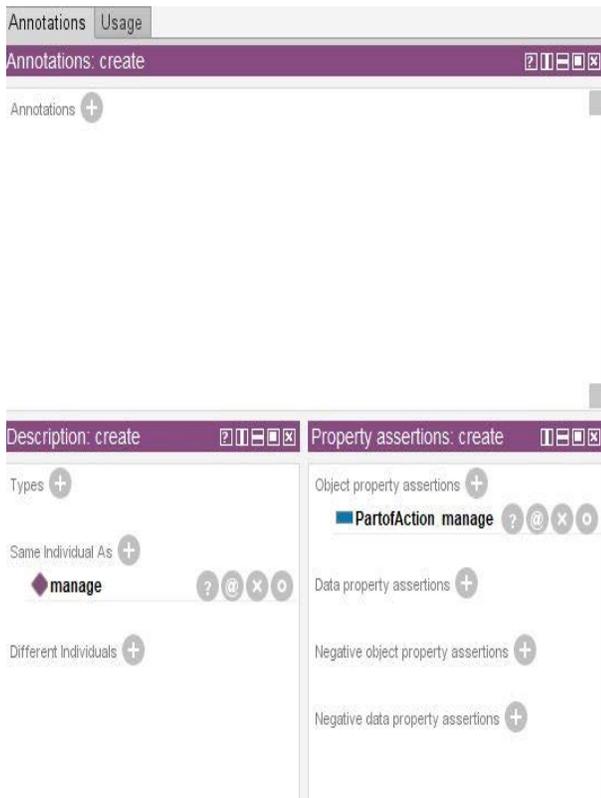


Fig. 4. Composition of Manage Action.

Our tool with the help of the Stanford NLP tool and prolog rules allows extracting two use cases: “change account information” and “edit account information”.

According to the ontology USon, the update action is part of Edit. In Fig. 3 the update action is a synonym of change action then change is part of the Edit action.

We can deduce from this combination between ontology and prolog rules that there is an inclusion between two use cases: “change account information” and “edit account information”.

IV. GENERATION OF UML DIAGRAMS

A. Generated Class Diagram

After extracting the design elements, the next task was to regroup these elements to constitute the class diagram. The developed tool generates an XMI file which is an Ecore file. Ecore file is the Eclipse Modeling Framework (EMF) meta-model, which illustrates the names of the classes, their attributes, and their types, as well as the methods and relationships with their classifications. Also, PlantUML API is used to visualize the class diagram. These treatments were done in python language. To implement our new approach, we used the same case studies^{1,2} from article [19] and compared the results. We found that in our old approach, the class "people" (case study¹ number 2) was detected, yet in our approach; there is an association of inheritance between "people" and all the

actors thanks to our ontology. This association has been added to the generated class diagram.

Regarding the first case study², there is redundancy in the operations obtained (filtrate (type) and choose (type)) which are at the origin of associations before refinement. Wordnet could not detect that these verbs are equivalent, so we used the ontology.

We noted that the associations obtained from the old approach are all obtained using the extended rules of our second approach.

B. Generated Package and use Case Diagrams

A package diagram offers many advantages to designers who want to create a graphical representation of their UML system or project. This diagram simplifies the complex class diagram into a tidy visual form. In our case, we used the package diagram to organize the class diagram.

After generating the class diagram, the next task was to generate the package and use case diagrams. To do this, we based on the design elements already extracted such as: classes, associations, and actors.

To extract a package, we first use the associations that link the actor and another class, and secondly, we add the term "manage" before each class to form a package. To detect the dependency between the packages, we take into consideration the relationship between the classes that make up the packages. The use cases are formed from the associations between classes provided that one of the classes is an actor. PlantUML API is used to visualize the package and use case diagrams. All treatments were done in python. Fig. 5 shows the generated package diagram of the case study¹ which represents inline course management: videos, quizzes, and others.

The generated package diagram is based on associations and classes of the class diagram. The red arrows between the packages represent the dependencies between them. Table II represents some use case diagrams for each package presented in Fig. 5.

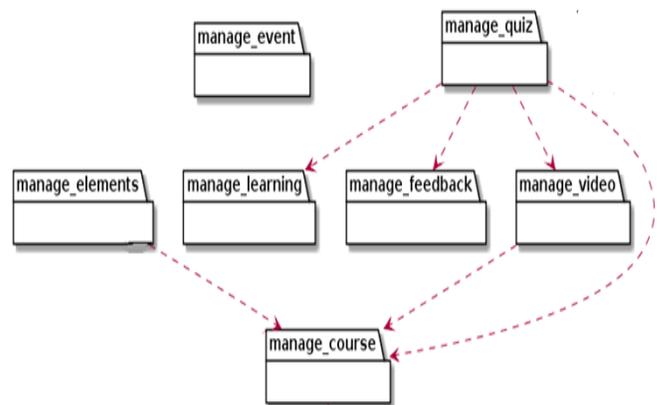


Fig. 5. The Generated Package Diagram.

¹ https://drive.google.com/file/d/1wk8yCa9wS12ooEwR0UjWDsK0_b4k aKG / view?usp=sharing

² https://github.com/MarcelRober/StoryMiner/blob/master/example_storie s.txt

TABLE II. USE CASE DIAGRAMS OF PACKAGES

Package name	Use case diagrams
Manage Course	
Manage Element	
Manage Quiz	
Manage Event	
Manage Video	

By comparing the results obtained manually with those which are automatic, we noted that our approach extracts 99% of the relationships. Generating the package and use case diagrams based on the associations of the class diagram has revealed its effectiveness.

Regarding the class diagram, the Uson ontology allows firstly detecting inheritance association between actors such as actor named People and other actors, secondly determining the synonyms of associations.

Regarding the use case diagram, the Uson ontology allows firstly determining the synonyms of use cases, secondly the inclusion relationships between two use cases.

Our approach remains very effective thanks to the strong point of the combination of the domain of ontology and prolog rules. We can say that the ontology we created complements the prolog rules in order to obtain better results.

Table III shows a comparison between my old approach [19] and my proposal. However, Table IV depicts a comparison between related work and my proposed.

Our approach is the unique method that defined extraction rules for associations of class diagram using prolog language. Subsequently, the association and use cases are analyzed and refined using our built ontology named "Uson". The associations are the key for building the UML diagrams: use case and package diagram.

TABLE III. COMPARISON BETWEEN MY OLD APPROACH AND MY PROPOSAL

	Association		Composition relationship	Inheritance relationship	Synonyms	Output
Old approach [19]		Python heuristic rules	Python heuristic Rule s: H1, H2, H3	Python heuristic H4 and H5	Wordnet: Synonym of associations (verb which link two classes)	Class diagram: XMI file
Proposal approach	NLP tool: Stanford coreNLP	- python language - Prolog rules	- Prolog rules - Ontology USon	Ontology USon	- Wordnet - Ontology USon: 1. synonym of use cases (action in use case) 2. synonym of associations	- Class diagram: XMI file and PNG image (using plant UML) - Package diagram: PNG image(using plant UML) - Use case diagram: PNG image(using plant UML)

TABLE IV. COMPARISON BETWEEN SOME RELATED WORK AND MY PROPOSED

Related work	Input	Output	Approach and tool
[7]	User stories	Business process ontology	Reuse of user stories
[8]	Business document	Concept-classes	- Wordnet - Linguistic rules
[9]	Engineering requirements	Formal Web Ontology	Evaluation and classification of Engineering requirements
[5]	User stories	Conceptual model without attributes	Visual narrator tool
[14]	Text requirement and user stories	Conceptual model	Visual C # language and Stanford CoreNLP
[15]	User stories	- Elimination of duplicate requirements, and identification conflicting requirements - not any generation of the UML diagram or conceptual model	- Prolog and python language - Ontology
[19]	User stories	Class diagram	Python language and Stanford CoreNLP
My proposal	User stories	- Class diagram - Package diagram - Use case diagram	- Standford coreNLP - Python and prolog language - Ontology

V. CONCLUSION

This paper have proposed an approach to automate the analysis phase in an agile context, to extract the design elements which are essential to constitute the generated UML diagrams: the class diagram, the package and use case diagrams.

Our approach is based on the combination of prolog rules and an ontology which present the user stories. The prolog rules used dependencies offered by Stanford core NLP. This combination is the strong point of our approach. The main advantages of the proposed technique are:

- Improvement of the results obtained from our previous approach by Applying artificial intelligence presented in prolog rules and ontology.
- Generation of three UML diagrams which facilitate the design of analytical tasks in the team.
- Refined classes have been obtained following a transformation of some classes into attributes using composition relationships, and some relationships to operations.
- Definition of prolog rules for detecting errors in relation extraction.

Our proposed approach is very useful to ease the analytical tasks in the design team. Next, minimize time and costs. The benefits of our approach are the utilization of agile requirement to automate them, these requirement named user stories are the best way to describe the engineering requirement. In the future, our work will be completed by generating user interfaces and code of the software.

REFERENCES

- [1] D. Pandey and V. Pandey, "Requirement Engineering: An Approach to Quality Software Development," Journal of G--lobal Research in Computer Science, vol. 3, no. 9, pp. 31-33, 2012.
- [2] M. Cohn, User Stories Applied: for Agile Software Development. Redwood City, CA, USA: Addison-Wesley Professional, 2004.
- [3] Y. Wautelet, S. Heng, M. Kolp, and I. Mirbel, "Unifying and extending user story models," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8484 LNCS, pp. 211–225, 2014, doi: 10.1007/978-3-319-07881-6_15.
- [4] I. K. Raharjana, D. Siahaan, and C. Fatchah, "User Stories and Natural Language Processing: A Systematic Literature Review," IEEE Access, vol. 9, pp. 53811–53826, 2021, doi: 10.1109/ACCESS.2021.3070606.
- [5] G. Lucassen, M. Robeer, F. Dalpiaz, J. M. E. M. van der Werf, and S. Brinkkemper, "Extracting conceptual models from user stories with Visual Narrator," Requir. Eng., 2017, doi: 10.1007/s00766-017-0270-1.
- [6] M. Elallaoui, K. Nafil, and R. Touahni, "Automatic Transformation of User Stories into UML Use Case Diagrams using NLP Techniques," in Procedia Computer Science, 2018, vol. 130, pp. 42–49, doi: 10.1016/j.procs.2018.04.010.

- [7] C. Thamrongchote and W. Vatanawood, "Business process ontology for defining user story," 2016 IEEE/ACIS 15th Int. Conf. Comput. Inf. Sci. ICIS 2016 - Proc., pp. 1–4, 2016, doi: 10.1109/ICIS.2016.7550829.
- [8] P. More and R. Phalnikar, "Generating UML Diagrams from Natural Language Specifications," Int. J. Appl. Inf. Syst., vol. 1, no. 8, pp. 19–23, 2012, doi: 10.5120/ijais12-450222.
- [9] A. Mukhopadhyay and F. Ameri, "An ontological approach to engineering requirement representation and analysis," Artif. Intell. Eng. Des. Anal. Manuf. AIEDAM, vol. 30, no. 4, pp. 337–352, 2016, doi: 10.1017/S0890060416000330.
- [10] H. Tan, M. Ismail, V. Tarasov, A. Adlemo, and M. Johansson, "Development and Evaluation of a Software Requirements Ontology," SKY 2016 - 7th Int. Work. Softw. Knowledge, Proc. - conjunction with IC3K 2016, pp. 11–18, 2016, doi: 10.5220/0006079300110018.
- [11] W.B. A. Karaa, Z. B. Azzouz, A. Singh, N. Dey, A. S. Ashour, H. B. Ghazala, "Automatic Builder of Class Diagram (ABCD): an Application of UML Generation From Functional Requirements," Journal of Software Practice and Experience, vol. 46, no.12, pp. 1443–1458, 2016, doi: 10.1002/spe.2384.
- [12] M. P. S. Bhatia, A. Kumar, and R. Beniwal, "Ontology driven software development for automated documentation," Webology, vol. 15, no. 2, pp. 86–112, 2018.
- [13] M. Javed and Y. Lin, "Iterative process for generating ER diagram from unrestricted requirements," ENASE 2018 - Proc. 13th Int. Conf. Eval. Nov. Approaches to Softw. Eng., vol. 2018-March, no. Enase, pp. 192–204, 2018, doi: 10.5220/0006778701920204.
- [14] J. S. Thakur and A. Gupta, "Automatic generation of analysis class diagrams from use case specifications," 2017.
- [15] M. S. Murtazina and T. V. Avdeenko, "Requirements analysis driven by ontological and production models," CEUR Workshop Proc., vol. 2500, pp. 1–10, 2019.
- [16] P. G. T. H. Kashmira and S. Sumathipala, "Generating Entity Relationship Diagram from Requirement Specification based on NLP," 2018 3rd Int. Conf. Inf. Technol. Res. ICITR 2018, pp. 1–4, 2018, doi: 10.1109/ICITR.2018.8736146.
- [17] L. Müter, T. Deoskar, M. Mathijssen, S. Brinkkemper, and F. Dalpiaz, "Re_nement of user stories into backlog items: Linguistic structure and action verbs," in Requirements Engineering: Foundation for Software Quality (Lecture Notes in Computer Science), vol. 11412. New York, NY, USA: Springer, 2019, pp. 109_116.
- [18] F. Gilson, M. Galster, and F. Georis, "Generating use case scenarios from user stories," in Proc. Int. Conf. Softw. Syst. Processes, Jun. 2020, pp. 31–40, doi: 10.1145/3379177.3388895.
- [19] S. Nasiri, Y. Rhazali, M. Lahmer, and N. Chenfour, "Towards a Generation of Class Diagram from User Stories in Agile Methods," Procedia Comput. Sci., vol. 170, pp. 831–837, 2020, doi: 10.1016/j.procs.2020.03.148.

Law Architecture for Regulatory-Compliant Public Enterprise Model: A Focus on Healthcare Reform in Egypt

Alsayed Abdelwahed Mohamed¹, A.Abdo³

Faculty of Computers and Artificial Intelligence
Helwan University, Egypt^{1,3}

Faculty of Computing, Arab Open University in Egypt³

Nashwa El-bendary²

Arab Academy for Science, Technology, and Maritime
Transport, College of Computing & Information
Technology, Aswan - Egypt

Abstract—Public business operations are governed by a set of legal sources, which regulate their implementation under administrative laws that are increasingly influencing software system design and development. Enterprise Architecture (EA) is a critical approach for driving business and Information Systems (IS) transformation in the public sector. On the other hand, EA frameworks lack representation schemas that support law models. Understanding of the law Architecture in the government domain is required for EA work and is thus the first architecture activity that must be completed. As EA approaches for Law compliance reviews are performed by legal experts, there is a gap between law experts and technical system architects. To cover these gaps, this paper proposes a novel framework for analyzing the administrative laws, extracting the legal policies and legal rules, identify their relationships with other EA domains, and identifying the law compliance requirements. Moreover, the integration of our proposed law architecture framework with existing EA frameworks to reach a law-compliant public enterprise model is identified. Finally, the applicability of the proposed framework is shown and validated through a case study. Moreover, subject matter experts of the legal domain also evaluated the extracted legal policies and rules during the implementation of our proposed framework.

Keywords—Law architecture; regulatory compliance; requirements engineering; enterprise architecture; law ontology; TOGAF

I. INTRODUCTION

Administrative laws are the law acts that regulate the operations of government agencies. The branch of law governs administrative agencies' development and operation. Administrative law, commonly known as (regulatory law), includes the rules and regulations that an administrative body promulgates and enforces. Administrative laws are compulsory for the institutional domain and are a benchmark for officials and architects in government and government digital transformation initiatives. Enterprise architecture (EA) is the method of aligning an authority's business with information technology, including the convergence of business procedures, Information Systems (IS), technology, and people[1]. EA has been in progress and deployment for the past years[2]. However, most authorities continue to have issues with EA practices; EA practice is not a simple task. Despite a large

number of current EA frameworks, public authorities have been unable to translate EA solutions to meet their needs [3].

In the study [1], the Gartner Group stated that a big percentage of EA implementations fail in the world because they start with modeling rather than defining business requirements first where the business requirements must meet the needs of IS. The business needs of the public agencies should comply with laws. Regulatory or law compliance management is a general concept that encompasses all practices and procedures used to ensure that a public authority adheres to all legal rules mandated by local or international laws. These laws are typically outlined in a natural language text that is difficult to comprehend for non-legal professionals. Existing laws pose a different set of problems. New digital transformation solutions may not be properly tested until they are put into operation and laws are developed to fix previous issues caused because of business and social shifts. To overcome the law compliance problem, several research studies have been conducted. Some researchers propose a law ontology approaches as a solution for law compliance. Some other researchers proposed regulatory compliance frameworks with the business process [4]. Moreover, other studies were conducted to regulatory compliance for requirements engineering [5] [6] [7] Enterprise architecture frameworks are used to guide the creation of enterprise models in terms of three domains namely business architecture, IS architecture, and technology architecture. Thus, extending the enterprise model to represent the policies identified by law can be obtained by defining the law architecture (our proposed framework) and integrating it with other enterprise architecture domains, as illustrated below in Fig. 1.

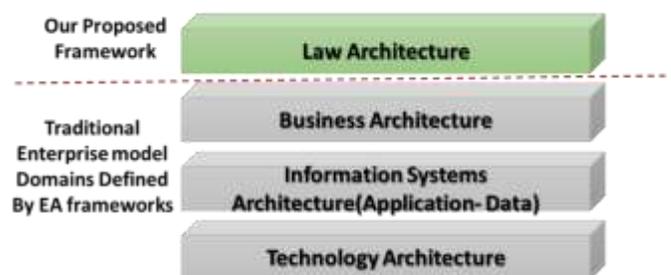


Fig. 1. Public Enterprise Model View after Integrating the Proposed Framework with EA Frameworks.

Knowledge of the law Architecture in the government domain is a prerequisite for architecture work. The law Architecture is also often necessary as a means of defining the governance layer and regulatory compliance requirements of following EA activities for all enterprise model domains.

Many emerging approaches to e-Government architecture, on the other hand, do not take legal sources into account. There is no formal structure for e-Government enterprise model enforcement with legal sources in these approaches. Compliance with regulatory issues and e-Government business models is a critical problem for governments. Thus, this study proposes a novel framework to address the following questions:

- How to manage compliance with laws across all enterprise model domains; business, IS, and Technology?
- How to design Government IS that complies with administrative laws?

- How to transfer laws policies into a machine-readable format?

In the next section, we present a state-of-the-art review of the related works. a detailed description for the proposed law architecture framework components and method including the integration with EA frameworks will be presented in Section III. Section IV presents the applicability of the proposed framework to real-life cases; the case study of the Egyptian Universal Health Insurance program is demonstrated. Finally, in Section V, a brief analysis of the obtained conclusions and future work are discussed.

II. STATE OF THE ART

Law compliance or regulatory compliance is a concept of acting in accordance with established laws, regulations, etc. In order to judge that a public authority is following the given legal rules or not, compliance auditing is conducted.

TABLE I. COMPARISON OF SURVEYED STATE-OF-THE-ART OF REGULATORY COMPLIANCE

Study		Extraction of legal rules	Ontology representation	Integration with enterprise model	Meta-model	Requirements engineering for SW	Legal modeling
Business process compliance	Jörg Becker,2015 [1]	√				√	√
	Christina Stratigaki ,2016 [2]	√			√		√
	Noel Peter,2017 [3]	√					
	Selja Seppala ,2017 [4]	√	√				√
	Hashmi Mustafa,2018 [5]	√					√
	Cesare Bartolini ,2019 [6]	√	√		√	√	√
	Raimundas M. , 2020 [7]	√			√		
	Tobias Seyffarth, 2021 [8]				√	√	
Law Ontology	Mirna El Ghosh,2017 [9]	√	√				
	Monica Palmirani,2018 [10]		√		√		√
	Galina Kurcheeva,2019 [11]	√	√		√		
	Valentina Leone, 2020 [12]	√	√				
Requirements Engineering	Ftemeh Zarrabi,2015 [13]		√		√	√	
	Ana Zalazar,2017 [14]					√	
	Sushant Agarwal,2018 [15]	√			√	√	
	Mahmood Alsaadi,2019 [16]	√				√	
	Michael Felderer,2020 [17]					√	
The proposed framework		√	√	√	√	√	√

A compliance review is an important activity of compliance auditing which depends on the extraction of regulatory compliance requirement from administrative laws. There have been an increasing number of research challenges to automate regulatory requirements extraction from legal acts.

Administrative laws, which drive the development of regulatory compliance solutions, are generally available as natural language text documents (semi-structured format), and understanding them is performed by legal domain expertise. Testing a given business model for enforcement is, unsurprisingly, a manual process performed by accredited domain experts. Moreover, effective regulatory compliance requires good knowledge of the enterprise model of public authority by the legal auditors and a good knowledge of the compliance framework by the enterprise architects. This makes the close collaboration of legal auditors with enterprise architects a necessity. The challenge of automated compliance checking can be seen as extending an enterprise model to include concepts defined by the regulatory compliance framework and enabling the needed automation in checking an enterprise model against the policies defined in a regulatory compliance framework.

This is because most legal acts are written in natural language, which a conventional computer system cannot understand. A legal ontology is a theory that describes the concepts that exist in the legal acts and how they are connected. Many kinds of research have been undertaken to develop ontological representations for legal documents to achieve law compliance [8] [9] [10] [11]. On other hand, a big number of studies exist for achieving law compliance through linking laws policies with Business Process Models (BPM) [12] [13] [14] [15] [16] [17] [18] [19]. Moreover, Regulatory compliance is a well-studied area in software requirements engineering, including research on how to model, check, analyze, extract, and validate compliance of software different requirements engineering techniques for achieving software compliance as discussed in [20] [21] [22] [23] [24] the below Table I outlines the comparison between different approaches for achieving regulatory compliance. Based on the comparison Table I, we argue that this literature survey could not find any contribution that addresses the topic of how to integrate law compliance solutions with the enterprise model.

III. PROPOSED FRAMEWORK

The proposed Law Architecture Framework (LAF) provides the baseline for mapping various legal policies related to regulatory requirements to business process, stakeholders, capabilities, IS, data concepts, and technology architecture. Regulatory compliance planning and subsequent deployment projects require mapping and tracking the evolution of the following artifacts of law architecture:

- Legal policies tied to business capabilities.
- Legal rules tied to business processes.
- Legal policies and rules tied to data concepts.
- Legal policies and rules tied to Information systems.
- Legal policies and rules tied to Technology.

The law architecture framework is a goal-oriented, law-driven framework aimed to define compliance requirements through which a particular public enterprise and its related information systems can comply with a given administrative law. We define the law architecture as a holistic representation, multidimensional views of legal policies and their derived legal rules, and identifying their relationships with the business, data, information systems, and technology domains, in addition to identifying the requirements of compliance with these rules and policies. Law architecture requires a complete framework to guide its development and practice. We argue that the effort required to accomplish the law architecture activities should be performed by a team that has both bits of knowledge of technology and law domains, a team with the following key roles:

- Law Architect is an information and communications technology professional that is responsible for analyzing administrative laws and extracting from them the regulatory compliance requirements of business, IS, and technology architectures. The law architect must be reasonably knowledgeable of the administrative laws and public sector domains to be familiar with public sector services and the operational environment.
- Legal Counsel: According to the BIZBOK Guide [25]" A Legal Counsel is a person who is legally qualified and licensed to represent an organization in a legal matter, such as lawsuits, policy formations, contract negotiation, and patent activities". A legal counsel regularly works in legal departments of the public authorities and is considered the subject matter expert during the law architecture cycle. The legal council is responsible for the analysis, and validation of extracted policies and rules. The objectives of the law architecture framework can be summarized as follows:
 - To align laws and regulations with business and IT Systems.
 - To identify and model legal policies and rules and linking them with the enterprise model domains namely; business architecture, information systems, and technology architecture.
 - To identify the relationship between legal rules and other architecture domains.
 - To identify the regulatory compliance criteria with the legal rules.
 - To manage any suggested law amendments and assess their impact on the enterprise model.
 - To develop the appropriate viewpoints, models, and matrixes that enable the demonstration of how the legal policies are addressed in the business, IS, and technology architectures.

As shown in Fig. 2, the law architecture framework consists of the following components:

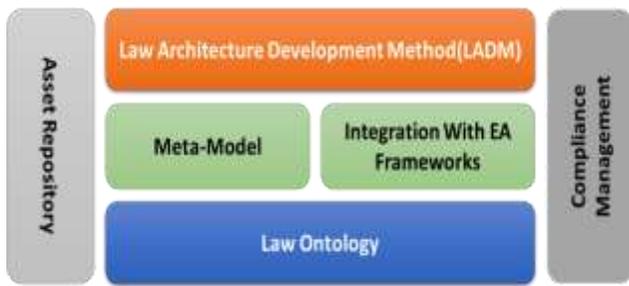


Fig. 2. Components of the Proposed Law Architecture Framework.

- Law-Architecture-Development-Method (LADM): describes a set of guidelines and activities for developing a law Architecture.
- Law Architecture Meta-model: defines a formal structure of core concepts and their relationships to ensure consistency within the law architecture.
- Asset Repository: outlines different taxonomies and guidelines to categorize and store the inputs and outputs of law-architecture effort such as laws catalogs.
- EA framework integration: describes the ability to integrate the law architecture framework with other enterprise architecture frameworks.
- Law Ontology: defines an ontological and semantic representation of the law architecture Meta-model for getting a machine-understandable model that is enabled easily for automation and model analysis.
- Regulatory Compliance Management: defines different approaches for checking a public enterprise model against the legal policies and rules defined within the LADM.

The details of some selected components of the proposed framework are detailed as follow:

A. Law-Architecture-Development-Method (LADM)

The LADM is the product of the continuous efforts of law architects and legal counsels. It describes a method for modeling and managing the compliance requirements of laws that govern the public enterprise and forms the core of the Law architecture framework. It describes the steps, guidelines, and tools to analyze, extract, and model the legal policies and rules from administrative laws. Moreover, it guides the definition of the correlation between extracted legal policies and rules with other public enterprise model domains, to align the public authority and its IT systems with the law.

Fig. 3 depicts the phases of the proposed LADM which consists of the following:

- Law Architecture Initiation.
- Phase (A): legal Policy Mapping.
- Phase (B): Legal Rule Mapping.
- Phase (C): Definition of legal Policies/ Business architecture correlation.

- Phase (D): Definition of legal Policies / IS architecture correlation.
- Phase (E): Definition of Legal Policies / Technology architecture correlation.
- Management of Suggested Law amendments.

The level of detail addressed in Law Architecture will depend on the scope and objectives of the overall law-architecture effort. The key steps in details of the LADM include the following:

1) *Law architecture initiation*: The Objectives of this phase include initiating the law Architecture Capability planned by the public authority, defining the law Architecture Principles, and Identifying the set of administrative laws that applies to the business performed by a public organization which is related to the IT systems' Scope. The LADM is a generic approach that can be used by a broad range of public bodies and, if necessary, in combination with a wide range of EA frameworks. The recommended steps within the initiation phase are as follows:

a) *Determine the Scope of the Public Agency Impacted*: Identify the public entity core units that would be most affected or gain the most benefit from the law architecture work.

b) *Define and Establish the Law Architecture Team*: The required team members that will develop the law architecture should be identified.

c) *Define Law Architecture Principles*: A principle is an agreed-upon fact that can guide one's analysis and reasoning. Core principles that apply to law architecture, in general, are:



Fig. 3. Law-Architecture-Development-Method (LADM) Phases.

- Legal policies are accessible and shared.
- Common Vocabulary and legal policy Definitions.
- Law architecture's scope is defined by a set of laws applied to the public authority.
- The law architecture process is iterative.
- Law architecture artifacts are reusable especially those extracted from domain-specific laws.

d) *Develop Law Catalogs*: Catalogs form the raw legal documents for the development of models and views and represent a key artifact for driving the law architecture activities. The following catalogs should be developed:

- Project-Related Acts Catalog.
- Domain-Related Acts Catalog.
- Legal Policy catalog.
- Suggested Amendments Catalog.
- Ontology Representation.
- Suggested Amendments Catalog.
- Regulatory Compliance Requirements.
- Approved Amendments Catalog.
- Suggested Amendments Catalog.
- Compliance Requirements Catalog.

2) *Phase (A) Perform legal policy mapping*: The objectives of this phase are analyzing architecture scope impacted acts and extract the legal policies representations, Reviewing extracted policies with different legal experts and ensuring their correctness and consistency, and finally updating the legal policy catalog. Administrative laws define the legal limits for public authority by enforcing requirements that specify which actions are legal and which must be prohibited during the conduct of daily operations. A legal policy is a statement or set of statements that describe the course of events and how a specific action should be carried out. A policy is defined to achieves a business goal. A legal policy has a subject, who is the stakeholder(s) who is/are responsible for carrying out the policy's actions. A legal policy requires a business capability to be executed, a business capability is defined as a set of resources and activities that integrate to deliver a value to an internal or external stakeholder (target). Legal Policy mapping is the core of law architecture and an excellent starting point. Describe the policies and legal rules associated with the law articles, the main idea in the law architecture is that legal policies define decisions on whether a state of the business activity is allowed or not. The extraction of legal policies and rules is a hardworking and time-consuming activity and is recommended to be carried out by legal architects (or business analysts) with the help of legal counsels. To extract the legal policies the architects must look for important information by

scanning the legal articles and try to answer the following questions:

- Why was the article written? The answer to this question identifies the objectives that the policy intends to achieve.
- What rules/actions intended by the policy? And what are the required capabilities required to achieve these rules?
- Who responsible for executing these rules/actions and who will be affected by them?
- When are these rules/actions allowed to be performed?
- Where these rules/actions are applicable? By answering when and where questions the event and scope of the legal policy can be identified.

A legal Policy can be classified as follow:

a) *Business Function*: Identified by article(s) that defines a course of business activities that should be performed in a predefined manner to achieve some business outcomes.

b) *Organization Unit*: Mapped by an article(s) that define an organizational structure and responsibilities.

c) *Technical Function*: Mapped by an article(s) that directly describe an IT resource or system.

Legal Policy Mapping Approaches include the following:

a) *Top-Down Approach*: The top-down approach to policy mapping is performed by scanning and analyzing the administrative law starting from the first article until the last article. For each article try to answer the five whys' questions listed above to extract the Legal policies essential to regulate the operation and services provided by the public enterprise.

b) *Goal-Oriented Approach*: the legal councils identify the main objectives of each administrative law, and for each goal, the team scans the law and extracts the related policies.

3) *Phase (B) Perform legal rules mapping*: The objectives of this phase are analyzing legal policies and laws then extract the intended legal rule, reviewing and validating extracted legal rules with different legal experts and ensuring their correctness and consistency, and finally, update the legal rule catalog. To perform legal rule mapping properly the following concepts should be considered:

- A legal rule is derived from one or many legal policies, in other words, many legal rules can be sequentially implemented to achieve a certain legal policy. A legal rule is also has a business goal and has a scope. A legal rule requires input data to process and has an input event that triggers the execution of the legal rule.
- A legal rule contains an action that is implemented in a business process usually in a form of a business decision.
- A constraint is an assertion that, by one of its properties or a relationship to another concept, such as as a role in

an activity, restricts or reduces the potential representation for a concept. For example, the statement “a beneficiary who under-poverty-line” constrains the set of all possible beneficiaries to the possibly smaller set of only those beneficiaries who are under the poverty line.

- A stakeholder is an entity afforded rights and/or obligations by the law.
- A right refers to one or more actions that deliver a defined stakeholder a certain value or benefit. If a stakeholder is not obligated to perform an action or to receive an action value, called an anti-obligation.
- An obligation refers to one or more actions that are required to be undertaken by a certain stakeholder.
- An anti-right refers to one or more actions that are not permitted to be performed by a certain stakeholder.
- A rule’s action is either a right, obligation or anti-right per our above definitions.

4) *Phase (C) Identify the correlation between extracted policies and business architecture:* Legal policies represent the governance layer that organizes the business functions, therefore linking legal policies with their related capabilities, processes, organization units, and stakeholder will simplify the alignment of these business objects with laws and will result in a law-compliant Information system when these business artifacts mapped to IS. In other words, will help the administrative entity to deliver the right service to the right citizen at the right time. Moreover, will enhance the efficiency and transparency of the public institutions. The correlation between legal policies and business architecture can be illustrated through the analysis of the legal policies catalog and developing the following matrixes:

- Legal policy Mapping to Business Capabilities describes.
- Legal Rule Mapping to Business Process.
- Legal policy Mapping to Organization Structure.
- Legal Rule Mapping to Stakeholder.

5) *Phase (D) Identify the correlation between extracted policies and Information Systems(IS) architecture:* Information systems designed to support business processes in public authorities must conform to the relevant legal acts and regulations. The correlations between business architecture and legal policies should be transferred into technical system requirements and IS should be designed to comply with these requirements. IS architecture is consists of data architecture and application architecture. After identifying the business regulatory compliance requirements through phase (C), the following steps should be performed in this phase to design, a law-compliant IS in a later detailed IS architecture definition cycle using any EA frameworks. Typical dimensions of the

correlations between extracted legal policies and IS architecture include the following:

a) *Define Data Concepts:*

- Extract data concepts from the legal rules.
- Extract the data modalities from the constrains.
- Classify the data concepts (internal or external).

b) *Define Role/Application Function Matrix:*

- Each participant stakeholder with the legal rules should be mapped to an application role.
- Each action assigned to a legal rule should be mapped to an application function.

c) *Define Application Interaction Matrix:*

- Stakeholders and data concepts defined in the legal rule catalog should be classified into internal or external (of the scope of the public authority).
- Each external concept defines a possible integration requirement with external systems.

6) *Phase (E) Identify the correlation between extracted policies and technology architecture:* Technology Architecture consists of software and hardware assets that are used to implement and realize information system solutions, legal policies, and rules that have a direct impact on technology architecture. The correlation between legal policies and technology architecture can be identified as follow:

a) *Availability of the Systems:* Based on the scope and type of services identified by legal policies the availability of the systems can be identified for example the availability of health services should be 24/7, the technology services should grantee the operation of the systems 24 Hours in 7 days each week. The availability of some other public services mandated by law could be only within the official working hours or for a certain time like examination systems or submission of tax declaration requests.

b) *Data Archiving and Retention Plans:* Data retention refers to the storage of transaction data for a certain time. The primary objective in government data retention is traffic analysis and judicial arbitration. Each country has its data retention policies and plans that are defined in its administrative laws.

c) *Network and Communications:* Based application interaction matrix defined in phase D, the requirement of communication lines between the public authority and external entities should be identified, the detailed specifications of these lines can be identified later in the detailed technology architecture cycle of EA practice. The scope of the legal policy also can identify the locations where the required IS should be accessed, thus also implies communication lines requirements.

d) *Security Requirements:* The access rights for data assets can be defined based on the Roles/ application function matrix developed in Phase C.

7) *Phase (F) Collect and communicate the suggested law amendments:* The object of proposing new law amendments is to improve the effectiveness, efficiency, and fairness of the administrative processes. Architecture practice implemented in Law, business, data, IS and technology phases may propose new procedures designed to afford more time for this function. It is hoped that delays and costs will be reduced and that new prestige will be provided for those who make the decisions. Suggested Law amendments should be handled as follow:

- a) Build the Suggested amendments Catalog.
- b) Update the Catalog with any suggested law amendments during architecture phases (Law, Business, data, IS, and Technology).
- c) Prioritize Law amendments and assess their impact.
- d) Communicate the amendments to the relevant Stakeholders.

B. Law Architecture Meta-Model

In the previous section, a method to create and manage the law architecture artifacts within the government enterprise was defined. At each step, a discussion of architectural work and artifacts was presented such as the legal policy map and legal rule map. The content Meta-model provided here defines a formal structure for the core concepts of the law architecture and their relationships to ensure consistency within the law architecture and to guide the implementation of the law architecture within a public enterprise. Fig. 4 depicts the proposed law architecture Meta-model.

C. Integrating Law Architecture with EA Frameworks

In the context of digital transformation in the public sector, the integration between law architecture and EA frameworks is a natural architectural alignment of two related disciplines. Law architecture represents the mapping of legal policies and rules related to IT architecture while EA provides a guiding framework for business and IT architecture alignment. According to the Federation of Enterprise Architecture Professional Organizations, EA “represents the holistic planning, analysis, design, and implementation for the development and execution of strategy by applying principles and practices to guide organizations through the integration and interoperation of all other architecture domains”[26]. Most EA approaches do agree on the foundational domains, which include:

- Business Architecture.
- Information Systems (Application- Data) Architecture.
- Technical Architecture.

For each architectural domain, there is an associated set of Law related concerns, each set can be described by law architecture models that define the legal rules governing that domain. Integrating Law architecture with EA offers the following benefits:

- Brings a robust, Law-centric focus to the practice of EA in the public sector.
- Integrates all aspects of law compliance through all layers of the IT solution.
- Facilitates the adoption and implementation of new law amendments in public enterprises.

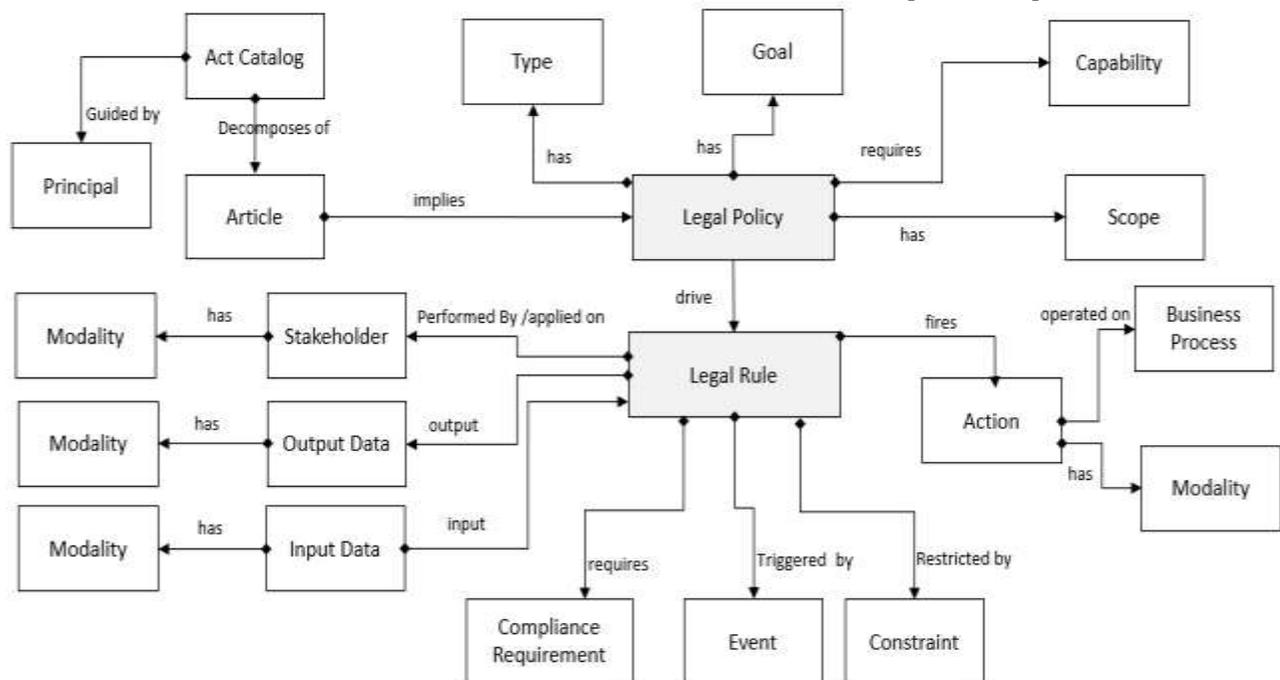


Fig. 4. Law Architecture Meta-Model.

TOGAF [27] is an architecture framework. TOGAF that is widely used by many EA practitioners within the public sector was selected here to demonstrate the capability of our proposed framework to integrate with EA frameworks. At the core of the TOGAF framework is the architecture development method (ADM) that “provides the methods and tools for assisting in the acceptance, production, use, and maintenance of enterprise architecture. It is based on an iterative process model supported by best practices and a re-usable set of existing architecture assets” [20]. The proposed Law Architecture Framework (LAF) here, has a Law-architecture-development-method (LADM) that provides the methods for assisting in the production and maintenance of law architecture. The ADM of the TOGAF can be adapted to integrate with the LADM of our proposed LAF.

Fig. 5 depicts the ADM in a larger form and Fig. 6 illustrates how TOGAF ADM can be integrated with LADM. The Suggested ADM Adaptation to be integrated with the proposed LADM can be listed as follow:

1) After the architecture vision (Phase A) of the ADM: Then the LADM should be initiated and completed according to the detailed description listed above.

2) After the end of the LADM: Then the TOGAF Business architecture phase should be performed.

3) The result from LADM Phase (A): Should be added as an input to the TOGAF Business architecture phase.

4) One of the objectives of the TOGAF business architecture phase is to develop the target business architecture, the target business architecture may have any suggested business process engineering that is based on some law amendments suggestions, these amendments should be linked with LADM Phase (F), Collect and Communicate the suggested Law amendments .

5) After the completion of the TOGAF business architecture phase, the information systems architecture phase should be initiated by adding the result from the LADM phase (D) to the inputs of that phase of the ADM.

6) One of the objectives of the TOGAF IS architecture phase is to develop the target IS architecture, the target IS architecture may have any suggested IS improvements based on some law amendments suggestions, these amendments should be linked with LADM Phase (F).

7) After the completion of the TOGAF IS architecture phase, the technology architecture phase should be initiated by adding the result from LADM Phase (E) to the inputs of the phase.

8) Any suggested technology improvements that require some law amendments suggestions, should be linked with LADM Phase (F)

9) The TOGAF ADM phases will be completed then as defined in TOGAF Standard.

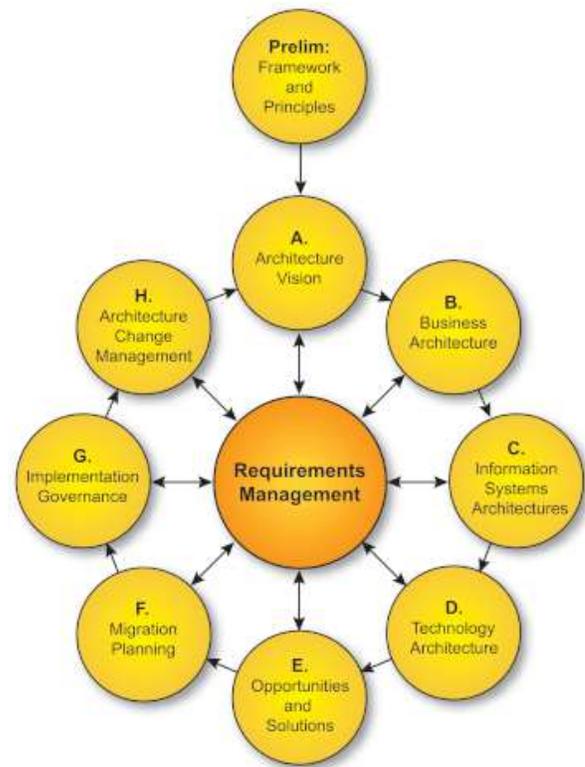


Fig. 5. ADM Process of the TOGAF Framework.

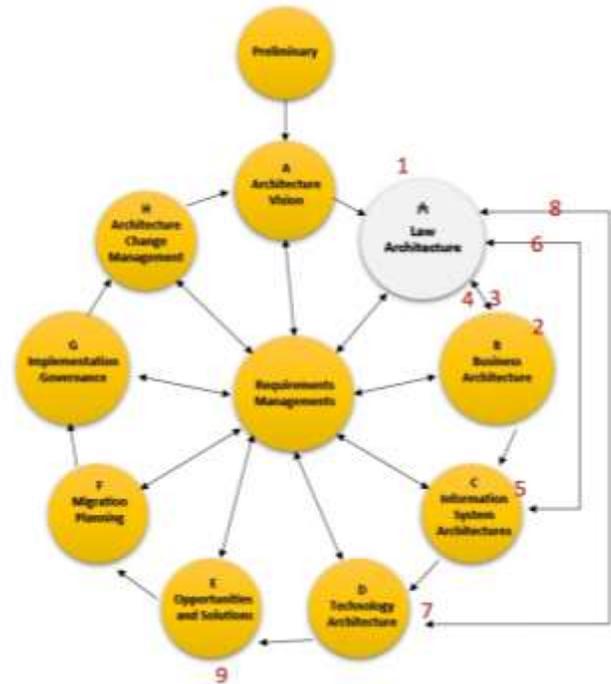


Fig. 6. Adapting the TOGAF ADM to Integrate with LADM.

IV. HEALTHCARE REFORM IN EGYPT THROUGH THE LENS OF THE LAW ARCHITECTURE: A CASE STUDY

TABLE II. UHI PROJECT-SPECIFIC ACTS CATALOG

A. Universal Health Insurance (UHI) Law at a Glance

The Egyptian Parliament endorsed a fresh law on universal health coverage that was formally promulgated on January 11, 2018, by the President. The Universal Health Insurance Law (2/2018) is regarded as an unprecedented effort to control the healthcare industry in Egypt, expanding the extensive protection of healthcare to every industry of culture. Healthcare was dealt with on a case-by-case ground before this law. The 27 governorates of Egypt will be split into six regional fields when applying the fresh law, and the law will then be implemented gradually over the next 15 years. Existing laws and regulations in the health insurance sector will stay in effect during the transformation era. Subscription to the new health insurance scheme will be compulsory for all Egyptians living in the Arab Republic of Egypt and free for Egyptians employed or remaining onboard with their relatives. Fees are laid by revenue and extra financing sources to include tariffs on the tobacco industry and other extra regions. The state has dedicated to offering the strategy to about 25 percent of the population who are unable to afford it free of cost.

B. Defining the UHI Law Architecture based on the LADM

The Ministry of Health has formed a technical committee including representatives from the Ministry of Communications and Information Technology (MCIT), the Ministry of Finance, and a consortium of experienced international and local experts. The mandate of this committee was to explore how to realize the objectives of the UHI law, study each use case defined in the law, and define the technical specification of the required systems. As the corresponding author of this manuscript was one of the enterprise architecture team assigned by MCIT to achieve the technical design and architecture of the required systems. The architecture team depended on analyzing the UHI law as the prime source of information in identifying the high-level system requirements. By following the steps and guidelines of the LADM proposed above and illustrated in Fig. 2, the team analyzed the UHI available laws and identified the law architecture of the UHI program before performing a detailed enterprise architecture activity, the extracted law architecture models resulted from this exercise can be illustrated as follow:

1) *Law architecture initiation*: During this phase, the law architecture work will follow the principles identified in section (III), and Table II represents the Project-Specific Acts Catalog that stores the core administrative laws related to the UHI project.

2) *Phase (A) Perform legal policy mapping*: The UHI law was analyzed, and 10 legal policies and 23 related legal rules were extracted as shown in Table III. To understand the main elements in the legal policy and their relationships to each other the legal policy model was suggested. For example, article (41) identified three important legal policies as shown in Fig. 7.

#	Act No.	Subject	Issuing Date	Issuing Authority
1	No.2 of 2018	The universal health insurance system law	Jan. 2008	The Parliament
2	Bylaw-No.909 of 2018	The executive regulations of the universal health insurance system law promulgated by law No. 2 of 2018	May.2008	The Prime Minister
3	No.707 of 2018	Setting criteria for targeting those below the poverty line in applying the provisions of the Universal Health Insurance Law	April 2018	The Prime Minister
4	No.2257 of 2020	Devolution of some medical assets in the governorates of Port Said and Luxor to the Public Health Care Authority in preparation for their inclusion in the Universal health insurance system	Nov. 2020	The Prime Minister
5	No.2273 of 2020	Determinants of patients with chronic diseases and tumors and controls for their exemption from Universal health insurance premiums	Nov. 2020	The Prime Minister

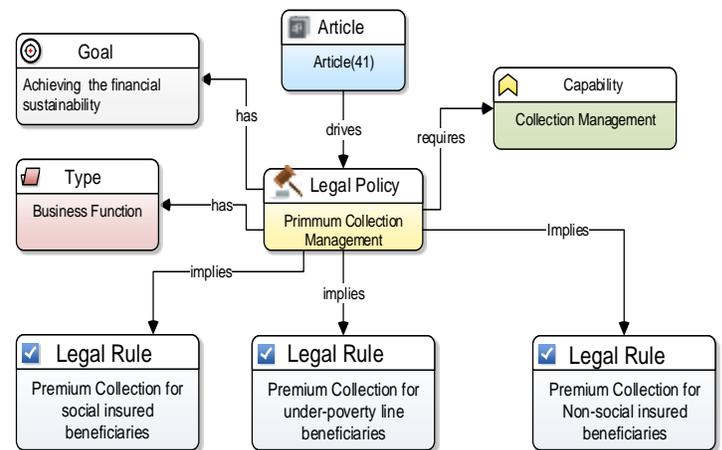


Fig. 7. Model of Premium Collection Management Policy.

3) *Phase (B) Perform legal rule mapping*: The legal policy catalog listed below in Table III was analyzed versus the UHI law to extract and model the related legal rules, Fig. 8 represents a sample legal rule model that represents “Premium Collection for Social Insured Beneficiaries”. Each extracted legal rule was modeled to provide a visualized artifact that can be easily understood by different business and IT stakeholders.

TABLE III. LEGAL POLICY CATALOG OF THE UHI PROJECT

Policy ID	Policy Name	Type	Ref. Article	Req. capability	Goal	Legal Rules
LP_01	Premium Collection Management	Business Function	41	Collection Management	Collection of Financial resources required to achieve financial sustainability	<ul style="list-style-type: none"> Premium Collection for social insured beneficiaries Premium Collection for under-poverty line beneficiaries Premium Collection for Non-social insured beneficiaries
LP_02	Premium calculation	Business Function	41	Collection Management	Collection of Financial resources required to achieve financial sustainability	<ul style="list-style-type: none"> Premium calculation for public employees Premium calculation for private-sector employees Premium calculation for Self Employed Premium calculation for pensioner premium calculation for pension entitled premium calculation for beneficiaries under the poverty line
LP_03	Board of Directors structure of UHI authority	organization structure	5	Program Management	Defining the structure and responsibility of the authority management	<ul style="list-style-type: none"> UHI authority board of directors Duration of the board
LP_05	Disability injuries policy	Business Function	13	Beneficiary management	Defining the process of issuing a disability certificate	<ul style="list-style-type: none"> disability certificate issuing responsibility disability certificate notification responsibility
LP_06	Co-payment policy	Business Function	table (3)	<ul style="list-style-type: none"> Health Service Management Collection Management 	Define the beneficiaries' contributions for health services	<ul style="list-style-type: none"> co-payment value for home visits co-payment value for medicine co-payment value for scans co-payment value for medical tests co-payment value for inpatient services
LP_07	Military recruits handling policy		53	Beneficiary management	Managing the services for the military people	<ul style="list-style-type: none"> Military recruits handling procedures
LP_08	UHI website and Database	Technical Function	57 - bylaw	IT Management	Building a website and DB for UHI Authority	<ul style="list-style-type: none"> building a website and defining access rights
LP_09	National Social Insurance Authority responsibility	Technical Function	57 - bylaw	IT Management	Feeding the authority's database with the necessary data for all those subject to the registered social insurance laws and their families.	<ul style="list-style-type: none"> feeding the authority's database with social insured people and their families
LP_10	work injuries policy	Business Function	12	Beneficiary Management	Regulate the work injury treatment process	<ul style="list-style-type: none"> work injury notification procedures work injury treatment

TABLE IV. LEGAL POLICY / CAPABILITIES MATRIX

Policy Name	Business capability
Premium Collection Management	Collection Management
Premium calculation	Collection Management
Board of Directors structure of UHI authority	Program Management
Disability injuries policy	Beneficiary management
Co-payment policy	Service Management
Military recruits handling policy	Beneficiary Management
UHI website and Database	IT Management
National Social Insurance Authority responsibility	IT Management

1) Phase (C) identify the correlation between legal policies business architecture:

a) *Develop legal Policy/Capabilities Matrix:* Table IV. depicts legal policy/capabilities matrix.

b) *Legal Rule/Business Process Matrix:* Table V depicts legal rule/business process matrix.

TABLE V. LEGAL RULE / BUSINESS PROCESS MATRIX

Legal Rule	Business Process
Premium Collection for social insured beneficiaries	Manage accounts receivable Funds
disability certificate issuing responsibility	Authorize Medical Service
co-payment value for home visits	Manage Rate setting
Military recruits handling procedures	Manage Treatment plan
Feeding the authority's database with social insured people and their families	Manage Integration with external entities
work injury notification procedures	Manage Treatment plan

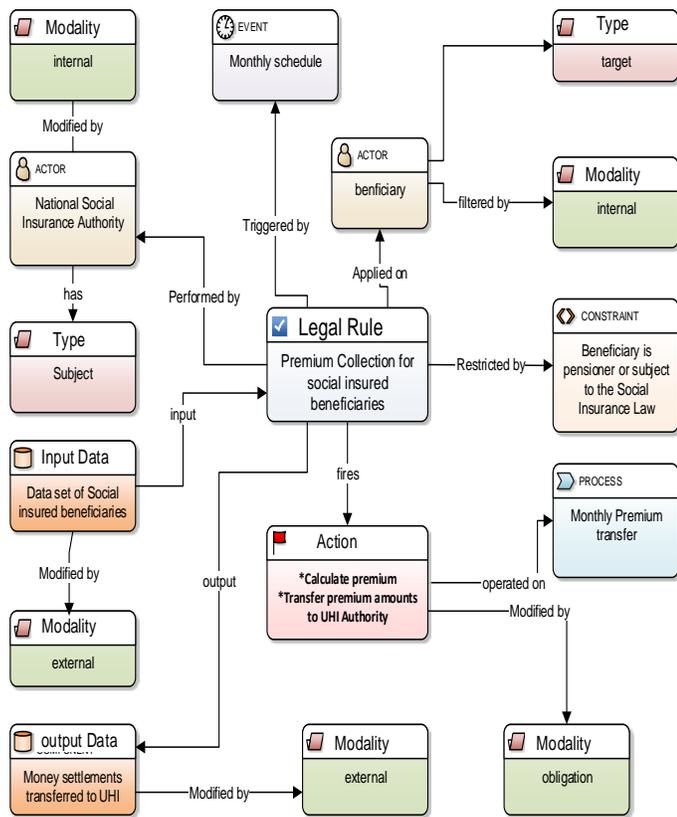


Fig. 8. Legal Rule Mode: Premium Collection for Social Insured Beneficiaries.

c) Stakeholder Mapping: By analyzing the legal policy and rules catalogs several system use cases can be identified in terms of obligations and rights as illustrated in the below Fig. 9.

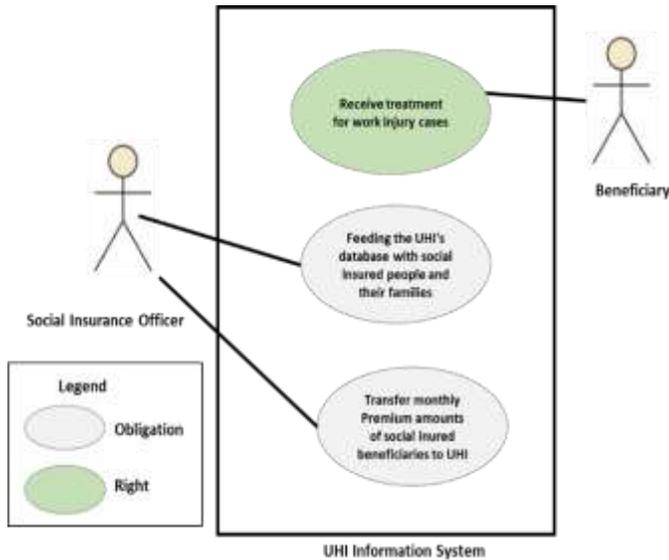


Fig. 9. A Legal Use Case Diagram of the UHI Project.

2) Phase (D) Legal Policy/Information systems Correlation:

a) Data Concepts Extracted from the Law: By analyzing the legal policy and rules catalogs the data concept catalog that is related to the IS was created, the most important data concept related to our extracted above policies is the beneficiary, and by analyzing the legal rule's actions and constraints we deduced that the beneficiary data should be segmented based on the following features shown in Fig. 10.

b) Application Interaction Matrix: By classifying the data concepts as internal or external to the UHI authority, the external data concepts define an integration requirement. Moreover, analyzing the actions related to the legal rules can also identify integration requirements. Following these procedures, the following integration requirements were identified as below in Fig. 11 and Table VI:

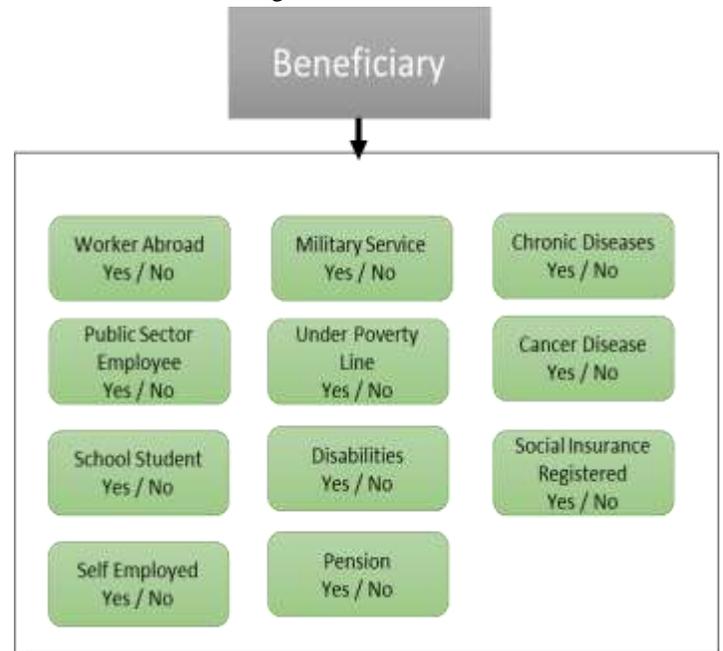


Fig. 10. Beneficiary Segmentation Features.

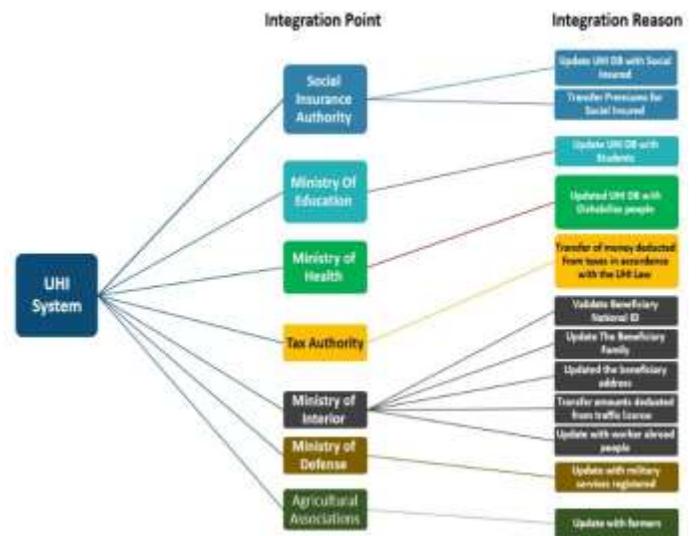


Fig. 11. Application Interaction Diagram.

TABLE VI. APPLICATION INTERACTION MATRIX

#	Target System (External Systems)	Integration Reasons
1	Social Insurance Authority	<ul style="list-style-type: none">• Update UHI DB with Social Insured• Transfer Premiums for Social Insured
2	Ministry of Education	<ul style="list-style-type: none">• Update UHI DB with Students
3	Ministry of Health	<ul style="list-style-type: none">• Updated UHI DB with Dishabilles people
4	Tax Authority	<ul style="list-style-type: none">• Transfer of money deducted from taxes under the UHI Law
5	Ministry of Interior	<ul style="list-style-type: none">• Validate Beneficiary National ID• Update the Beneficiary Family• Updated the beneficiary address• Transfer amounts deducted from traffic license• Update with worker abroad people
6	Ministry of Defense	<ul style="list-style-type: none">• Update with military services registered
7	Agricultural Associations	<ul style="list-style-type: none">• Update UHI system with farmers and their families
8	Ministry of Social Solidarity	<ul style="list-style-type: none">• Update UHI System with under-poverty-line people and their families

3) Phase (E) Legal Policy / Technology architecture correlation: The following technology requirements were extracted based on the legal policy catalog and the regulatory compliance requirements catalog was updated as following Table VII:

TABLE VII. SAMPLE OF REGULATORY REQUIRMENTS CATALOG

#	Requirement	Type	Rational	Architecture Domain
1	The system and servers should operate at 24/7 uptime.	availability	The system provides healthcare services	Technology
2	all changes to a financial database must be recorded in a trace file with before and after values	Audit Tracing	Enable financial transactions tracing	Technology
3	All communications between the UHI system and Social Insurance System	Communication	Respond to system integration requirements	Technology
4	Keep a record of all transactions for 5 years	Data Retention	Mandated by law	Technology

V. DISCUSSION

As presented above, in the introduced case study, the UHI law was the prime source of information that the technical committee depended on for extracting the high-level specifications of the digital transformation solution required for the UHI program. EA frameworks lack a detailed methodology for analyzing and modeling laws, thus our proposed law architecture framework suits these types of problems where the

core business processes and services are highly regulated by laws. The law architecture framework offers the ability to visualize what decoupling and divesting laws and how it impacts various aspects of the public ecosystem. This was illustrated through the use case as the legal policies and rules were extracted from the UHI law, in addition to the relationship between these rules and business capabilities, processes, IS, data concepts, and technical specifications based on the proposed LADM that represents the heart of the proposed law architecture framework. The application of the proposed framework supported the UHI program in extracting the regulatory compliance requirements from the law and supported the design of regulatory-compliant system.

Law architecture does not stand alone in a public enterprise. Several complementary or ancillary frameworks or disciplines may exist within a public enterprise. Each of these frameworks may have its focus. The success and maturity of law architecture are influenced by its ability to integrate and align with each of these disciplines and frameworks. Another limitation in our proposed framework is the LADM proposed for extracting the legal rules and policies and their correlations with other artifacts in the public enterprise model is a manual task. This is because ambiguity exists in the laws that require the investigation and validation of legal experts.

VI. CONCLUSIONS

In recent years, there have been various contributions to deal with law compliance problem, None of the previously surveyed work in the state of the art section takes an (i) generic approach to law compliance management, (ii) seeks to integrate with EA frameworks, (iii) allows declarative modeling of policies and rules while separating these two concepts and, (iv) introduces a new concept and formal definition for the law architecture and, (v) supports for traceability by a set of links to establish between legal policies and enterprise models. Our claim through this work is that our proposed framework allows for achieving these goals and overcomes these limitations:

- Domain-independence: a solution to the problem of IS requirements engineering to comply with all applicable administrative laws.
- Domain-variability: a solution that takes into account the variability of possible legal policies and rules applied in different public sector services.
- Flexible-integration: a solution that takes into consideration the integration with other EA frameworks to achieve a law-compliant public enterprise model.
- Machine-readable: a solution that can be easily translated into an ontology model.
- Knowledge-gap: a solution that covers the knowledge gap between the technical architects and the legal experts.

Finally, the applicability of our proposed framework is shown and validated through a case study. Subject matter experts of the legal domain also evaluated the extracted legal policies and rules during the implementation of the case study.

VII. FUTURE WORK

In this paper, we did not cover all components of the proposed law architecture framework, future work will include a detailed description of other framework components such as compliance management and ontology representation of the framework metamodel. Moreover, the capability of our framework to integrate with other EA frameworks will be investigated.

ACKNOWLEDGMENT

The activities described in our case study were supervised by Dr. Khaled El-Attar, Vice Minister - The Egyptian Ministry of Communications and Information Technology, as part of the Universal Health Insurance program implementation. We also dedicate this work to the deceased Prof. Hawaf Abdel-Hakim for his kindness, help, and assistance.

REFERENCES

- [1] N. A. A. Bakar, S. Harihodin, and N. Kama, "Assessment of Enterprise Architecture Implementation Capability and Priority in Public Sector Agency," in *Procedia Computer Science*, 2016, vol. 100, pp. 198–206, doi: 10.1016/j.procs.2016.09.141.
- [2] G. L. Richardson, B. M. Jackson, and G. W. Dickson, "A principles-based enterprise architecture: Lessons from Texaco and Star Enterprise," *MIS Quarterly: Management Information Systems*, vol. 14, no. 4, pp. 385–402, 1990, doi: 10.2307/249787.
- [3] B. van der Raadt, M. Bonnet, S. Schouten, and H. van Vliet, "The relation between EA effectiveness and stakeholder satisfaction," *Journal of Systems and Software*, vol. 83, no. 10, pp. 1954–1969, Oct. 2010, doi: 10.1016/j.jss.2010.05.076.
- [4] A. Elgammal and O. Turetken, "Lifecycle Business Process Compliance Management: A Semantically-Enabled Framework," in *2015 International Conference on Cloud Computing, ICC3 2015*, Jul. 2015, pp. 1–8, doi: 10.1109/CLOUDCOMP.2015.7149646.
- [5] M. Ochodek and S. Kopczyńska, "Perceived importance of agile requirements engineering practices – A survey," *Journal of Systems and Software*, vol. 143, pp. 29–43, Sep. 2018, doi: 10.1016/j.jss.2018.05.012.
- [6] S. D. Ringmann, H. Langweg, and M. Waldvogel, "Requirements for legally compliant software based on the GDPR," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11230 LNCS, pp. 258–276, doi: 10.1007/978-3-030-02671-4_15.
- [7] A. M. Mustapha, O. T. Arogundade, A. Abayomi-Alli, O. J. Adeniran, K. Adesemowo, and C. Y. Alonge, "A Systematic Method for Extracting and Analyzing Cloud-Based Compliance Requirements," in *2020 International Conference in Mathematics, Computer Engineering and Computer Science, ICMCECS 2020*, Mar. 2020, pp. 1–7, doi: 10.1109/ICMCECS47690.2020.240839.
- [8] M. el Ghosh, H. Naja, H. Abdulrab, and M. Khalil, "Ontology learning process as a bottom-up strategy for building domain-specific ontology from legal texts," in *ICAART 2017 - Proceedings of the 9th International Conference on Agents and Artificial Intelligence*, Feb. 2017, vol. 2, pp. 473–480, doi: 10.5220/0006188004730480.
- [9] G. Kurcheeva, M. Rakhvalova, D. Rakhvalova, and M. Bakaev, "Mining and indexing of legal natural language texts with domain and task ontology," in *Communications in Computer and Information Science*, Nov. 2019, vol. 947, pp. 123–137, doi: 10.1007/978-3-030-13283-5_10.
- [10] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini, and L. Robaldo, "PrOnto: Privacy ontology for legal reasoning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Sep. 2018, vol. 11032 LNCS, pp. 139–152, doi: 10.1007/978-3-319-98349-3_11.
- [11] V. Leone, L. di Caro, and S. Villata, "Taking stock of legal ontologies: a feature-based comparative analysis," *Artificial Intelligence and Law*, vol. 28, no. 2, pp. 207–235, Jun. 2020, doi: 10.1007/s10506-019-09252-1.
- [12] J. Becker, J. vom Brocke, M. Heddier, and S. Seidel, "In Search of Information Systems (Grand) Challenges: A Community of Inquirers Perspective," *Business and Information Systems Engineering*, vol. 57, no. 6, pp. 377–390, Dec. 2015, doi: 10.1007/s12599-015-0394-0.
- [13] C. Stratigaki, M. Nikolaidou, P. Loucopoulos, and D. Anagnostopoulos, "Business process elicitation from regulatory compliance documents: An E-government case study," in *Proceedings - CBI 2016: 18th IEEE Conference on Business Informatics*, Dec. 2016, vol. 2, pp. 8–13, doi: 10.1109/CBI.2016.43.
- [14] N. Peter Mrope, G. Simiyu Namusonge, and M. Amuhaya Iravo, "Does Compliance with Rules Ensure Better Performance? An Assessment of The Effect of Compliance with Procurement Legal and Regulatory Framework on Performance of Public Procurement in Tanzania," *European Journal of Logistics, Purchasing and Supply Chain Management*, vol. 5, no. 1, pp. 40–50, 2017.
- [15] S. Seppälä, M. Ceci, H. Huang, L. O'Brien, and T. Butler, "SmaRT Visualisation of Legal Rules for Compliance," *undefined*, pp. 73–85, 2017.
- [16] M. Hashmi and G. Governatori, "Norms modeling constructs of business process compliance management frameworks: a conceptual evaluation," *Artificial Intelligence and Law*, vol. 26, no. 3, pp. 251–305, Sep. 2018, doi: 10.1007/s10506-017-9215-8.
- [17] C. Bartolini, A. Calabró, and E. Marchetti, "Enhancing business process modelling with data protection compliance: An ontology-based proposal," in *ICISSP 2019 - Proceedings of the 5th International Conference on Information Systems Security and Privacy*, 2019, pp. 421–428, doi: 10.5220/0007392304210428.
- [18] R. Matulevičius, J. Tom, K. Kala, and E. Sing, "A Method for Managing GDPR Compliance in Business Processes," in *Lecture Notes in Business Information Processing*, Jun. 2020, vol. 386 LNBIP, pp. 100–112, doi: 10.1007/978-3-030-58135-0_9.
- [19] T. Seyffarth and S. Kuehnel, "Maintaining business process compliance despite changes: a decision support approach based on process adaptations," 2021, doi: 10.1080/12460125.2020.1861920.
- [20] A. S. Zalazar, L. Ballejos, and S. Rodriguez, "Analyzing requirements engineering for cloud computing," in *Requirements Engineering for Service and Cloud Computing*, Springer International Publishing, 2017, pp. 45–64.
- [21] S. Agarwal, S. Steyskal, F. Antunovic, and S. Kirrane, "Legislative compliance assessment: Framework, model and GDPR instantiation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Jun. 2018, vol. 11079 LNCS, pp. 131–149, doi: 10.1007/978-3-030-02547-2_8.
- [22] M. Alsaadi, A. Lisitsa, and M. Qasaimeh, "Minimizing the ambiguities in medical devices regulations based on software requirement engineering techniques," in *ACM International Conference Proceeding Series*, Dec. 2019, pp. 1–5, doi: 10.1145/3368691.3368709.
- [23] F. Jorshari and R. H. Tawil, "A High-Level Scheme for an Ontology-Based Compliance Framework in Software Development," *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*, pp. 1479–1487, 2015.
- [24] M. Usman, M. Felderer, M. Unterkalmsteiner, E. Klotins, D. Mendez, and E. Alégroth, "Compliance Requirements in Large-Scale Software Development: An Industrial Case Study," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2020, vol. 12562 LNCS, pp. 385–401, doi: 10.1007/978-3-030-64148-1_24.
- [25] Business Architecture Guild, *A Guide to the Business Architecture Body of Knowledge (BIZBOK Guide)*, Version 6.5. 2018.
- [26] "The Federation of Enterprise Architecture Professional Organizations (FEAPO)." <https://feapo.org/> (accessed May 24, 2021).
- [27] OMG: The Open Group, *The Open Group Architecture Framework (TOGAF)*, Version 9.2. 2018.

Microorganisms: Integrating Augmented Reality and Gamification in a Learning Tool

Ratna Zuarni Ramli¹

Faculty of Computer and Mathematical Science
Universiti Teknologi MARA
Kuala Pilah, Malaysia

Nor Athirah Umairah Marobi²

Noraidah Sahari@Ashaari³
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia, Bangi Malaysia

Abstract—Microorganisms is a Year 6 Science subject that primary school students find considerably less attractive because of the enormous facts that require them to have good imaginary skills to understand. Only limited applications are available on smartphones as tools to learn subjects, especially Science, Technology, Engineering and Mathematics (STEM). Since the young generation is very much into current technology, there is a need to develop an application to gain students' interest and improve their understanding of microorganisms. Therefore, a microorganism learning application that combines augmented reality (AR) and gamification called Microorganisms was developed. The Microorganisms prototype includes two modules; learning and training. The learning modules use AR technology that scan marker images to display a digital layer of microorganisms in three dimensions (3D). Meanwhile, the training module is represented through gamification consisting of quiz questions, a timer, and a score. The application was designed using Agile methodology and developed using various software such as Unity, Autodesk 3ds Max, Vuforia, and Firebase. Ten respondents, nine students and one teacher in a primary school, assessed the prototype through experimental testing. The results showed that, on average, the user satisfaction value was 4.6 out of 5. Thus, the Microorganisms application based on AR and gamification can be considered a good learning tool for primary school students to learn about micro-organisms.

Keywords—Augmented reality; digital game-based learning (DGBl); game; learning tool; microorganisms; usability testing

I. INTRODUCTION

Students have been learning the Science subject from a young age, not only as a requirement in school but also for student comprehension. Microorganisms are one of the sub-topics that Year 6 students learn in primary school. The students need to realise that microorganisms, such as viruses, are tiny and can only be seen through a microscope and cannot be seen with the naked eye. This concept becomes more difficult for students to understand and appreciate the design and structure of the microorganism itself. Most students and even parents consider that the Science subject or field is not easy to understand [1]. Moreover, scientific notes are dense with writing and facts, which causes students to lose interest in learning and understanding the topic of microorganisms. Thus, it is essential to create a learning tool that can grab students' attention and realise that Science is an exciting subject to learn. Factors that can motivate students to learn include presenting the content in various ways, such as sound, visualisation, and

animation [2]. Specifically, it is more stimulating for the students to get to see the microorganisms in 3D.

Augmented reality (AR) is a current technological medium that offers a unique opportunity that combines the physical and virtual worlds [3]. AR works interactively and in real-time, aligns real and virtual objects [4]. Many applications relating to education have applied AR in their design. For example, the AR Marine Scientific program aims to create awareness of marine environments among lower-grade primary school students. It was identified that the application could help low academic achievers improve their learning performance [5]. Besides, AR (as a learning supplement) has also been used in research to evaluate and compare Biology students learning via AR and printed books, in which it was found that both methods indeed helped enhance the learning of Biology [6]. Furthermore, some AR applications include a game element to attract learners' attention. Gaming is significantly related to the young generation, which generally consists of avid gamers. Therefore, a good learning tool that combines AR technology and gamification in the application design may help students learn better [7].

Thus, this research aims to design and develop an application as a learning tool to learn the microorganism topic, especially for primary school students, with the integration of AR technology and gamification in the design. First, this research paper discusses the current tools in the market for learning the Microorganism topic and the AR applications that have been developed by other researchers in a similar area, followed by the present research methodology, including the framework and the development of the Microorganism application prototype. Sections 3 to 6 discuss the process of evaluating the application, Microorganisms. Finally, the paper reviews the evaluation results and discusses the development of the Microorganisms application.

II. LEARNING TOOL FOR MICROORGANISMS

Learning Science subjects can be frustrating for some people because they cannot see or imagine how certain scientific phenomena work. Hence, many learning tools have been created to help enhance learning activities, especially science subjects or topics. Some tool designs may suit particular aims and can be used by others facing similar issues in learning.

One study compared the learning performance between students learning biology using an AR application and a

printed book [6]. Their study showed no difference in understanding and remembering factors between the students in both categories. However, the students that used the AR application showed a better learning attitude than the students who used printed books. However, the study gave no detailed explanation of the application design.

Meanwhile, another work [8] proposed a visual aid for students and practitioners to learn about antimicrobials in Microorganisms in the form of a video. The support was aimed at adult learners, and based on the result analysis, the tool indeed helped improve user understanding. Moreover, the video presentation is an excellent method to assist students in learning the topic. Nevertheless, it can still be improved by using 3D or even four dimensions so that the organism could be seen from a different angle.

Another study created a learning tool for Microorganisms specifically to learn the mushroom structure. The study applied AR technology to display the design in 3D [9]. Based on the evaluation process, it was identified that the tool offered a reasonable satisfaction rate among users. Nevertheless, the application focused on AR and did not include any game element in the design. Also, the application content is not similar to the primary school syllabus in Malaysia. Still, it can be used as a reference for local students to learn the topic beyond that covered in school.

Applications with AR features have been developed since the 2000s. Many AR applications help students to improve their learning performance, especially in subjects considered challenging or less interesting; for example, ATHYNOS, a serious AR game for children with Dyscalculia, was developed to help them learn mathematics. ATHYNOS was identified to be able to help students with calculation disorders and, based on the evaluation results, the students took less time to understand the subject [10]. Another application was also designed and developed to learn Kanji, considered one of the most challenging writing systems. Called Dragon Tale, the adventure game with AR features was designed to help students learn Kanji interactively [11]. Similar research also aimed to help students in Thailand to learn the same writing system [12]. Hence, the study designed an AR application that improved Kanji skills by more than 41%.

According to some studies [6], [8], [9], a learning application should consider including multiple media to attract students' attention, such as video, audio, graphic, text and animation. The lengthy and dense text should be avoided as much as possible. The element of 3D that is activated through AR features can be one of the fun factors in learning [13]. It is also essential to assess the student's comprehension; thus, questions or quizzes should be attached at each lesson in the application [6], [8], [14]. Since gaming is very much connected to the young generation, the assessment can be presented in a game where students can compete with other players. The current problem is identified based on real scenarios at local schools where students are having a tough time learning the subject. Since there are limited learning materials that suit the syllabus, a research project is conducted to develop a learning application for primary school students to learn Microbiology.

III. METHODOLOGY

The research methodology was divided into three main phases: 1) Conceptual, 2) Design and Development, and 3) Evaluation. The conceptual phase involved the process of analysing the current problems and project requirements. The second phase, Design and Development, comprised designing the application and developing the prototype. At the same time, the last phase involved evaluating the prototype and verifying the application design. The overall research method was adapted from the agile methodology that is usually applied in software development. In mobile software engineering, this methodology is essential since the software is constantly changing and evolving based on users' immediate needs [15]. Table I shows the phases and research activities for this study.

In phase 1, current issues are identified and based on past studies, the essential elements are recognised. Thus, the elements that are included in the application are multimedia like text, graphic, audio, video and animation in 3D. The application also includes AR and gaming feature as fun factors to draw students' interest.

In phase 2, the application is design and developed using multiple software and briefly describe in the following subtopic. While in phase 3, the process of evaluation is discussed concisely in Section IV.

A. Framework

There are two modules in the Microorganisms application, learning and training, which are indirectly interconnected. The student can browse through the learning modules or choose to answer questions in the training modules at any time. The learning module is aimed to represent the content of Microorganisms in 3D with the support of Augmented Reality technology. Five topics were included in the learning modules: Fungi, Bacteria, Viruses, Protozoa, and Algae. Students can choose any topic to learn by clicking on the listed buttons labelled with the topic's name. In addition, a button labelled AR is also included in each topic to represent the content of the specific topic in 3D. The student must scan a specialised marker to activate the AR function. Besides, a video presentation of each topic is also included in the learning module.

TABLE I. RESEARCH ACTIVITY

Phase	Research Activity	Output
1. Conceptual	Analyse the current problem Get initial information from primary school teachers and parents Analyse similar studies	Component
2. Design and Development	Design the proposed framework Identify software requirements Develop the application prototype	Application prototype
3. Evaluation	Functional testing Usability testing	Verified application framework

The training module is aimed at testing student knowledge on the same topics as the learning module. In the training module, some questions are grouped into four difficulty levels to imitate the gamification structure. The students are given a time limit to answer the questions, and at the end of the training session, the student's individual score is displayed together with their ranking. Fig. 1 shows the application framework for Microorganisms.

The application prototype for Microorganisms was designed based on an application framework. Adobe Photoshop and Unity software were used to create the application interface design. Meanwhile, 3D models for the Augmented Reality modules were developed using Autodesk 3ds Max software. The Firebase database was used to store user data and the scores obtained in the training module. The languages used to program the application were C# and PHP. Fig. 2 to Fig. 9 shows the interface design for Microorganisms.

B. Prototype

The application prototype for Microorganisms was designed based on an application framework. Adobe Photoshop and Unity software were used to create the application interface design. Meanwhile, 3D models for the Augmented Reality modules were developed using Autodesk 3ds Max software. The Firebase database was used to store user data and the scores obtained in the training module. The languages used to program the application were C# and PHP. Fig. 2 to Fig. 9 show the interface design for Microorganisms.

Fig. 2 shows the main page of the application that consists of three buttons, Masuk (Login), Tetapan (Settings), and Keluar (Exit) buttons. Users need to select the login button to go to the category selection page. Users can also choose the settings button to set the voice settings and see the how-to-play guide while pressing the exit button to stop the application.

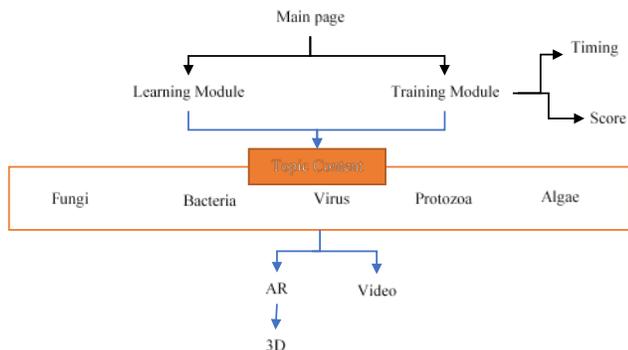


Fig. 1. Application Framework.



Fig. 2. Main Page.

Fig. 3 shows the category page where users can select the Belajar (Note) button to view notes related to Microorganisms or the Kuiz (Quiz) button to answer quizzes. Users are also allowed to return to the main page by clicking the home button at the top of the page.

Fig. 4 shows the Notes page where the users can select the Fungi, Bacteria, Viruses, Protozoa, and Algae buttons to view notes related to the chosen type of microorganism. There are also buttons for Jenis (Type), Proses Hidup (Processes), Kegunaan (Uses), Keburukan (Disadvantages), and Pencegahan (Prevention). Users can also view the 3D model in AR by clicking on the AR button. There is also a back button to return to the category selection page.

Fig.5 shows a video page with similar buttons in the Notes page to allow users to view related video descriptions in video format. There is also a back button to return to the category selection page.

The application can access the users' camera to allow users to scan the designed marker. Fig. 6 shows the page that displays the 3D model where users can select the Fungi, Bacteria, Viruses, Protozoa and Algae buttons to view 3D models related to the chosen type of microorganism. There is also a back button to return to the Notes page.



Fig. 3. Category Page.



Fig. 4. Notes Page.



Fig. 5. Video Page.



Fig. 6. 3D Model in Augmented Reality.

Fig. 7 shows the game level where users have to answer the quiz questions in order to move to a higher rank.



Fig. 7. Level of Difficulties in the Game.

There is a score on the left side of each question. All questions are multiple-choice, like in Fig. 8. There is a time limit to answer each question.



Fig. 8. Example of Quiz.

Fig. 9 shows the scoreboard page to displays the five highest user scores in ascending order.



Fig. 9. Scoreboard.

IV. EVALUATION

Evaluation was conducted via two sets of questionnaires to obtain feedback from the users of the Microorganisms application. The testing process involved ten respondents, consisting of nine students and one Science teacher. The questionnaire consisted of three parts: to test the function of the application and a non-functional variable, usability.

A. Functional Testing

In functional testing, Microorganism was tested to identify any problems in operating the application. The functions in the learning modules, including browsing the content using the buttons, the scan markers to test the AR function, and watching the 3D and video content, were tested. Meanwhile, in the training modules, the testing was done by browsing through the game questions of different difficulties, answering some questions, and getting feedback and the game score. Functional testing was done via a simple walkthrough method to ensure that all features such as the buttons, notes, video, 3D model, and AR marker worked well. Four significant phases were included in the functional testing, as described in Table II, while Table III shows one of the detailed procedures in the functional testing. The functional testing procedure was adapted from a past study [16]. Each team member in the project did the testing. No issue was found in all four phases of functional testing. The status column in Table II indicates a pass to show that the application features are working well.

TABLE II. PHASES IN FUNCTIONAL TESTING

Phase	Description	Process	Status
1	Display notes and video	Manual	Pass
2	Display 3D model	Manual	Pass
3	Play game	Manual	Pass
4	Display scoreboard	Manual	Pass

TABLE III. EXAMPLE OF DETAILED FUNCTIONAL TESTING PROCEDURE IN PHASE 2

Criterion	Description
Objective	To ensure the 3D module works well. It tests whether an application can display a 3D model of the microorganism and respond correctly to user interactions.
Pre-requirement	Markers for AR for users to scan.
Testing Procedure	<ol style="list-style-type: none"> 1. Install the application on a mobile device. 2. Run the application. 3. Select the "Login" button – The application displays the learn and quiz menu page. 4. Select the "Learn" button – The application displays the learning notes page. 5. Select the "AR" button available on each type of microorganism – The application displays the camera mode. 6. Scan the marker image as a target image – The application displays the model and name of the microorganism.
Exit steps	<ol style="list-style-type: none"> 1. Press the home button. 2. The application returns to the main page. 3. Press the "Exit" button.

B. Usability and Effectiveness Evaluation

Ten respondents were asked to evaluate the application and then answer a set of questions using a Likert scale to assess the usability of the Microorganisms application. According to [17], responses from five to ten participants are enough in experimental evaluation. The testing was conducted via an online medium to obtain feedback from the respondents about the usability of Microorganisms. Then, .apk files were uploaded onto Google Drive and the URL distributed to the respondents. Feedback from the respondents was collected via Google Forms online questionnaire. The usability evaluation was done to ensure that the application is in line with what users want by observing how users use the application. There are ten questions in total, with the items adapted from a past study [18]. Based on the descriptive analysis, the total mean value for usability was 4.26 and 4.56 for effectiveness. The detail of the usability and effectiveness result is shown in Table IV.

TABLE IV. USABILITY AND EFFECTIVENESS RESULT

No.	Factor	Item	Mean
1	Usability	User-friendly	4.3
2	Usability	Clear instruction	4.5
3	Usability	Easy to use	3.9
4	Usability	Button works well	4.2
5	Usability	Attractive interface	4.4
6	Effectiveness	Easy to understand	4.4
7	Effectiveness	AR is interesting	4.5
8	Effectiveness	Questions in quizzes are understandable	4.5
9	Effectiveness	Enjoyable	4.8
10	Effectiveness	Suitable	4.6

V. DISCUSSION AND CONCLUSION

The initiative to begin the research project came from a teacher in a local primary school who had trouble teaching the microorganism topic. The teacher had problems showing the details of the microorganism to students since the school had no lab equipment. Reference books were the only resource and non-interactive, so it wasn't easy to attract students' interest in learning the topic. Thus, Microorganisms' application was developed as a learning tool to help the students understand the topic and the teacher to teach the subject. It was purposely developed in a mobile platform since most of the students' parents have mobile gadgets that the students can use not only at school or with the teacher but anywhere. The teacher specifically prepared the content of the notes, and the representation of the notes was redesigned by the researcher with the teacher's approval and verification.

To make the notes more interesting, visualising the topic can be a good method to apply [19]. Thus, features of AR and 3D models were added to the mobile application. These features have been identified to hold students' attention in the subject matter [11], [12], [20], [21]. The researchers argued that AR with 3D models creates new experiences and provides new teaching methods, learning, and research [13], [21].

Hence, the AR feature with a 3D model was designed as part of the application to grasp students' attention and interest to learn the topic. Since the topic is part of Science, Technology, Engineering and Mathematics (STEM), the application can also be considered as one of the initiatives to expose students to STEM concepts based on real-world situations.

Designing for a generation exposed to current technology and addiction, gamification has now become a compulsory element in learning tools [22]. Many learning tools that integrate gaming elements in their content aim to increase enjoyment in learning and measure student comprehension of the subject [10], [11]. Also, some studies emphasised the vital role of games in knowledge acquisition and content understanding [14], [22]. The study found that many learning tools have succeeded in delivering knowledge by incorporating gaming into their devices. Consequently, Microorganisms was developed with AR and layering of 3D models, with game element integration.

The process of designing and developing Microorganisms took seven months, while another three months were required to evaluate its features, usability, and effectiveness. The evaluation of the application features, known as functional testing, was done via a simple walkthrough method, while the evaluation of usability and effectiveness was done via an online questionnaire. Based on the results of the assessments, the Microorganisms application received good feedback from all respondents, with high mean values for all items. Therefore, this application can be used in class as a learning tool for teachers to teach their students the same topic.

There are some limitations to this research project; firstly, many respondents are negligible due to an imbalance in the number of students and teachers. In the future, the number of respondents will be increased to balance between both types of users (student and teacher). Secondly, the set of questions used to evaluate Microorganisms' usability is simple, so the researchers were unable to obtain detailed feedback from the respondents that could have further enhanced the application's design.

VI. FUTURE WORKS

Based on the discussion above, this research will continuously improve the diverse representation of the microorganism notes, adding more gamification elements, such as group competitions, and setting the application in an iOS platform. At the same time, future research will focus more on the importance of suitable methods and sufficient respondents to evaluate the application. The proper method of evaluation will uncover a significant or insignificant aspect of any application [23], [24], [25].

ACKNOWLEDGMENT

This study was funded by the Universiti Kebangsaan Malaysia Research Grant (GPK-P&P-2020-005).

REFERENCES

- [1] Sumintono, B. (2017). Science education in Malaysia: challenges in the 21st century. *Jurnal Cakrawala Pendidikan*, 36(3).
- [2] Nielsen, B. L., Brandt, H. & Swensen, H. (2016). Augmented reality in science education—affordances for student learning. *Nordic Studies in Science Education* 12(2): 157.

- [3] Kesim, M. & Ozarslan, Y. (2012). Augmented Reality in Education: Current Technologies and the Potential for Education. *Procedia - Social and Behavioral Sciences* 47(222): 297–302.
- [4] Salmi, H., Kaasinen, A. & Kallunki, V. (2012). Towards an Open Learning Environment via Augmented Reality (AR): Visualising the Invisible in Science Centres and Schools for Teacher Education. *Procedia - Social and Behavioral Sciences* 45: 284–295.
- [5] Lu, S.-J., & Liu, Y.-C. (2015). Integrating augmented reality technology to enhance children's learning in marine education. *Environmental Education Research*, 21(4), 525–541. doi:10.1080/13504622.2014.911247.
- [6] Weng, C., Otanga, S., Christianto, S. M., & Chu, R. J.-C. (2019). Enhancing Students' Biology Learning by Using Augmented Reality as a Learning Supplement. *Journal of Educational Computing Research*, 073563311988421. doi:10.1177/0735633119884213.
- [7] Suprpto, N., Nandyansah, W., & Mubarak, H. (2020). An Evaluation of the "PicsAR" Research Project: An Augmented Reality in Physics Learning. *International Journal of Emerging Technologies in Learning (iJET)*, 15(10), 113-125.
- [8] Athauda, G., Toonkel, R. L., & Kashan, S. B. (2018). Developing a visual learning tool to aid the studying of antimicrobial spectrum of activity. In *Proceedings for Annual Meeting of The Japanese Pharmacological Society WCP2018 (The 18th World Congress of Basic and Clinical Pharmacology)* (pp. PO1-7). Japanese Pharmacological Society.
- [9] Aottiwerch, N., & Kokaew, U. (2017). Design computer-assisted learning in an online Augmented Reality environment based on Shneiderman's eight Golden Rules. In *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-5). IEEE.
- [10] Avila-Pesantez, D. F., Vaca-Cardenas, L. A., Delgadillo Avila, R., Padilla Padilla, N., & Rivera, L. A. (2018). Design of an Augmented Reality Serious Game for Children with Dyscalculia: A Case Study. *Technology Trends*, 165–175. doi:10.1007/978-3-030-05532-5_12.
- [11] Plecher, D. A., Eichhorn, C., Kindl, J., Kreisig, S., Wintergerst, M., & Klinker, G. (2018). *Dragon Tale - A Serious Game for Learning Japanese Kanji*. *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts - CHI PLAY '18 Extended Abstracts*. doi:10.1145/3270316.3271536.
- [12] Thongchum, K., & Charoenpit, S. (2018). A conceptual design of Kanji mobile application with augmented reality technology for beginner. *2018 5th International Conference on Business and Industrial Research (ICBIR)*. doi:10.1109/icbir.2018.8391183.
- [13] Zainal Rasyid Mahayuddin, A F M Saifuddin Saif (2020) Augmented Reality Based AR Alphabets Towards Improved Learning Process in Primary Education System. *Journal of Critical Reviews*, 7 (19), 514-521. doi:10.31838/jcr.07.19.66.
- [14] Ishaq, K., Zin, N. A. M., Rosdi, F., Jehanghir, M., Ishaq, S., & Abid, A. (2021). Mobile-assisted and gamification-based language learning: a systematic literature review. *PeerJ Computer Science*, 7, e496.
- [15] Kaleel, S. & Harishankar, S. (2013). Applying Agile Methodology in Mobile Software Engineering: Android Application Development and its Challenges. *Computer Science Technical Reports* 11.
- [16] Knott, D. (2015). *Hands-on mobile app testing: a guide for mobile testers and anyone involved in the mobile app business*. Addison-Wesley Professional. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (Seven ed.)*. Upper Saddle River, NJ Prentice Hall: Pearson. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (Seven ed.)*. Upper Saddle River, NJ Prentice Hall: Pearson.
- [17] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (Seven ed.)*. Upper Saddle River, NJ Prentice Hall: Pearson. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (Seven ed.)*. Upper Saddle River, NJ Prentice Hall: Pearson. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (Seven ed.)*. Upper Saddle River, NJ Prentice Hall: Pearson.
- [18] Fkrudin, A., Yusoff, M., Romli, A. B., Pengajian, J., Politeknik, A. & Omar, U. (2018). Kebolegunaan Aplikasi Mudah Alih (Mobile Apps) Bagi Kursus Sains, Teknologi Dan Kejuruteraan Dalam Islam (M-Istech) Di Politeknik Malaysia (Usability of Mobile Application (Mobile Apps) in The Course of Science, Technology and Engineering in Islam (M-IST. *Malaysian Online Journal of Education* 2(1): 18–28.
- [19] Ramli, R. Z., Kapi, A. Y., & Osman, N. (2015). Visualisation makes array easy. In *Proceedings of the 2015 International Conference on Testing and Measurement: Techniques and Applications (TMTA'15)*. CRC Press (pp. 381-384).
- [20] Chen, P., Liu, X., Cheng, W., & Huang, R. (2016). A review of using Augmented Reality in Education from 2011 to 2016. *Lecture Notes in Educational Technology*, 13–18. doi:10.1007/978-981-10-2419-1_2.
- [21] Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education*, 14(1). doi:10.1186/s41239-017-0062-1.
- [22] Alkhatabi, M. (2017). Augmented Reality as E-learning Tool in Primary Schools' Education: Barriers to Teachers' Adoption. *International Journal of Emerging Technologies in Learning (iJET)*, 12(02), 91. doi:10.3991/ijet.v12i02.6158.
- [23] Alsubhi, M. A., Sahari, N., & Wook, T. T. (2020). A Conceptual Engagement Framework for Gamified E-Learning Platform Activities. *International Journal of Emerging Technologies in Learning (iJET)*, 15(22), 4-23.
- [24] Khaleel, F. L., Ashaari, N. S., & Wook, T. S. M. T. (2020). The impact of gamification on students learning engagement. *International Journal of Electrical and Computer Engineering*, 10(5), 4965.
- [25] Zaini, N. A., Noor, S. F. M., & Wook, T. S. M. T. (2019). The Model of Game-based Learning in Fire Safety for Preschool Children. *Editorial Preface From the Desk of Managing Editor*, 10(9).

Emotional Evocative User Interface Design for Lifestyle Intervention in Non-communicable Diseases using Kansei

Noor Afiza Mat Razali¹, Normaizeerah Mohd Noor², Norulzahrah Mohd Zainudin³
Nur Atiqah Malizan⁴, Nor Asiakin Hasbullah⁵, Khairul Khalil Ishak⁶
National Defence University of Malaysia, Kuala Lumpur, Malaysia^{1, 2, 3, 4, 5}
Management and Science University, Selangor, Malaysia⁶

Abstract—The advancement of technology has led to the development of an artificial intelligence-based healthcare-related application that can be easily accessed and used to assist people in lifestyle intervention for preventing the development of non-communicable diseases (NCDs). Previous research suggested that users are demanding a more emotional evocative user interface design. However, most of the time, it has been ignored due to lack of a model that could be referred in developing emotional evocative user interface design. This creates a gap in the user interface design that could lead to the ineffectiveness of content delivery in the NCD domain. This paper aims to investigate emotion traits and their relationship with user interface design for lifestyle intervention. Kansei Engineering method was applied to determine the dimensions for constructing emotional evocative user interface design. Data analysis was done using SPSS statistic tool and the result showed the emotional concepts that are significant and impactful towards user interface design for lifestyle intervention in NCD domain. The outcome of this research shall create new research fields that incorporate multi research domain including user interface design and emotions.

Keywords—Emotion; Kansei; non-communicable diseases; lifestyle intervention; user-interface design

I. INTRODUCTION

Most technologies in the healthcare domain user interface design are equipped with a clear and simple interaction approach that includes a text and button. However, the new display technology paradigm has shifted to touch-screen or semi-transparent display, allowing a new approach for visual human-machine interaction which more complex and efficient. The emergence of the new display technologies and artificial intelligence (AI) resulted in the adaptation of devices interface navigation that depending on the user's behaviour [1]. Recently, the chatbot has seemed to be a potential tool in communicating the context of healthcare effectively and improved patient education and treatment compliance[2] with an abundance of technologies that have been developed to assist people on a daily basis. However, affective interaction principles are not able to be realized yet [3]. Previous research showed that users are demanding for more emotional evocative user interface design [4]. Nevertheless, most of the time, it has been ignored due to the lack of a model that could be referred to in developing user interface design (UID) that is

emotionally evocative[5]. This creates a gap in the user interface design that could lead to the ineffectiveness of content delivery by the personal healthcare application.

Assessment of emotion related to UID is vital for the utmost user experience and effectiveness in using the application to achieve the desired usage objectives. Researchers are focusing mainly on how to design and redesign UID by examining the possible UID cognitive features such as design colour, characteristic, size, contents and messages [6]. The process of redesigning UID has been however carried out on a trial-and-error basis, which led to difficulties in selecting the most suitable design that can increase usage efficiency. Although the emotional embedded design is preferable by users, there are inadequate studies done on this topic.

To improve the human factor in interacting with a machine, the existing methodology namely Kansei Engineering (KE) is capable of guiding UID designers in developing UID that is in line with the complexity of dimension in a lifestyle intervention and user experience-based application design. Thus, in the UID context, various dimensions in a lifestyle intervention can be mapped to establish a joint representation of an optimum function of lifestyle intervention. Since KE is effective enough in addressing all dimensions needed for the user interface design, this paper proposed Emotional Evocative User Interface Design for Lifestyle Intervention in Non-communicable diseases (NCDs) domain using KE.

This paper is organized as follows: In Section II we present a brief overview of our research including advice and recommendation in avoiding the development of NCDs, the criteria of chatbot UID in smartphone and an overview of Kansei Engineering. Section III follows with details of our pilot study that utilize chatbot interfaces to evaluate the relationship between chatbot UID and user's emotion trait. Finally, we conclude in Section IV with a discussion on the result investigation and determined the significant emotions for the lifestyle intervention for NCDs development.

This research is fully supported by the National Defence University of Malaysia (UPNM) under Short Grant UPNM/2020/GPJP/ICT/6. The authors fully acknowledged UPNM and Ministry of Higher Education Malaysia (MOHE) for the approved fund, which made this research viable and effective.

II. RELATED WORK

A. Non-Communicable Diseases

Non-communicable disease (NCD) cases are a global health problem that continues to rise globally. The pandemic Covid-19 that is forcing people to be at home when there is no need to go out contributed to less exercise and overeating, causing the increase of NCD developments[7]. The development of NCDs do not result from germs and viruses, but individual risky behaviours such as tobacco use, alcohol consumption, unhealthy diet and inactive lifestyle [8],[9],[10]. Therefore, it is important to undergo a regular health screening to help to identify any early signs of chronic disease to reduce the socio-economic burden on an individual [11] and turn it to reduce morbidity and mortality as well as improve health and prevent disease [12], [11], [13].

A healthy lifestyle is classified into two categories [14], which are non-modified (age, family history) and modified (no obesity, healthy diet, physical activity and no smoking) variables that contribute to having a longer life with good health [13], [15]. This study utilised modified variables to measure user's emotion for lifestyle intervention compliance including food and nutrition intake. According to the previous research, healthy eating is based on the right amount of nutrients (protein, fat, carbohydrate, vitamins and minerals) and proportions of foods recommended by the food pyramid for daily requirement [16]. Meanwhile, the best amount of food serving depends on the types of exercise done by an individual. For instance, higher intensity activity increases energy expenditure and fuel for muscles. Besides, physical activity can balance energy by changing appetite and cutting food intake [17]. According to WHO, adults aged between 18 to 24 should do at least 150 minutes per week for mild physical exercise or at least 75 minutes per week for intense physical activity or both physical activities[18].

With the advancement of technology, lifestyle intervention can be done using the application to assist the users in their lifestyle. Nowadays, technology is widely used as a healthcare tool in spreading awareness and promoting health since it is effective, convenient, and cost-effective. Recently, several chatbots such as Quro, Shihbot and Mandy have been used in public health to reduce the burden of low and middle-income countries[19]. Since a decade ago, smartphones have been gaining popularity with mobile health (mHealth) applications used to conduct lifestyle interventions in gaining better health at a lower cost [20]. However, most of the applications are not considering the user's emotion during the designing and development phase that creates a gap in the established UID. Therefore, an emotional embedded user interface design for lifestyle intervention to prevent NCDs development should be developed to ensure the effectiveness of the applications.

B. Significance of Emotion in Lifestyle Intervention Compliance

Emotions are one of the essential properties of human-to-human interaction and complex process that consists of feeling, behaviour and psychology influenced by surrounding[21]. According to Legg et al., physicians ability in determining the current patient's emotion will improve the

level of compliance to the instruction given [22] due to the patient's satisfaction and adherence to treatment influenced by the level of understanding towards the instruction of lifestyle interventions. Anger, frustration and irritation due to negative experience caused by inappropriate health service providers will lead to non-adherence behaviour by the patients[23]. Legg et al. also emphasized that anxious emotions (worry and nervousness) trigger difficulty for individuals to understand instruction given by health providers and impede people's decision making[22].

Emotional issues cannot be ignored and should be taken into consideration while implementing any type of lifestyle intervention and more research is needed in this area [22]. In addition, Lockner & Bonnardel also highlighted the importance of studying emotions especially in the context of UID [24]. Therefore, it is vital to determine the emotional expression of an individual before implementing any lifestyle intervention to avoid the development of NCD effectively. As a result, people will be able to comply with instruction provided by the healthcare provider and lead a healthy lifestyle [23].

C. Kansei Methodology

Kansei Engineering is the combination of technology between Kansei and engineering realms, which is frequently used for product development that includes components such as desired need, emotion or sense [25]. Kansei can be utilised to establish a new product design based on users' evaluation of a product emotionally. Kansei integrates the product based on knowledge of what the users feel. Thus, Kansei is useful in influencing implicit users' insight related to the product design element and incorporating these insights with a new product[26]. Recently, people have shown interest in the application of Kansei in various fields such as industrial products, healthcare, education and e-commerce[27]–[29]. Furthermore, Kansei has been proven successful in measuring human emotions toward services and products such as in designs of eyewear, popup box and e-commerce websites desired by users[30], [31]. In the realm of information security, Kansei is utilised to design information security-related UID [32].

Kansei expresses human's feeling towards artefact, situation, or environment. Kansei refers to an external process that is tacit or known as a function of the brain at a higher level. This means that Kansei cannot be measured directly but can be measured partially or indirectly through a quantitative method using a self-reporting system like Semantic Differential (SD) scale, Different Emotional Scale (DES) or free labelling system. The measurement is done using Kansei Checklist, a form of a questionnaire that includes emotional keywords or known as Kansei Words (KW). KW is a word representing a user's emotions [6]. Commonly, the selection of KW is done based on the literature review and experts' advice. Normally, Kansei is used to measure the emotion of a user towards a product design. In this study, Kansei was applied using a quantitative method with KW, a form of questionnaires distributed to participants to evaluate related evocative user interface design based on users' insight towards artefact.

In this work, we proposed that Kansei can be utilised as a methodology to measure users' emotion in a chatbot design where the result can serve as a guideline on how to design a UID that can be associated with emotional evocative user interface design focusing on a lifestyle intervention for NCD. In this study, a chatbot was used as the UID. Generally, a chatbot can be used to spread awareness, promote health and act as a medium of interaction in lifestyle intervention to ensure a healthy lifestyle practice towards preventing the development of NCD.

D. Emotion Evocative User Interface Design (UID)

The most important factors in designing chatbot UID for lifestyle intervention are: 1) to make sure that the chatbot interaction with the target user is pleasant and comfortable to be used emotionally, 2) the chatbot could use a similar approach in understanding the target user like the one used by the health provider and 3) to create a relationship with target user to use the chatbot continuously. The dependency on technology has increased in people's daily activities whether in education, entertainment, business, or healthcare. Hence, the use of technology in assisting individuals in lifestyle modification for NCD intervention is convenient and accessible as smartphones are easily acquired today. Previous studies have shown the significance of technology in lifestyle intervention, which can help people with lifestyle modification and motivation [33], [34].

To increase emotional evocative user interface design in chatbot, user preferences are used as a guideline in designing UID and are the main concern of designers [35], [36]. Researchers have suggested that gender, age preference and emotion factors have to be included in the technology that addresses user's feelings, process and behaviour [37]. Moreover, embedded culture in UID affects product development [38], [39]. Thus, the influence in designing the users' interface and understanding attraction depends on how the users use the technology.

Before starting to develop the chatbot UID, it is important to decide the personality of the bot. According to Pricilla et al., the personality should be developed based on the bot's function, target audience, and type of task that the bot must complete [35]. Creating the right personality for the bot will boost the user experience and interaction. Besides, female-gendered chatbots have a huge effect on gaining trust for both women and men because female bots are perceived to be friendlier and trustworthy [36].

To provide a good user experience, the bot may also need to support rich interactions such as audio, pictures, and maps. Suggestion criteria in designing chatbot followed by user preference, gender, age and culture preferences include typefaces[37], colour [38], emoticon[1], avatar[5], tone[5] and voice[5].

1) *Typefaces*: People's impressions of books, packages, signs, and screen interfaces are affected by typeface elements. The typefaces used have a huge impact on how users view a chatbot and appreciate the emotion evoked while reading the contents. Fig. 1 shows two types of typefaces based on the finding by Candello et al. that stated, a robot-like typeface

(OCR) was more identified as a machine in a chat, rather than script typefaces (Bradley). Typefaces (Bradley) were less perceived as human in a chat. These show that script typefaces give an impression of more human like emotional conversation and have higher ratings of trust than robot-like speech.[37] Besides, the author suggested, designers can select different typefaces based on target users' familiarity with chatbots to give a favourable first impression. People who have had less experience with chatbots need additional elements or content reinforcement in order to view chatbots as less emotionless machines and designers' role in decreasing the typeface bias are vital.

Typeface 1: A robot-like typeface (OCR-A neutral typefaces such as Helvetica and Georgia)

abcdefghijklmnop
qrstuvwxyz

Typeface 2: Mimics handwriting (Bradley)

abcdefghijklmnop
qrstuvwxyz

Fig. 1. Type of Typeface [37].

2) *Colour*: Colours are also associated with other feelings and emotions. Colours not only affect our attitudes and perceptions toward a product but are also correlated with variations in UI trusting behaviour. Most studies agree that blue is perceived as a cold colour, while red is perceived as a warm colour. The author in [38] believes that the warmth of the user interface has a positive effect on trusting behaviour. However, pleasant colours and a joyful atmosphere in a chatbot should be surrounded by cheerful and lively design settings for a more positive effect. The colours design and atmosphere of a chatbot can result in a positive effect on user's attitudes toward it, which in turn influences compliance to the advice given by the chatbot in having a healthy lifestyle.

3) *Emoticon and Gifs*: Emojis are defined as the typographic display of a facial representation that used to express communication feelings and emotions[39]. Emojis have the ability to change the tone of a message. Furthermore, emojis are added with text and other platforms such as stickers and GIFs to expand and complement the text's expressive message. Unlike plain text, which is informative and brings the meaning of a message within the text, emojis are richer in terms of meaning that show stronger emotional behaviour. For example, the text "That not funny" and "That not funny 😞💔" convey different meanings. The first sentence sounds serious whereas the second one looks sad and heartbroken[1].

4) *Avatar*: An avatar may incorporate "joy-of-use" into the conversation and inspire the user with friendly encouragement and the right answers to promote a positive

work climate. A character's trustworthiness is also important in natural communication. As a result, each character's exterior appearance, speech characteristics, gesture, and general movement can vary. Fig. 2 shows alarm messages which require immediate actions. The red colour and the fireman avatar was representing the warning and alert towards the situation[5].



Fig. 2. Avatar for Alarm Message [5].

This criterion is a significant element in measuring users' emotional trait towards interaction with a chatbot using all the senses of vision, hearing, sensitivity and scent that resulted in a reaction established and stored in the human brain by experiencing all the impressions. Thus, this created a reaction of human behaviour towards the design of a chatbot. In this study, we proved that the data gathered using Kansei Engineering methodology are capable of providing a guideline to UID designers in developing UID that is able to address the emotion complexity dimension for NCDs lifestyle intervention and user experience-based application design. Thus, in the UID context, various dimensions in NCDs can be mapped to establish a joint representation of an optimum function of NCDs lifestyle intervention from the perspective of chatbot design.

III. EMOTIONAL EVOCATIVE USER INTERFACE DESIGN FOR LIFESTYLE INTERVENTION IN NON-COMMUNICABLE DISEASES

Emotional evocative UID can be realized if there are enough data gathered to explain the relationship between design and how it will affect user emotion in a particular domain. Emotion embedded UID will enable the optimum joint representation for lifestyle intervention to avoid NCDs development. Communication using chatbot has the potential to deliver message according to user's need if the chatbot design considers the emotion that capable to increase the positive effect on user's attitudes. This study adopted a model proposed by [40] to determine the emotional trait that will increase the adherence to instructions related to lifestyle intervention to prevent the development of NCDs. However, in this study, some modifications were made to this model to the suitability of the study scope. This study was conducted in four phases, which were 1) Identifying instrument of lifestyle intervention in NCD 2) Emotional measurement 3) Emotional Conceptualisation and 4) Design Requirement Formulation. This paper reports the result of a pilot study performed to

determine the reliability of the research methods and test the subject recruitment strategy.

For this study, the process for determining emotion assessment for chatbot interface for NCDs lifestyle intervention is shown in Fig. 3. The first phase was the process of identifying the artefact of the chatbot application interface from the Google App store and journals. The chatbot UID is defined as a visual layout design of artefacts that can interact with users [41].

Based on the findings, 8 out of 15 chatbots were selected using a matrix approach. The matrix approach is used to validate a specimen by checking the features of design and values that make up the appearance of each specimen [31]. Table I show the extraction of chatbot UID with the corresponding category, item and value, respectively.

A. Methodology

In the second phase, 71 KW was synthesised from a literature review and models including Panas-X that connected to the lifestyle intervention domain. In addition, a Kansei Affinity Cluster (KAC) was adopted for the process of synthesising KW and the confirmatory phase of KW. It was recommended that the researcher choose and pick only certain words that are important to the study [42]. Therefore, based on the 71 KW discovered, the words were filtered and selected based on the KAC model since KAC words are very suitable for representing emotions in chatbot UID. The result of this selected process was 16 KW. Finding from this process is as shown in Fig. 4.

After the KW was determined, Kansei Checklist was developed to evaluate emotions and was represented as a form of a questionnaire that consists of KW. The developed Kansei Checklist is shown in Fig. 5.

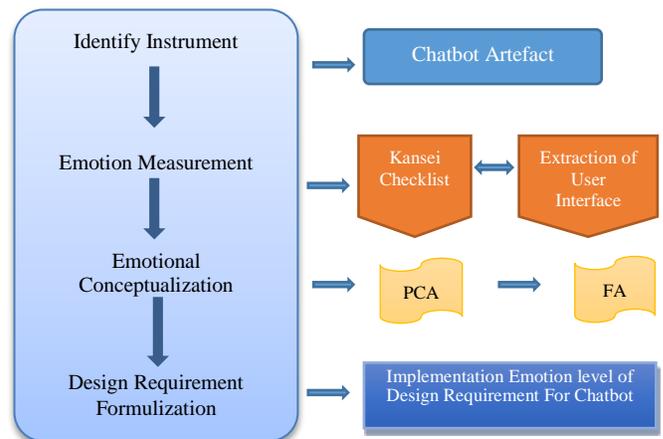


Fig. 3. Process for Emotion Assessment for Chatbot Interface using KE [40].

Frustration, Disgusting, Irritating, Annoying,
Concern, Painful, Stressful, Joyful, Cheerful,
Energetic, Satisfaction, Exciting, Exhausting,
Hopeful, Panic, Lively

Fig. 4. The Kansei Words [42].

Sample: _____

	5	4	3	2	1	
Frustration	<input type="checkbox"/>	Not Frustration				
Disgusting	<input type="checkbox"/>	Not Disgusting				
Irritating	<input type="checkbox"/>	Not Irritating				
Annoying	<input type="checkbox"/>	Not Annoying				
Concern	<input type="checkbox"/>	Not Concern				
Painful	<input type="checkbox"/>	Not Painful				
Stressful	<input type="checkbox"/>	Not Stressful				
Joyful	<input type="checkbox"/>	Not Joyful				
Cheerful	<input type="checkbox"/>	Not Cheerful				
Energetic	<input type="checkbox"/>	Not Energetic				
Satisfaction	<input type="checkbox"/>	Not Satisfaction				
Exciting	<input type="checkbox"/>	Not Exciting				
Exhausting	<input type="checkbox"/>	Not Exhausting				
Hopeful	<input type="checkbox"/>	Not Hopeful				
Panic	<input type="checkbox"/>	Not Panic				
Lively	<input type="checkbox"/>	Not Lively				

Fig. 5. The Kansei Checklist.

To conduct the evaluation experiment, KW was assembled into a 5-point Semantic Differential (SD) scale to form a checklist. Ten participants were recruited based on different demographic factors to evaluate KW appeal for eight artefacts that were selected in Phase 1. The samples of the artefact are as shown in Fig. 6. The artefacts' categories were made

according to the literature review obtained. However, not all categories and items were included in this pilot evaluation process due to the limitation on the allocation of timeframe and process delivery.

This evaluation was made by the selected participants with interest in lifestyle intervention using chatbot to prevent development of NCDs. The process was conducted in a close environment to prevent the emotion changes of the participants between five to six hours.

In this study, the demographic factors included sex, race and age in between 18 and 30. In this study, 10 smartphones were given to all subjects to show the artefacts simultaneously by the facilitator. The participants were asked to give a score based on their feelings on the Kansei Checklist for each sample product chatbots (artefacts).

B. Result

Result analysis started by determining the Cronbach's alpha of Kansei Checklist. Cronbach's alpha is a convenient test used to reach a reasonable reliability standard [50]. Cronbach's alpha reliability commonly ranges from 0 (very unreliable) to 1.0 (perfect reliability). The nearer the Cronbach's alpha to 1.0, the higher is the internal consistency of the scale item. Based on Cronbach's alpha rules of thumb recommended value; above 0.9 is excellent, above 0.8 is good and above 0.7 is acceptable[51]. This value also relies on the number of items that affect whether the scale reliability will be higher or lower. Therefore, in this study, Cronbach's alpha was performed in SPSS Statistics using Reliability Analysis.

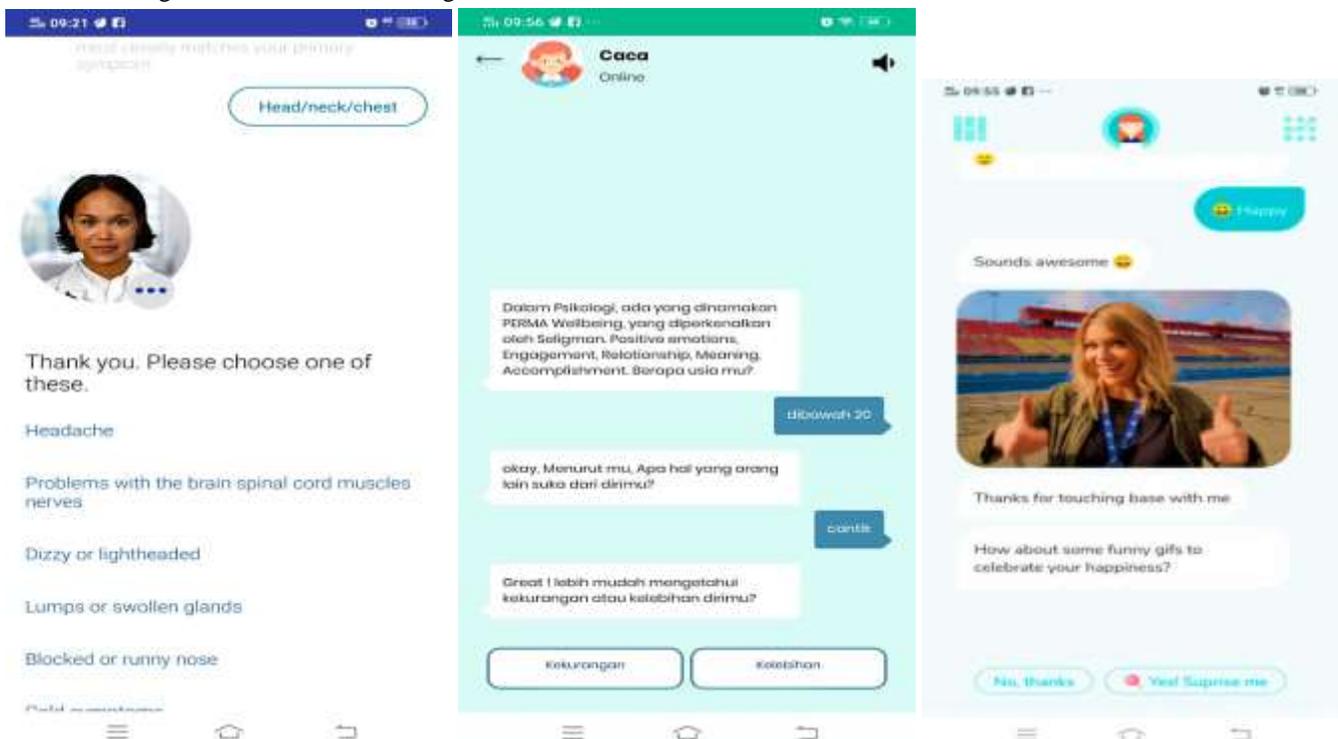


Fig. 6. The Kansei Checklist.

TABLE I. EXTRACTION OF CHATBOT UID

Category	Item	Value
Timing [40],[43]	Appear of notification	Seconds, Minutes, Hours
	Speed of reply	Fast, Medium, Slow
	Appear of location box	Above, Below, Centre, Right, Left
Typography[37]	Font Type	Regular, Italic, Bold
	Font Size	Small, Medium, Large
	Font Colour Reply	Red, Blue, Green, White, Black, Grey, Others
Emoticon[39],[44]	Type of emoticon	Unicode, Graphic
Colour[45]	Colour of header	Red, Blue, Green, White, Black, Grey, Others
Size[46]	Size of screen box	Half, Full
Button [47]	Size of button reply	Small, medium, large
	Colour of button reply intensity	Light, Dark
	Shape of button reply	Round, Rectangle, Oval, Square
Profile Picture [5]	Location of button reply appear	Centre, Right, Left, Down, Up
	Frame of Avatar	Oval, Round, Square,
Media[48]	Type of Avatar	Robot, Human, others
	Image	Less, More
	Gif	Yes, No
Animation [49]	Motion path of text	Appear, Fade, Fly-in, Float in, Wipe
Voice tone	Type of conversation	Robotic, Natural

TABLE II. EXTRACTION OF CHATBOT UID

Cronbach's Alpha	Cronbach's Alpha for Standardized Items	Number of Items(N)
0.771	0.757	128

Based on Table II, the alpha coefficient for 128 items was 0.771, which indicated that the items have high internal consistency. Meanwhile, the Cronbach's Alpha standard was 0.757. Alpha coefficient values between 0.60 and 0.70 have been shown to be acceptable in exploratory research, whereas range values between 0.70 and 0.95 are known to be satisfactory to a good degree of reliability [52]. Thus, it can be concluded that the reliability of the Kansei checklist was verified.

Then, the analysis was performed using principle loading analysis (PCA). PCA is an exploratory data analysis method that requires a technique to reduce a large dataset to a small set, but still maintains the amount of information in a large set[53]. There are three types of PCA namely PC Loading, PC Score and PC Vector. In this study, PC Loading was used, which represented the evaluation of how KW affects specimens. Fig. 7 shows the result of PC Loading for Component 1 and Component 2.

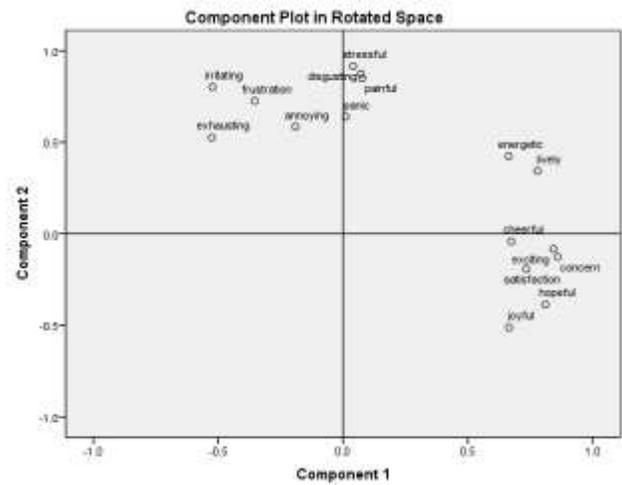


Fig. 7. PC Loading for Pilot Study.

From Fig. 7 above we observed that a good distribution of KW existed on both axes, which indicates an efficient evaluation. Based on a previous study, the loading can range -1.0 to 1.0. Furthermore, loading closer to -1 or 1 means that the component is highly influenced by the variable. While loading close to 0 means that the variable has a lower influence on the component [54].

In this loading, plot "Concern", "Hopeful", and "Exciting" have a large positive PC loading on component 1 (x-axis), whereas "Exhausting" has a moderate negative loading, making this component to be marked as an "Attractiveness". This study assumed that chatbots with a high score on component 1 of the PC axis would have a greater sense of attraction and vice versa.

Meanwhile, "Stressful" and "Disgusting" have a large positive PC Loading on component 2 (y-axis) near to 0; so, this component was marked as a "Hostility". This study assumed that chatbots with a higher score on component 2 of the PC axis would tend to feel hostility and vice versa.

Then, Factor Analysis (FA) was conducted to identify significant KW variables for the chatbots that reflect adherence of people in the chatbot features in a lifestyle intervention. The result was used to refine the outcome of PCA. FA is a technique of data reduction to describe ambiguity in terms of unobserved explanatory variables named factors [55]. Table III shows the result of FA after varimax rotation.

Based on this table, the first factor indicates the majority contribution with 28.109%, followed by the second factor with 24.216%. The majority of factors represented both variables. Therefore, as can be observed in this table, both Factor 1 and Factor 2 have a dominant KW affect, representing 28.109% and 52.325% = of the variability. Meanwhile, the third and fourth-factor contributions were 19.36% and 15.243%, respectively. Thus, Factor 2, Factor 3 and Factor 4 were seen with 52.325%, 71.693% and 86.936% of variability, respectively.

TABLE III. CONTRIBUTION AN ACCUMULATED CONTRIBUTION

Factor	Contribution (%)	Accumulated Contribution (%)
1	28.109	28.109
2	24.216	52.325
3	19.368	71.693
4	15.243	86.936

The results of factor loading after rotation of the varimax are shown in Table IV. In chatbot design, variables with a higher score are considered as important factors. This study defined 0.7 as the standard score. However, a slightly lower score can be also known as a significant concept[56].

Factor 1 consisted of “Disgusting”, “Painful”, “Frustration”, “Stressful”, “Panic”, and “Irritating”. These variables were labelled as the concept of “Hostility”. Factor 2 consisted of “Satisfaction”, “Hopeful”, “Concern” and “Lively”. These variables were labelled as the concept of “High-spirited”. Factor 3 comprised “Cheerful”, “Exhausting” and “Concern”, which were labelled as the concept of “Complicated”. Factor 4 consisted of “Energetic”, “Annoying” and “Exciting”. These variables were labelled as the concept of “Hyperactive”. This study reflected the common procedure in KE using representative terms for naming each factor group, the one that could effectively define the factor group. There is nothing right or wrong in naming group factors, except that the name must match the variables in factors.

TABLE IV. FACTOR LOADING FOR KW

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Disgusting	0.908	0.125	-0.209	
Painful	0.896			
Frustration	0.796		-0.414	-0.324
Stressful	0.794	-0.1	-0.116	0.439
Panic	0.68	-0.381	0.32	0.156
Irritating	0.668	-0.602	-0.356	0.124
Satisfaction		0.923	0.101	
Hopeful	-0.204	0.847	0.408	
Concern	-0.158	0.748	0.226	0.485
Lively	0.43	0.697	0.228	0.27
Cheerful	0.206	0.295	0.835	
Exhausting	0.387	-0.117	-0.816	
Joyful	-0.34	0.336	0.792	
Energetic	0.238	0.297	0.134	0.842
Annoying	0.209	-0.41	-0.431	0.754
Exciting	-0.177	0.471	0.405	0.704

From the experiment results, we determine that the concept of emotion in Chatbot UID is structured by four variables as evident from the outcome of FA, which was Hostility, High-Spirited, Complicated and Hyperactive. However, it can be concluded that Factor 1 and Factor 2 were more relevant concepts of Kansei for Chatbot due to the

52.325% of data variance. Thus, to obtain an optimal outcome, two factors were used namely Hostility and High-spirited, which were very significant emotion concepts. Whereas the other two factors were also significant, which were Complicated and Hyperactive, but have a poor impact to be used in designing the chatbots that include targeted emotions as supporting elements.

C. Discussion

This paper has described our research to establish emotions evocative chatbot UID in the healthcare domain. This finding will be useful in future research concerning the use of KE in the measurement of emotion, which integrates the user's subjective perception in a UID. In this study, we had extracted 16 KW related to lifestyle intervention from previous studies and eight chatbots artefacts have been identified using multilinear techniques. For this pilot study, the significant KW has been chosen using FA. Participants were recruited from different demographic factors that include a variety of analysis and a more precise relationship to reflect the specific regional field of NCDs. FA revealed more detailed findings compared to PCA. However, PCA illustrated a larger picture of the emotion structure, where it is shown that emotion was strongly affected by the first PC, which was “Attractiveness”, whereas the second main component “Hostility” has a poorer effect. From this pilot study, it has been determined that the significant emotions for designing UID chatbot for the lifestyle intervention in NCDs domain are hostility and high-spirited. These emotions will be the pillars in guiding the chatbots' design requirement that could influence emotion level for NCDs lifestyle interventions.

Based on the extraction of chatbot UID that has been chosen, the results show that evaluation using the Kansei Engineering method have revealed that two significant factors which are “Hostility and “High-spirited” as likely to have strong healthcare awareness and influences to individual's behaviour in using a chatbot. An individual is comfortable, pleasant and confident to comply with the chatbot guidance and instruction when supported with these two emotional factors. In this study, we selected UID from the applications in the Google App store as our artefacts since many applications are related to the healthcare domain and free to access.

IV. CONCLUSION AND FUTURE WORK

As a conclusion, this paper has conceptualised emotion assessment in UID for lifestyle intervention especially NCDs domain using KE methodology. The results were presented as a pilot study to determine the reliability of its data collection towards a small number of participants.

The next step is to perform activities with a larger group of participants to obtain more comprehensive data to determine emotion assessment in the NCDs domain for lifestyle intervention using Kansei Engineering method. The result could be utilised to create the guidelines on the representation of emotions towards the design of chatbot in the lifestyle intervention for NCDs domain. The guideline can be used by developers and software designers in designing emotionally evocative chatbot to guide lifestyle intervention and contribute to reducing the risk factor in developing NCDs. The most

important factor in designing a chatbot for lifestyle intervention is to make sure that chatbot interaction is pleasant and comfortable to be used emotionally and users have a tendency to share and follow the instructions guided by the chatbot to prevent the development of NCDs. This is to ensure that the users are comfortable and willing to cooperate upon receiving instructions of any lifestyle intervention in NCDs.

This research was able to offer insight into people's attitudes, sensitivity, and knowledge of lifestyle interventions for preventing NCD development issues. As a result, the findings of this study are recommended to be used as a foundation for future research into the relationship between emotion and UID using a chatbot or other conversation agent interfaces as an artefact. Nevertheless, a thorough investigation involving comprehensive evaluation and measurement of emotion is needed to expand the research findings and to extend more promising results.

ACKNOWLEDGMENT

This research is fully supported by the National Defence University of Malaysia (UPNM) under Short Grant UPNM/2020/GPJP/ICT/6. The authors fully acknowledged UPNM and the Ministry of Higher Education Malaysia (MOHE) for the approved fund which makes this research viable and effective.

REFERENCES

- [1] A. Fadhil, G. Schiavo, Y. Wang, and B. Yilma, "The Effect of Emojis when interacting with Conversational Interface Assisted Health Coaching & Information Tracking," 2018.
- [2] A. Palanica, P. Flaschner, A. Thommandram, M. Li, and Y. Fossat, "Physicians' Perceptions of Chatbots in Health Care: Cross-Sectional Web-Based Survey," *J. Med. Internet Res.*, vol. 21, no. 11, pp. 10.2196/12887, 2018.
- [3] S. A. Crossley and D. S. McNamara, "Adaptive educational technologies for literacy instruction," *Adapt. Educ. Technol. Lit. Instr.*, pp. 1–310, 2016, doi: 10.4324/9781315647500.
- [4] K. Lin, F. Xia, W. Wang, D. Tian, and J. Song, "System Design for Big Data Application in Emotion-Aware Healthcare," *IEEE Access*, vol. 4, pp. 6901–6909, 2016, doi: 10.1109/ACCESS.2016.2616643.
- [5] D. Ziegeler and D. Zuehlke, "EMOTIONAL USER INTERFACES AND HUMANOID AVATARS IN INDUSTRIAL ENVIRONMENTS," *IFAC Proc. Vol.*, vol. 38, no. 1, pp. 106–111, 2005, doi: <https://doi.org/10.3182/20050703-6-CZ-1902.01420>.
- [6] N. A. M. Razali, N. J. A. M. Saad, K. K. Ishak, and N. A. Hasbullah, "Fear Appeal Inducement in Pop up Design for IS Procedure Compliance," in *2019 2nd International Conference on Communication Engineering and Technology (ICCET)*, 2019, pp. 17–22.
- [7] The Lancet, "COVID-19: a new lens for non-communicable diseases," *Lancet*, vol. 396, no. 10252, p. 649, 2020, doi: 10.1016/S0140-6736(20)31856-0.
- [8] M. P. Kelly and F. Russo, "Causal narratives in public health: the difference between mechanisms of aetiology and mechanisms of prevention in non-communicable diseases," *Sociol. Heal. Illn.*, vol. 40, no. 1, pp. 82–99, 2018, doi: 10.1111/1467-9566.12621.
- [9] W.-Y. Low, Y.-K. Lee, and A. L. Samy, "Non-communicable diseases in the Asia-Pacific region: Prevalence, risk factors and community-based prevention," *Int. J. Occup. Med. Environ. Health*, vol. 28, no. 1, pp. 20–26, 2015, doi: 10.2478/s13382-014-0326-0.
- [10] J. Baird et al., "Developmental Origins of Health and Disease: A Lifecourse Approach to the Prevention of Non-Communicable Diseases," *Healthcare*, vol. 5, no. 1, pp. 1–14, 2017, doi: 10.3390/healthcare5010014.
- [11] I. N. Ojong, A. D. Nsemo, and P. Aji, "Routine Medical Checkup Knowledge, Attitude and Practice among Health Care Workers in a Tertiary Health Facility in Calabar, Cross River State, Nigeria," *Glob. J. Health Sci.*, vol. 12, no. 8, pp. 27–37, 2020, doi: 10.5539/gjhs.v12n8p27.
- [12] A. B. AL-Kahil, R. A. Khawaja, A. Y. Kadri, S. M. Abbarh, J. T. Alakhras, and P. P. Jaganathan, "Knowledge and Practices Toward Routine Medical Checkup Among Middle-Aged and Elderly People of Riyadh," *J. Patient Exp.*, p. 237437351985100, 2019, doi: 10.1177/2374373519851003.
- [13] A. V. Khera et al., "Genetic risk, adherence to a healthy lifestyle, and coronary disease," *N. Engl. J. Med.*, vol. 375, no. 24, pp. 2349–2358, 2016, doi: 10.1056/NEJMoa1605086.
- [14] B. Ghosn, S. Benisi-Kohansal, S. Ebrahimpour-Koujan, L. Azadbakht, and A. Esmailzadeh, "Association between healthy lifestyle score and breast cancer," *Nutr. J.*, vol. 19, no. 1, pp. 1–11, 2020, doi: 10.1186/s12937-020-0520-9.
- [15] A. Marques, M. Peralta, J. Martins, V. Loureiro, P. C. Almanzar, and M. G. de Matos, "Few European Adults are Living a Healthy Lifestyle," *Am. J. Heal. Promot.*, vol. 33, no. 3, pp. 391–398, 2019, doi: 10.1177/0890117118787078.
- [16] C. Report et al., "INTERNATIONAL JOURNAL OF NUTRITION ISSN NO : 2379 - 7835 Rising Interest Interactions Between the Water , Energy and Food Security Sectors Energy Mix," no. 3, pp. 17–21, 2020, doi: 10.14302/issn.2379.
- [17] N. Zhang and S. Bi, "Effects of physical exercise on food intake and body weight: Role of dorsomedial hypothalamic signaling," *Physiol. Behav.*, vol. 192, no. March, pp. 59–63, 2018, doi: 10.1016/j.physbeh.2018.03.018.
- [18] Y. J. Yang, "An Overview of Current Physical Activity Recommendations in Primary Care," *Korean J. Fam. Med.*, vol. 40, no. 3, pp. 135–142, May 2019, doi: 10.4082/kjfm.19.0038.
- [19] C. Huang, M. Yang, C. Huang, Y. Chen, M. Wu, and K. Chen, "A Chatbot-supported Smart Wireless Interactive Healthcare System for Weight Control and Health Promotion," in *Conference: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2018, pp. 1791–1795.
- [20] S. E. Bonn, M. Löf, C. G. Östenson, and Y. Trolle Lagerros, "App-technology to improve lifestyle behaviors among working adults - The Health Integrator study, a randomized controlled trial," *BMC Public Health*, vol. 19, no. 1, pp. 1–8, 2019, doi: 10.1186/s12889-019-6595-6.
- [21] F. Agrafioti, D. Hatzinakos, S. Member, and A. K. Anderson, "ECG Pattern Analysis for Emotion Detection," vol. 3, no. 1, pp. 102–115, 2012.
- [22] A. M. Legg, S. E. Andrews, H. Huynh, A. Ghane, A. Tabuenca, and K. Sweeny, "Patients' anxiety and hope: Predictors and adherence intentions in an acute care context," *Heal. Expect.*, vol. 18, no. 6, pp. 3034–3043, 2015, doi: 10.1111/hex.12288.
- [23] J. Heyhoe, Y. Birks, R. Harrison, J. K. O'Hara, A. Cracknell, and R. Lawton, "The role of emotion in patient safety: Are we brave enough to scratch beneath the surface?," *J. R. Soc. Med.*, vol. 109, no. 2, pp. 52–58, 2016, doi: 10.1177/0141076815620614.
- [24] D. Lockner and N. Bonnardel, "Emotion and Interface Design How to measure interface design emotional effect?," *Int. Conf. KANSEI Eng. Emot. Res. Emot.*, no. June 2015, pp. 10–25, 2014.
- [25] A. M. Lokman, A. A. Awang, A. R. Omar, and N. A. S. Abdullah, "The integration of quality function deployment and Kansei Engineering: An overview of application," *AIP Conf. Proc.*, vol. 1705, no. 1, pp. 20004–20012, 2016, doi: 10.1063/1.4940252.
- [26] A. Lokman, "KE as affective design methodology," *Proceeding - 2013 Int. Conf. Comput. Control. Informatics Its Appl. "Recent Challenges Comput. Control Informatics"*, pp. 7–13, Nov. 2013, doi: 10.1109/IC3INA.2013.6819139.
- [27] F. Shi, N. Dey, A. S. Ashour, D. Sifaki-Pistolla, and R. S. Sherratt, "Meta-KANSEI Modeling with Valence-Arousal fMRI Dataset of Brain," *Cognit. Comput.*, vol. 11, no. 2, pp. 227–240, 2019, doi: 10.1007/s12559-018-9614-5.
- [28] Z. Li et al., "Rule-based back propagation neural networks for various precision rough set presented KANSEI knowledge prediction: a case study on shoe product form features extraction," *Neural Comput. Appl.*, vol. 28, no. 3, pp. 613–630, 2017, doi: 10.1007/s00521-016-2707-8.

- [29] H.-C. Shen and K.-C. Wang, "Affective product form design using fuzzy Kansei engineering and creativity," *J. Ambient Intell. Humaniz. Comput.*, vol. 7, no. 6, pp. 875–888, 2016, doi: 10.1007/s12652-016-0402-3.
- [30] N. Afiza and M. Razali, "Fear Appeal Inducement in Pop up Design for IS Procedure Compliance," pp. 17–22, 2019.
- [31] N. K. Chuan, A. Sivaji, M. M. Shahimin, and N. Saad, "Kansei Engineering for e-commerce Sunglasses Selection in Malaysia," *Procedia - Soc. Behav. Sci.*, vol. 97, no. November, pp. 707–714, 2013, doi: 10.1016/j.sbspro.2013.10.291.
- [32] A. M. Lokman, "Design and emotion: The Kansei Engineering the Definition of Kansei," vol. 1, no. 1, pp. 1–14, 2010.
- [33] S. Reddy, J. Fox, and M. Purohit, "Artificial intelligence-enabled healthcare delivery," *Artif. Intell. Healthc. Deliv.*, vol. 112, p. 014107681881551, Dec. 2018, doi: 10.1177/0141076818815510.
- [34] B. Liu and S. S. Sundar, "Should Machines Express Sympathy and Empathy? Experiments with a Health Advice Chatbot," *Cyberpsychology, Behav. Soc. Netw.*, vol. 21, no. 10, pp. 625–636, Oct. 2018, doi: 10.1089/cyber.2018.0110.
- [35] C. Pricilla, D. P. Lestari, and D. Dharma, "Designing Interaction for Chatbot-Based Conversational Commerce with User-Centered Design," *ICAICTA 2018 - 5th Int. Conf. Adv. Informatics Concepts Theory Appl.*, pp. 244–249, 2018, doi: 10.1109/ICAICTA.2018.8541320.
- [36] D. C. Toader et al., "The effect of social presence and chatbot errors on trust," *Sustain.*, vol. 12, no. 1, pp. 1–24, 2020, doi: 10.3390/SU12010256.
- [37] H. Candello, C. Pinhanez, and F. Figueiredo, "Typefaces and the perception of humanness in natural language chatbots," in *Conference on Human Factors in Computing Systems - Proceedings*, May 2017, vol. 2017-May, pp. 3476–3487, doi: 10.1145/3025453.3025919.
- [38] F. Hawlitschek, L.-E. Jansen, E. Lux, T. Teubner, and C. Weinhardt, *Colors and Trust: The Influence of User Interface Design on Trust and Reciprocity*. 2016.
- [39] M. Shahbaz, T. Razzaq, and M. R. Gohar, "Emoji Usage by Android Consumers: A Gender-Based Inquiry Keywords: Gender, Kikaboard, Consumer and Emoji Introduction Every researcher has a purpose behind choosing an area for research work or investigation. For carrying out this study, researcher h," *Orient Res. J. Soc. Sci.*, vol. 2, no. 1, pp. 1–26, 2017.
- [40] N. A. M. Razali, K. K. Ishak, N. J. A. MdSaad, N. M. Zainudin, N. Hasbullah, and M. F. M. Amran, "Conceptualization of User's Rage Assessment Using Chatbot Interface by Implementing Kansei Engineering Methodology for Information Security BT - Proceedings of the 8th International Conference on Kansei Engineering and Emotion Research," vol. 1256, pp. 184–193, 2020.
- [41] S. Durrani and Q. S. Durrani, "Applying Cognitive Psychology to User Interfaces," in *Proceedings of the First International Conference on Intelligent Human Computer Interaction*, Springer India, 2009, pp. 156–168.
- [42] A. M. Lokman and K. A. Kamaruddin, "Kansei affinity cluster for affective product design," *Proc. - 2010 Int. Conf. User Sci. Eng. i-USER 2010*, no. January, pp. 38–43, 2010, doi: 10.1109/IUSER.2010.5716719.
- [43] Q. V. Liao, M. Davis, W. Geyer, M. Muller, and N. S. Shami, "What can you do? Studying social-agent orientation and agent proactive interactions with an agent for employees," in *DIS 2016 - Proceedings of the 2016 ACM Conference on Designing Interactive Systems: Fuse*, Jun. 2016, pp. 264–275, doi: 10.1145/2901790.2901842.
- [44] Y. Tang and K. F. Hew, "Emoticon, emoji, and sticker use in computer-mediated communication: A review of theories and research findings," *Int. J. Commun.*, vol. 13, no. May, pp. 2457–2483, 2019.
- [45] M. Sillic, "Improving warning messages adherence: can Maya Security Bot advisor help?," *Secur. J.*, vol. 33, no. 2, pp. 293–310, 2020, doi: 10.1057/s41284-019-00185-7.
- [46] Restyandito and E. Kurniawan, "The effect of screen size and interaction style on mobile device usability," *Eng. Lett.*, vol. 25, no. 4, pp. 354–359, 2017.
- [47] A. Følstad and P. Brandtzaeg, "Chatbots and the new world of HCI," *interactions*, vol. 24, no. 4, pp. 38–42, Jun. 2017, doi: 10.1145/3085558.
- [48] S. Djamasbi, T. Tullis, J. Hsu, E. Mazuera, K. Osberg, and J. Bosch, *Gender Preferences in Web Design: Usability Testing through Eye Tracking*, vol. 7. 2007.
- [49] J. Huhtala, A. H. Sarjanoja, J. Mäntyjärvi, M. Isomursu, and J. Häkkinä, "Animated UI transitions and perception of time: A user study on animated effects on a mobile screen," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2, no. January, pp. 1339–1342, 2010, doi: 10.1145/1753326.1753527.
- [50] N. F. Habidin, A. F. M. Zubir, N. M. Fuzi, N. A. M. Latip, and M. N. A. Azman, "Sustainable Performance Measures for Malaysian Automotive Industry," *World Appl. Sci. J.*, vol. 33, no. 6, pp. 1017–1024, 2015, doi: 10.5829/idosi.wasj.2015.33.06.257.
- [51] K. S. Taber, "The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education," *Res. Sci. Educ.*, vol. 48, no. 6, pp. 1273–1296, 2018, doi: 10.1007/s11165-016-9602-2.
- [52] A. Md Ramli, M. S. Mohd Zahari, M. Z. Suhaimi, and S. Abdul Talib, "Determinants of Food Heritage towards Food Identity," *Environ. Proc. J.*, vol. 1, no. 1, p. 207, 2016, doi: 10.21834/e-bpj.v1i1.217.
- [53] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 374, no. 2065, p. 20150202, Apr. 2016, doi: 10.1098/rsta.2015.0202.
- [54] C. Syms, "Principal components analysis," *Enycl. Ecol.*, vol. 3, no. 1, pp. 566–573, 2018, doi: 10.1016/B978-0-12-409548-9.11152-2.
- [55] A. G. Yong and S. Pearce, "A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis," vol. 9, no. 2, pp. 79–94, 2013.
- [56] S. A. Kadir, A. M. Lokman, T. Tsuchiya, and S. M. Shuhidan, "Analysing Implicit Emotion and Unity in Propaganda Videos Posted in Social Network," *J. Phys. Conf. Ser.*, vol. 1529, no. 2, pp. 25–27, 2020, doi: 10.1088/1742-6596/1529/2/022018.

Exploring the Socio-economic Implications of Artificial Intelligence from Higher Education Student's Perspective

Sarah Areef¹, Lobna Amouri², Nahla El-Haggar³, Aishah Moneer⁴
Computer Department, Imam Abdulrahman Bin Faisal University
Community College, Dammam, KSA

Abstract—As a result of the instability of oil prices, the economic prospects of the Gulf region are increasing their focus on new technologies. Thus, Saudi Arabia has demonstrated a strong commitment towards the development and implementation of Artificial Intelligence (AI) technologies as alternative sources for revenue and growth in line with globalisation, development, and the vision 2030. This paper examines the impact of AI in the Saudi Arabia community, especially for social and economic evolution. Special focus on the use of smart cars and smart cameras to monitor intelligently traffic, public services and national security is explored. A total of 424 participants from Eastern Province took part in this study. Analysis and discussion of the obtained results are also presented. The findings showed that 75.71% of participants mostly highly agreed about the AI economic impact leading to an increase in both government and business financial incomes. Whereas only 59.84% of participants mostly highly agreed about the social impact of AI as they are worried about AI ethical concerns, job loss and the changing workforce.

Keywords—Artificial intelligence; Saudi community; data analysis; social efficiency impact; economic productivity impact

I. INTRODUCTION

AI is based on the use of big data, artificial learning and robot technology applications that will have far-reaching implications for the economy (Clark, 2020) [1], jobs, social fabrics and culture sectors (Elsaadani et al., 2018) [2]. For the last few decades, the energy sector's presence in the Saudi economy has decreased dramatically, from almost two-thirds of GDP to 43 percent by 2015 ((Elsaadani et al., 2018) [2] (Stephens et al., 2019) [3]). The Kingdom's government has responded with a series of contiguous strategies based especially on economic diversification that aimed to expand non-oil industries to ensure sustainable development, recruitment and to decrease the dependency of the public sector on jobs. According to Accenture analysis (Elsaadani et al., 2018) [2], the common point between all these strategies is the determined and exhaustive adherence to AI. This analysis demonstrated that AI has the capacity to considerably increase economic development rates by contributing \$215 billion in annual gross value added (GVA) to the economy of Saudi Arabia by 2035. So far, the main idea beyond the Kingdom's

Vision 2030 national growth strategy is the technology sector (in government, business, services, industry, healthcare, education) which is rapidly empowered by AI. So far, the Traditional Economic Growth model shown in Fig. 1 demonstrates that capital, labour factors and new technologies are the key drivers of economic growth. Nevertheless, according to Accenture analysis, the new growth model treats AI as a new factor of economic growth rather than simply another technology-based productivity enhancer.

The objectives of the study are:

- To study the Saudi Arabia community's acceptance and willingness to deal with AI.
- To compare the future situation of Saudi Arabia's community in an AI scenario to a scenario without AI.
- To explore the AI impact on health, security, privacy (ethical considerations), safety, employment, education as well as public services of Saudi Arabia's population.

The remainder of this study is structured as follows: Section II explores a literature review about the implication of AI on social and economic development. Section III introduces the suggested methodology. Section IV clarifies the results and discussion. Finally, Section V illustrates conclusions, future work and limitations.

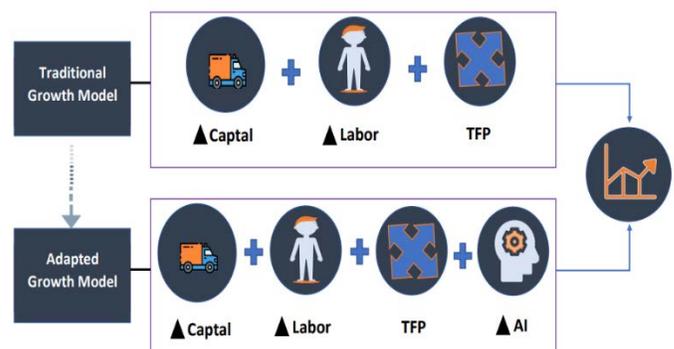


Fig. 1. Economic Growth Model (Elsaadani et al.,2018).

II. LITERATURE REVIEW

This section present synthesizes the existing research papers that treated first with the AI applications for smart cities (smart cars and cameras). Then, we introduced a literature review on the social and economic AI implications at the national and international level.

A. AI Applications in Smart Cities

A. Beg, et al. (2020) [4] analysed transportation vulnerabilities and emergency response systems; proposed a smart, autonomous UAV-enabled solution; and demonstrated the system in a simulated environment. Various traffic control and police system conditions were taken into consideration in the simulation: traffic light abuses and accident detection, mobile speed traps and automatic alerts, crowds.

Impedovo, et al. (2020) [5] studied 21 peer-reviewed papers that contributed to the broad range of artificial intelligence applications and demonstrated that under a more resource-efficient management and economy than ordinary cities, smart cities work. As such, advanced business models around smart cities have grown, leading to the development of smart companies and organisations that rely on advanced technologies.

S. Hoofar, et al. (2020) [6] presented a new approach to supporting smart cars by exploring Vehicle Ad Hoc Networks (VANETs), which are intelligent interconnected vehicle networks that can provide real-time traffic parameters such as the type, speed, direction and location of each vehicle. The studied intelligent transport system is commonly used in various applications such as traffic flow control, automatic parking systems and security compliance.

B. Social AI Impact

The researchers in [7] demonstrated that AI has unfavourable effects on disadvantaged segments of populations, and the effects on individual AI applications can vary. Some other researchers in [8] examined what type of human labour will be a replacement versus a complement to emerging technologies in the field of AI. Besides, researchers in [9] studied the impacts of rapid progress in artificial intelligence technologies on society. They discussed one such potential impact: the issue of machine learning system accidents, identified as unintended and harmful actions that could result from poor design of real-world AI systems. Researchers in [10] have examined the effect of the use of AI in medicine at three levels. For clinicians, mainly by quick, precise interpretation of images; for health systems, through enhancing workflow and the ability to minimise medical errors; and for patients, through enabling them to access their own health-promoting data. They also argued the existing limitations of using AI in medicine including bias, privacy and security, and lack of clarity, along with the potential directions for these applications. While researchers in [11], [12] [13], [14] and [15] used AI to build various systems for people with disabilities and aged persons who require support in their everyday lives or in recovery from emergency situations.

The proposed systems identified vital signs (Body Temperature, Blood Pressure, Blessing, Heart Pulse Rate, and Consciousness) as well as location/attitude monitored by the

sensor network. Besides, researchers in [16] focused on AI-based learning by trying to analyse and investigate how AI can be fruitfully exploited by universities in Saudi Arabia. In addition, they centralised on challenges of teaching/learning procedure and advantages in the implementation of Information and Communication Technologies (ICT) in a classroom, leading to delegate an academic committee to an AI-based context. The researchers in [17] created a short questionnaire and circulated it to four expert groups in 2012/2013. For instance, they detected that there was, in some areas, concern over high-level artificial intelligence growing up in a few decades and addressing its serious dangers for humanity. In other areas, these concerns have been overlooked or deemed science fiction, what the likelihood that the best experts allocated to high-level machine intelligence coming up within a given time frame, which threats they have seen in this growth, and how rapidly they have seen them evolve.

C. Economic AI Impacts

Approximately 47% percent of the all jobs in the U.S. are at risk from computerisation, according to Frey and Osborne (2013) [18]. Author (2015) [19] argues, however, that automation and technological advancement have not led to the obsolescence of human labour. Automation and labour are highly complementary, particularly for staff who are adaptable, innovative, and solution oriented.

E. Ernest, et al. (2019) in [20] demonstrated how AI has produced widespread worry of job loss and further increases in unfairness. The authors called, through the paper, for a reasonable outlook on the opportunities and risks from AI. For this, they discussed how skills policies and new ways of digital economy regulation are necessary to prevent further increases in market concentration, ensure adequate data protection and privacy and help redistribute the gains of productivity growth by combining wealth sharing, capital taxes and a decline in working time.

On the other hand, G. C. Allen (2019) in [21] addressed China's political thinking towards AI and national security and highlighted that China's leadership claims that being at the forefront of AI technologies is vital to the future of global military and economic-power rivalry.

M. Uzun (2020) in [22] addressed growing awareness about the negative impacts of AI among countries and proposed a solution modelling universal norms, international institutions and policies that ensure beneficial discovery of AI in order to provide a strategic equilibrium between major powers, international economic stability, and equitable economic control at national and international levels. Besides, O. Hassan (2020) in [23] described how the new mega-city Neom, currently under construction in Kingdom of Saudi Arabia, is targeting to integrate robotics and AI into various aspects of citizens' lives in order to generate earnings from key economic sectors for the future. The authors discussed also that this transition from an economy based on oil to AI targets to secure the survival of the Saudi community and meet the growing challenges of constructing a state around oil.

III. METHODOLOGY

This current study was prepared at the Computer Department of the Community College of Dammam, which is a college in Imam Abdulrahman Bin Faisal University that provides diploma graduates for both female and male students. All participants that took part in questionnaires answered closed-ended items which are likely to receive short answers. A Likert scale which is a close-ended, forced-choice scale is used in the questionnaire to provide a series of answers. This study used a 4-point Likert scale ranging from S.A (Strongly Agree), A (Agree), D (Disagree) and S.D (Strongly Disagree) to explore the participants' acceptance and willingness to deal with AI. Six criteria were used to develop the questionnaire including acceptance, security, ability, increase productivity, approval, and efficiency. A set of items were prepared for each domain according to five demographic profile items given at the beginning of the questionnaire. The demographic profile items include gender, identity, age-category, education, and standard of living. Besides, the conducted questionnaire explores the effect of specific AI applications in Saudi Arabia on social and economic developments. As well as, AI applications were discussed that include utilising smart cars and smart cameras for intelligent monitoring of traffic, public services, and national security.

A. Participants

A total of 424 participants took part in the questionnaire. The participants include students and teachers of Imam Abdulrahman Bin Faisal University, some other faculty members and employees in public and private sectors working outside the university. The groups we asked were categorised referring to the following criteria:

- Gender: Male or female.
- Identity: Saudi, non-Saudi.
- Age-Category: from 15 to 34 years, from 35 to 60 years.
- Education: High School Degree, Undergraduate Degree, Postgraduate Degree.
- Standard of living: Low, Moderate, High.

B. Data Collection

The conducted questionnaire is a cross-sectional one as collecting information from the respondents was conducted over a single period which was from 28 of February 2020 to 6 March 2020. It was prepared using the Google Forms tool and shared with participants by means of a link (see Appendix A). The sharing process took advantage of the best mobile messaging applications like WhatsApp, Viber, Snapchat and Facebook messenger that offer a superior substitute to email and text messaging thanks to built-in social networking characteristics and improved security. The questionnaire was appropriately configured to describe the views, concerns, obstacles and ideas of the participants in order to identify the measures needed to further strengthen the actual condition of AI in Saudi Arabia.

C. Data Analysis

Several steps were taken to analyse the data. Data was collected in an excel file and pictures were sorted out; percentages of the answers were calculated and studied thoroughly. The percentage for each item was determined using the mathematical formula 1, out of a total of 424 respondents who answered the questionnaire.

$$S_{i,j} = \frac{\text{Frequency of answers in each scale}_{i,j}}{\text{Total number of respondents}} * 100 \quad (1)$$

Where i is the number of items from 1,2,3,4...16 and j is the number of scales from 1 to 4. Note that the scale number 1 refers to scale Strongly Agree (S.A), 2 refers to Agree (A), 3 refers to Disagree (D) and 4 refers to Strongly Disagree (S.D).

The mean score ($M_{i,j}$) was also calculated for each item, to describe the strength of each one. In fact, weights (w) were assigned for each Likert scale. The larger number (4) was assigned to the most positive scale (S.A), 3 was assigned to (A), 2 to (D) and 1 to (S.D). The standard deviation was then applied to estimate the dispersion of each item relative to its mean. In fact, a second mean ($M_{0i,j}$) with the newest weights (W) (1, 4, 9 and 16) was explored to calculate the standard deviation $STDEV_{i,j}$. The mathematical formula 2, 3 and 4 represents respectively the first mean ($M_{i,j}$), the second mean ($M_{0i,j}$) and Standard Deviation $STDEV_{i,j}$:

$$M_{i,j} = \sum_{i,j} S_{i,j} * w_1 ; w_1 = w_j = 4, 3, 2, 1 \quad (2)$$

$$M_{0i,j} = \sum_{i,j} S_{i,j} * w_2 ; w_2 = w_j^2 \quad (3)$$

$$STDEV_{i,j} = \sqrt{M_{0i,j} - M_{i,j}^2} \quad (4)$$

Respectively, the mean score (M_c) and the standard deviation $STDEV_c$ were also calculated for each criterion c , where c is from 1 to 6 as shown in formulas 5 and 6. Then we used equations the formulas 7 and 8 to calculate the mean score (M_d) and the standard deviation $STDEV_d$ for each domain d , where d is from 1 to 2.

$$M_c = \frac{\sum M_{i,j} \text{ for each criterion}_c}{\text{Number of items for each criterion}_c} \quad (5)$$

$$STDEV_c = \frac{\sum STDEV_{i,j} \text{ for each criterion}_c}{\text{Number of items for each criterion}_c} \quad (6)$$

$$M_d = \frac{\sum M_c \text{ for each domain}_d}{\text{Number of criterion for each domain}_d} \quad (7)$$

$$STDEV_d = \frac{\sum STDEV_c \text{ for each domain}_d}{\text{Number of criterion for each domain}_d} \quad (8)$$

IV. RESULTS AND DISCUSSION

This section demonstrates and recognises the beneficial effects of adopting AI on the social domain and on specific criteria such as acceptance, security, ability, approval and efficiency. In addition, it identifies the advantages of AI in the economic domain to find out its effect on increasing the competitiveness of the Saudi community by using smart cars as well as implementing AI in working areas (companies, universities, hospitals...) to explore factors influencing the Saudi community's satisfaction with AI.

A. Demographics Profile Results

Table I shows the results of the questionnaire’s items regarding respondents’ demographic profile where the size of the sample is 424 Saudi community members. The demographic results signify 85% of the respondents to be female while 15% were male respondents. This can be explained by the fact that the college was originally a girl’s college, and the boy’s branch has recently been added, in addition to the lack of awareness between the sexes about the impact of artificial intelligence on society. According to the item “Age-Category”, most of the respondents (373 or 88%) of the sample population fall with 15- 34 years category while 51 respondents were between 35-60 years, representing 12% of the total sampled population. As for the item “Education”, the majority of the respondents (260 or 61%) have an Undergraduate Degree and 157 (37%) respondents have Postgraduate Degree. The remaining 7 (2%) of the total sampled population have a High School Degree. As for the item “Standard of Living”, most of the respondents (336 or 79%) have a Moderate Standard of Living, 73 of the respondents have a High Standard of Living representing (17%), and 15 of the respondents have a Low Standard of Living representing 4% of the total sampled population.

TABLE I. RESPONDENTS’ DEMOGRAPHIC PROFILE

Demographic Information’s	Options	Frequency	Percentage (%)
Gender	Male	63	15
	Female	361	85
Identity	Saudi	348	82
	Non-Saudi	76	18
Age-Category	15-34	373	88
	35-60	51	12
Education	High School Degree	7	2
	Undergraduate Degree	260	61
	Postgraduate Degree	157	37
Standard of Living	Low	15	4
	Moderate	336	79
	High	73	17

B. Analytical AI Impacts on Social and Economic Domains based on Item

The conducted questionnaire is a cross-sectional one as the collection information from the respondents was conducted over a single period of time, which was from 28 of February 2020 to 06 March 2020. It was prepared using the Google Forms tool and shared with participants by means of a link. The sharing process took advantage of the best mobile messaging applications like WhatsApp, Viber, Snapchat and Facebook messenger that offer a superior substitute to email and text messaging thanks to built-in social networking characteristics and improved security. The questionnaire was appropriately configured to describe the views, concerns, obstacles and ideas of the participants in order to identify the measures needed to further strengthen the actual condition of AI in Saudi Arabia.

Table II provides the analytical breakdown of the impacts of AI on social and economic domains on the Saudi community. Also, it illustrates the feedback obtained from

respondents regarding the social and economic criteria factors affecting the satisfaction of the AI Saudi community.

There are 14 items for the social domain classified as 6 items for acceptance, 2 items for security, 3 items for ability, 2 for approval, and 1 item for efficiency. Also, there are 2 items for the economic domain that examine the productivity criterion. The mean score for each item was used to describe the strength of each one. In the social domain, there are 4 items’ mean scores higher than 3 and 1 item for the economic domain. The obtained result demonstrates that some respondents highly agree about the impact of AI on security, approval, efficiency and productivity items.

Furthermore, 9 items’ mean scores are higher than 2 in the social domain and 1 item in the economic domain. Thus, most respondents mostly agree about the impact of AI on acceptance, security, ability and approval’ items. On the other hand, there is only 1 item’s mean score lower than 2 in the social domain in relation to being photographed without people’s knowledge. The result could be explained by the fact that the Saudi community is so far considered as conservative families. Fig. 2 and 3 provides the graphical interpretation of the information provided in Table II.

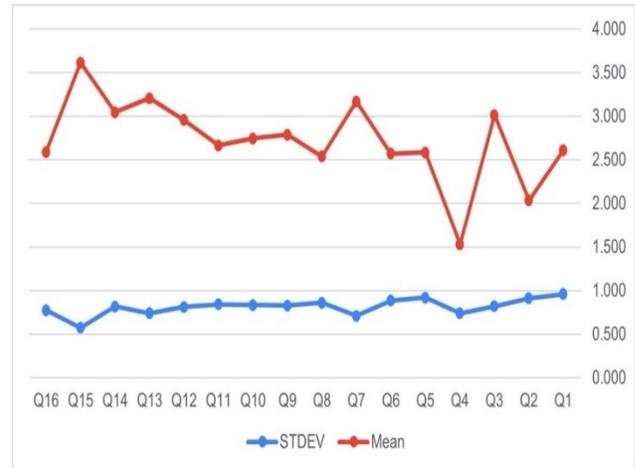


Fig. 2. Means and Standard Deviation for each Item.

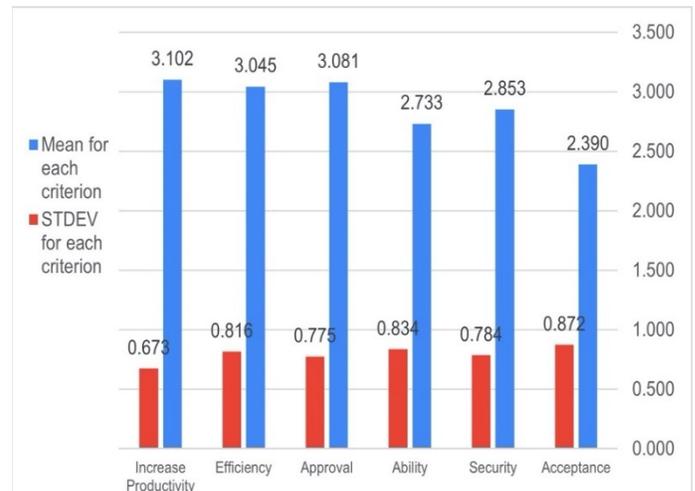


Fig. 3. Means and Standard Deviation for each Criterion.

TABLE II. PERCENTAGE DISTRIBUTION OF SOCIAL AND ECONOMIC DOMAINS FACTORS AFFECTING SATISFACTION OF AI SAUDI COMMUNITY

Domain	Criteria	Items	S.A%	A%	D%	S.D%	Mi	STDEVi
Social	Acceptance	1.Do you accept to be controlled by smart cameras while shopping, driving, walking, getting a public service (bank, ministry, school, university, hospital)?	19.1	37.97	27.83	15.09	2.611	0.960
		2.To you accept that your personal information will be saved including your face patterns, your actions, your movements, your car number?	8.49	17.69	42.45	31.37	2.033	0.911
		3.Do you think that using smart cameras could increase the efficiency of social services (like booking ticket using face patterns only in airports)?	28.77	49.06	16.75	5.42	3.012	0.82
		4.Do you accept to be photographed without your knowledge	3.07	5.66	32.55	58.73	1.531	0.739
		5.Do you accept being surveyed almost completely by government authorities?	15.75	41.75	28.07	14.62	2.583	0.92
		6.Do you accept boarding in a smart car and not driven by a human being?	15.09	38.68	34.43	11.79	2.571	0.885
	Security	1.Do you think that using smart cars in Saudi Arabia is much secure than traditional driving?	31.84	56.13	8.96	3.07	3.167	0.709
		2.Do you think that using smart surveillance cameras in Saudi Arabia could release a secure life better than your actual life?	14.62	234.43	41.04	9.91	2.538	0.86
	Ability	1.Do you think that the smart car is better than human being driver?	19.1	47.41	26.65	6.84	2.788	0.828
		2.Do you think that using smart cars will reduce the accident rate in Saudi Arabia?	20.52	38.21	36.56	4.72	2.745	0.833
		3.Do you think that smart cameras are able to work continuously (while raining, winding, a troubleshoot in electrical power)?	17.45	38.44	37.26	6.84	2.665	0.842
	Approval	1.Do you agree about using artificial intelligence generally in Saudi Arabia	26.42	47.41	21.7	4.48	2.958	0.812
		2.Do you think that the advantages of artificial intelligence are more than its disadvantages?	36.08	51.89	8.49	3.54	3.205	0.738
	Efficiency	1.Do you think that smart cameras will collect data efficiently?	31.6	45.28	19.1	4.01	3.045	0.816
	Total Distribution for Each Scale			20.55	39.29	27.27	12.89	
Economic	Increase Productivity	1.Do you think that using smart cars increases people's productivity?	64.62	33.02	0.94	1.18	3.615	0.572
		2.Do you think that using smart cameras in working areas (companies, universities, hospitals...) could increases employee's productivity?	11.56	42.22	39.86	6.37	2.59	0.775
	Total Distribution For Each Scale			38.09	37.62	20.4	3.77	

C. Analytical AI Impact for each Social Criterion

- Generally, the research found that the total respondents mostly agreed the AI social impact is 59.84% which is the sum of S.A (20.55%) and A (39.29%). This result will be detailed below among different social criteria.

Acceptance criteria: The research found that the Saudi community (the respondents) agreed most highly upon item 3 that using smart cameras could increase the efficiency of social services (like booking tickets using face patterns only in airports had an average of $m= 3.12$ ($STDEV= 0.82$) . The most commonly agreed upon acceptance of controlling by smart cameras while shopping, driving, walking, getting a public service (bank, ministry, school, university, hospital had an average of $m=2.611$ ($STDEV = 0.960$) and monitoring and tracking the personal information and behaviour without knowledge) had an average of $m= 2.39$ ($STDEV = 0.872$). Also, the most commonly agreed upon approval of completely surveyed by government authorities, showed an average of $m= 2.583$ ($STDEV = 0.92$). Onboarding in a smart car and not driven by a human being, had an average of 2.571 ($STDEV = 0.885$) and the agreement of saving the extracting personal features which include face patterns, actions, movements and car number, showed an average of $m= 2.033$ ($STDEV= 0.911$).

- Security criteria: The research found that the Saudi community agreed most highly upon item 1 with an average for using smart cars being much secure than traditional driving with an average of $m= 3.167$ ($STDEV=0.709$). They mostly agreed upon item 2 with an average $m= 2.538$ ($STDEV=0.86$) for using smart surveillance for cameras in Saudi Arabia to release a secure life better than your actual life.
- Ability criteria: The research found that the Saudi community mostly agreed upon item 1 with an average of $m= 2.788$ ($STDEV=0.828$) for the ability of the smart car to do better than a human being driver. They mostly agreed upon item 2 with an average of $m= 2.745$ ($STDEV= 0.833$) for the ability to use smart cars to reduce the accident rate in Saudi Arabia. Also, they mostly agreed upon item 3 with an average of $m= 2.665$ ($STDEV= 0.812$) for the ability of smart cameras to work continuously (while raining, winding, a troubleshoot in electrical power).
- Approval criteria: The research explored that the Saudi community agreed most highly upon item 2 with an average of $m= 3.205$ ($STDEV= 0.738$) for approving that the artificial intelligence has a great advantages on Saudi society more than its disadvantages and they mostly agreed upon item 2 with an average of $m= 2.958$ ($STDEV=0.812$) for utilising artificial intelligence generally in Saudi Arabia.
- Efficiency criteria: The research revealed that the Saudi community agreed most highly upon item 1 with an average of $m= 3.045$ ($STDEV= 0.816$) for the efficiency of using smart cameras for collecting data for approving that artificial intelligence has great

advantages on Saudi society more than its disadvantages. They mostly agreed upon item 2 with an average of $m= 2.958$ ($STDEV=0.812$) for utilising artificial intelligence generally in Saudi Arabia.

D. Analytical AI Impact for Economic Criterion

- Increasing Productivity: The research found that the Saudi community agreed most highly with an average of $m= 3.615$ ($STDEV=0.572$) for increasing people’s productivity by using smart cars. They mostly agreed upon item 2 with an average of $m= 2.59$ (0.775) for increasing employee’s productivity by using smart cameras in working areas (companies, universities, hospitals, etc.). Thus, the total respondents that agreed most highly the economic AI impact is 75.71% which is the sum of S.A (38.09%) and A (37.62%).

E. Overview of AI Impact on Domains

In Table III, generally, the research found that the Saudi community agreed most highly about the impact of AI upon the economic domain and specifically for the criteria “increasing people’s productivity” by using smart cars and implementing AI in working areas (companies, universities, hospitals...) with the total average $m=3.102$ ($STDEV=0.673$). Besides, the research found that the Saudi community mostly agreed with the impact of AI on the social domain, with the total average $M=2.82$ ($STDEV=0.816$). Specifically, for each criterion of the social domain, the respondents agreed most highly about the impact of AI upon the approval representing a total average of $m= 3.081$ ($STDEV= 0.775$) and upon the efficiency criteria of using smart cameras for collecting data with a total average $m= 3.045$ ($STDEV= 0.816$). The respondents most commonly agreed on the impact of AI upon the security that represents a total average $m= 2.853$ ($STDEV= 0.784$), upon ability representing the total average of $m= 2.733$ ($STDEV= 0.834$), and acceptance representing the total average $m= 2.39$ ($STDEV= 0.872$). Furthermore, Fig. 5 and 6 explore the analytical AI impacts on social and economic domains based on demographic profile. Fig. 5 shows firstly, that the “standard of living” highly influences the respondents' point of view upon the implication of AI on social and economic development as the least scales starting from mostly agree to agree were obtained for that demographic property. This result could be explained by the fact that most respondents have a “moderate standard of living” so that they have a fear of losing their jobs because of AI development.

TABLE III. AI IMPACT ON DOMAINS

Domain	Social					Economic
Feature s	Acceptan ce	Securit y	Abilit y	Approv al	Efficien cy	Productivi ty
Mc	2.39	2.853	2.733	3.081	3.045	3.102
STDEV c	0.872	0.784	0.834	0.775	0.816	0.673
Md	2.82					3.102
STDEV d	0.816					0.673

Secondly, the figure demonstrates also that the “age category” has a strong implication on the respondent’s approval criteria as most of the questionnaire participants being between 15 and 34 years old, they have a great curiosity to discover and explore all new AI technologies. As well as Fig. 6 demonstrating that the demographic property “education” has the highest scale (highly agreed) upon AI economic impact, whereas the same property has the least Likert scale (mostly agree) about AI social impact. The result could be explained by the fact that the higher the education level, the higher the respondents’ awareness about the influence of AI on society criteria including employment, security, privacy, and poverty.

Generally, from the research results shown in Fig. 4, 5 and 6, it is clear that almost all respondents are fully in support of the intervention of this technology in their life and their feedbacks reveal that AI has a greater impact on the economic domain than the social domain.

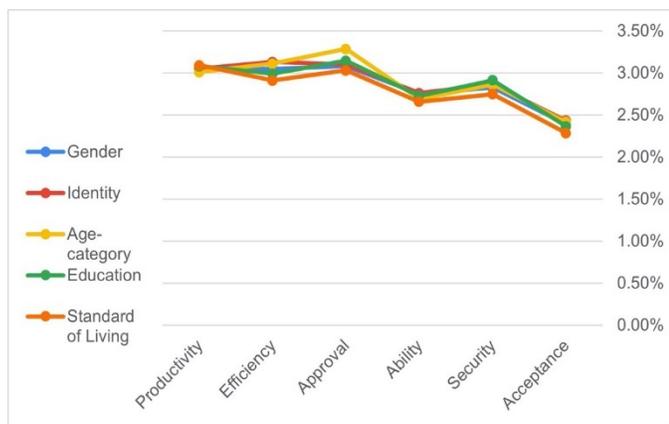


Fig. 4. AI Impacts on Social and Economic Domains.

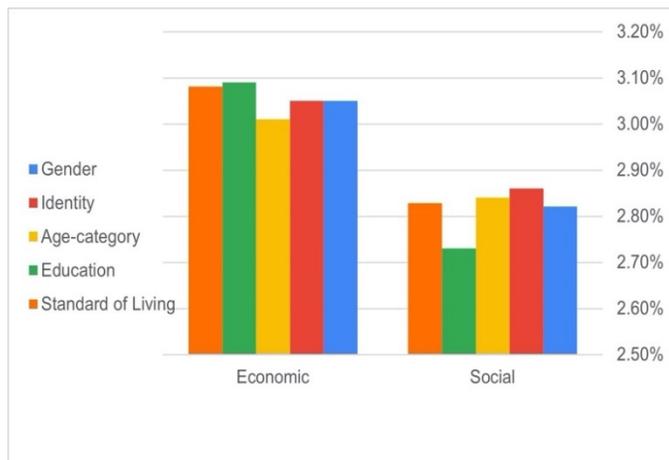


Fig. 5. Distribution of Relation’s Perspective between each Domains’ Criteria and Demo- Graphic’s Profile.

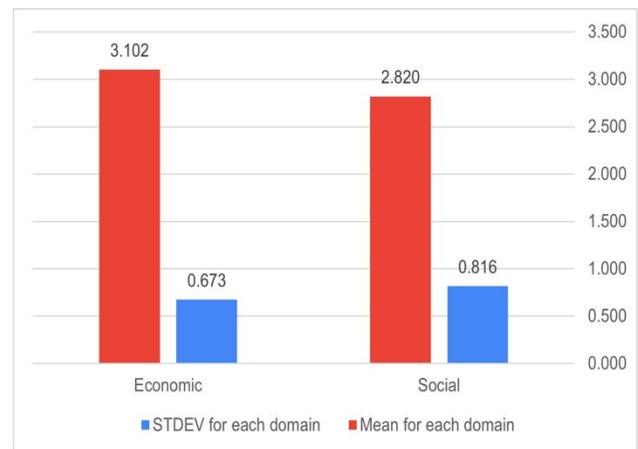


Fig. 6. Distribution of AI Impact on the Relationship between Social, Economic Domains and Demographic’s Profile.

V. CONCLUSION, FUTURE WORK AND LIMITATIONS

As Saudi Arabia continues to diversify away from oil dependence, the Kingdom has adopted a series of procedures considered to be adherent to AI technology. In this paper, the AI implications upon social and economic development have been presented. A special focus on the use of smart cars and smart cameras to monitor intelligently traffic, public services and national security has been studied. Understanding how the Saudi community is reacting during this economic pattern change is a significant challenge for social/economic experts and policymakers. The work outlined in this paper has resulted in evidence-based work, considerable worries of social AI implications, including ethical concerns, job loss and the changing workforce.

Responses also indicate that there is not enough general knowledge of AI, including what it can help accomplish in business and society and the consequences of not implementing it. Therefore, there needs to put in place adequate mechanisms to help companies and government negotiate the social and economic implications of AI. Several issues may be further investigated to highlight the concerns, behaviours and problems of the numerous participants’ demographic profile so that we could enhance the social and economic impact of AI in Saudi Arabia.

A. Limitations

The number of participants in this study is acceptable. Nevertheless, the participation rate could not be calculated for the online questionnaire as the prediction of adherent number for each mobile messaging application is impossible. Furthermore, the questionnaire takes only 2–3 minutes, and we tried to encourage participants through several ways of official communication. The male contribution was lower than expected due to the difficulty to contact them, as the college was originally devoted to females. Finally, there is no local data addressing matching rate or demographic profile distribution of AI impacts upon social and economic development in Saudi Arabia to compare with.

ACKNOWLEDGMENT

This paper is published under the socio-economic Implications of Artificial Intelligence project, as a part of Smart Campus implementation in Imam Abdulrahman bin Faisal University, Community College (CCD). The authors would like to special thanks to the Computer Department for supporting this work by giving the authorisation to share the survey through all students and faculty members.

REFERENCES

- [1] Clark, C. M., & Gevorkyan, A. V. (2020). Artificial Intelligence and Human Flourishing. *American Journal of Economics and Sociology*, 79(4), 1307-1344.
- [2] E. H. A. Elsaadani, M. Purdy, How Artificial Intelligence can Drive Diversification in the Middle East. Accenture, 2018.
- [3] M. Stephens, "Sheikha shamma bint sultan bin khalifa al nahyan and christopher m. schroeder (2019). 'perspective—future disruptive governments: Catching up with technological advancements and new horizons', future governments (actions and insights-middle east north africa, vol- ume 7)," 2019.
- [4] A. Beg, A. R. Qureshi, T. Sheltami, and A. Yasar, "Uav-enabled intelligent traffic policing and emergency response handling system for the smart city," *Personal and Ubiquitous Computing*, pp. 1–18, 2020.
- [5] D. Impedovo and G. Pirlo, "Artificial intelligence applications to smart city and smart enterprise," 2020.
- [6] H. Shokravi, H. Shokravi, N. Bakhary, M. Heidarrezai, S. S. Rahimian Kolor, and M. Petr, "A review on vehicle classification and potential use of smart vehicle-assisted techniques," *Sensors*, vol. 20, no. 11, p. 3274, 2020.
- [7] Huhtamo, E. (2020). The Self-Driving Car: A Media Machine for Posthumans?. *Artnodes*, (26), 1-14.
- [8] A. Agrawal, J. S. Gans, and A. Goldfarb, "Exploring the impact of artificial intelligence: Prediction versus judgment," *Information Economics and Policy*, vol. 47, pp. 1–6, 2019.
- [9] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mane', "Concrete problems in ai safety," arXiv preprint arXiv:1606.06565, 2016.
- [10] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature medicine*, vol. 25, no. 1, pp. 44–56, 2019.
- [11] K. Arai, "Vital sign and location/attitude monitoring with sensor networks for the proposed rescue system for disabled and elderly persons who need a help in evacuation from disaster areas," *sensors*, vol. 3, no. 1, 2014.
- [12] L. Amouri, M. Jallouli, C. Novales, G. Poisson, and N. Derbel, "Evaluation of user interface performing a dvz-fuzzy logic pilot for powered wheelchair.," *Journal of Automation, Mobile Robotics & Intelligent Systems*, vol. 13, no. 3, 2019.
- [13] L. Amouri, C. Novales, M. Jallouli, G. Poisson, and N. Derbel, "An effective dvz-fuzzy logic pilot for a mobile robot using generic architecture," *International Journal of Vehicle Autonomous Systems*, vol. 12, no. 3, pp. 201–220, 2014.
- [14] A. Hussein, M. Adda, M. Atieh, and W. Fahs, "Smart home design for disabled people based on neural networks," *Procedia Computer Science*, vol. 37, pp. 117–126, 2014.
- [15] A. Ghorbel, M. Jallouli, L. Amouri, and N. B. Amor, "A hw/sw implementation on fpga of absolute robot localization using webcam data," *Sensors, Circuits & Instrumentation Systems*, vol. 2, p. 75, 2017.
- [16] Saveliev, A., & Zhurenkov, D. (2020). Artificial intelligence and social responsibility: the case of the artificial intelligence strategies in the United States, Russia, and China. *Kybernetes*.
- [17] V. C. Müller and N. Bostrom, "Future progress in artificial intelligence: A survey of expert opinion," in *Fundamental issues of artificial intelligence*, pp. 555–572, Springer, 2016.
- [18] C. B. Frey and M. A. Osborne., "The future of employment: How susceptible are jobs to computerisation?," *Mimeo*, Oxford Martin School, 2013.
- [19] H. J. J. O. E. P. David, "Why are there still so many jobs? the history and future of workplace automation," *Journal of economic perspectives*, vol. 29, no. 3, pp. 3–30, 2015.
- [20] E. Ernst, R. Merola, and D. Samaan, "Economics of artificial intelligence: Implications for the future of work," *IZA Journal of Labour Policy*, vol. 9, no. 1, 2019.
- [21] G. C. Allen, *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*. Center for a New American Security Washington, DC, 2019.
- [22] Huhtamo, E. (2020). The Self-Driving Car: A Media Machine for Posthumans?. *Artnodes*, (26), 1-14.
- [23] O. Hassan, "Artificial intelligence, neom and saudi arabia's economic diversification from oil and gas," *The Political Quarterly*, vol. 91, no. 1, pp. 222–227, 2020.

Feistel Network Assisted Dynamic Keying based SPN Lightweight Encryption for IoT Security

Krishna Priya Gurumanapalli¹

Research Scholar, Department of Computer Science and
Technology
Sri Krishnadevaraya University
Anantapuramu, India

Nagendra Muthuluru²

Professor, Department of Computer Science and
Technology
Sri Krishnadevaraya University
Anantapuramu, India

Abstract—In the last few years Internet-of-Things (IoT) technology has emerged significantly to serve varied purposes including healthcare, surveillance and control, business communication, civic administration and even varied financial activities. Despite of such broadened applications, being distributed, wireless based systems, IoTs are often considered vulnerable towards intrusion or malicious attacks, where exploiting the benefits of loosely connected peers, the attackers intend to gain device access or data access un authentically. However, being resource constrained in nature while demanding time-efficient computation, the majority of the classical cryptosystems are either computationally exhaustive or limited to avoid attacks like Brute-Force, Smart Card Loss Attack, Impersonation, Linear and Differential attacks, etc. The assumptions hypothesizing that increasing key-size with higher encryption round can achieve augmented security often fail in IoT due to increased complexity, overhead and eventual resource exhaustion. Considering it as limitation, in this paper we proposed a state-of-art Generalized Feistel Network assisted Shannon-Conditioned and Dynamic Keying based SSPN (GFS-SSPN) Lightweight Encryption System for IoT Security. Unlike classical cryptosystems or even substitution and permutation (SPN) based methods, we designed Shannon-criteria bounded SPN model with Generalized Feistel Network (SPN-GFS) model that employs 64-bit dynamic key with five rounds of encryption to enable highly attack-resilient IoT security. The proposed model was designed in such manner that it could be suitable towards both data-level security as well as device level access-credential security to enable a “Fit-To-All” security solution for IoTs. Simulation results revealed that the proposed GFS-SSPN model exhibits very small encryption time with optimal NPCR and UACI. Additionally, correlation output too was found encouragingly fair, indicating higher attack-resilience.

Keywords—Internet-of-Things; dynamic programming; lightweight encryption; generalized Feistel network; substitution and permutation network

I. INTRODUCTION

The exponential rise in advanced software computing and low-cost hardware has broadened the horizon for industries to provide more-efficient and productive systems, serving healthcare, monitoring and surveillance, home-automation, smart-city planning and management, business communication, smart-factory, etc. Amongst the major such innovations the recently evolved IoT technology has gained wide-spread momentum to serve major aforesaid application

environment. With low-cost, decentralized communication and control ability, IoT has become one of the most sought-after technology to meet the major aspects of human activities and expectations [1]. Contemporarily, it not only enables optimal decision making, productivity but also facilitates luxurious life signifying accomplished human objectives in modern era. In general, the inception of IoT can be hypothesized to be with machine-to-machine (M2M) communication systems that later evolved as internet enabled query driven systems or services [1]. Architecturally, IoT systems comprise multiple distributed (autonomous) nodes and a sophisticated software computing enabled data-processing unit, where the first collects the real-world data, while the later executes query driven tasks or instructions to meet expected demands [1]. Interestingly, in early days of evolution, though no sophisticated communication platforms were defined; however, the contemporary IoT eco-systems operate over certain wireless channel with unavoidable network dynamism and uncertainty [2]. Undeniably, in the era of IoT-ecosystem(s), the interaction between human and machines is more frequent and increase with high-pace. This is because the machines getting smarter and capable to perform many human tasks on the basis of certain targeted instructions provided. This as a result has been broadening the horizon of IoT systems across industries as well as socio-scientific periphery [1][2]. Putting a glance around, one can find that IoT has been applied in varied critical purposes as well including healthcare, finance, manufacturing, research, surveillance etc.; however, being operated over certain unpredictable network condition its reliability and data-infra-procedural (DIP) safety always remains questionable. Ironically, with increase in high-pace IoT demands, the events of security breaches too have increased severely. The security breaches in IoT can be at the different level, i.e., device level, infrastructure level, network level (say, channel) and data level [3][4][6][38]. Such security breaches can cause detrimental impacts in terms of financial loss, data-loss, tack-manipulation and procedural destruction [2]. Consequently, such IoT-security breaches can have adverse impact on major aspects of human life and its decision systems [5].

Recalling the typical architecture of the IoT systems that encompasses Low-power Lossy (sensor) Networks (LLNs), gateways, transmission channel and data-warehouses or storage possessing undeniably security related vulnerabilities. Typically, the poorly interfaced nature of IoT devices and their inferior connection, especially operating online can undergo

security breaches and resulting losses [5-7]. On the other hand, attack probability gets increased manifolds over channel where different man-in-the middle attack (MITM), linear and differential attacks, impersonation, etc. are more common. Similarly, once compromising the security access level of IoT devices due to poorly configured setup or keys, an intruder can manipulate the device and its functioning, thus obstructing overall performance [7]. However, looking into depth it can be found the key reason behind such attack possibility is poor encryption systems with limited keying capacity and inferior cipher-space [8]. Such limitations are common in IoT systems where to meet computational efficiency demands; authors often apply lower-sized keys with smaller round of encryption. In fact, the threat of higher computation due to large key size and more complex confusion or cipher randomness limit efficacy of such systems [8]. It indicates the need of a robust (say, attack-resilient) and lightweight encryption system for IoTs [5][6]. An interesting fact that majority of the classical security systems employ different encryption modalities to perform device (access-level) security and data security; however, such systems under real-time (hardware) realization imposes significantly large computational overheads and resource (power and memory) exhaustion. Practically, the methods addressing IoT-network security or access-level security are different than the one used for data-security. Approaches pertaining to the data security predominantly consider attack-resilience, irrespective of the cost of computation (over large data size) and allied hardware realization [9]. On the contrary, the security measures designed towards access-control focus on retaining large key-size and encryption rounds to enable attack-resilience [10-12]. In sync with above inferences, to cope up with contemporary or even NextGen IoT ecosystems, designing a robust and "Fit-To-All (FTA) lightweight encryption system is inevitable [12]. Such security models are required to retain higher attack-resilience even at reduced computational cost, time, and hardware utilization.

Considering above the potential approaches available, despite of the different cryptosystems or crypto-algorithms like AES [13], RSA [14], DES [18], Huffman Coding, Homomorphic encryption etc. IoT demands more effective and more specifically lightweight solution [12][15-17][19]. Majority of these encryption methods focus on increasing key-size and corresponding encryption rounds to retain higher level of attack-resilience; however, fail in hardware realization due to significantly large computational overheads, memory and power exhaustion [21]. Similarly, a few researches suggested to use multi-factor authentication (MFA) by strategically amalgamating multiple cryptosystems to achieve IoT-security [1][11][13][14][17][19][32]. Such approaches often fail when delivering resource efficient and computationally efficient performance, especially with hardware realization [11][20]. Unlike classical cryptosystems, in the last few years lightweight encryption systems [12][13][16-18][33] have emerged as a potential solution for IoT security. The ability to retain higher level of confusion in cipher even at the significantly lower computation enables lightweight encryption system a viable solution towards IoT-security. Amongst the major approaches towards lightweight encryption based IoT-security, SPN-based algorithms are well-known [15][21]; however, retaining higher level of confusion even at reduced

encryption cycle and relatively lower key-size has remained a challenge [8][22]. Most of the existing SPN encryption models employs 64, 128 and even larger size keys which operate over tens of encryption rounds to achieve higher cipher-confusion [23-28]. Unfortunately, these all approaches turned inferior due to higher demands of gate-elements (GE), power consumption and computational time [21][33][37]. Moreover, such methods mainly focused on retaining higher level of confusion (in cipher), irrespective of the fact that doing arbitrarily it may impact bit error rate or correlation performance. This as a result can impact different IoT-enabled services such as healthcare sector, device access control, etc. which are highly sensitive to the single-bit and /or pixel changes. In such circumstance, developing a robust lightweight encryption system with conditional confusion (say, limit constrained confusion in SPN) with lower key-size and minimum encryption rounds can be of great significance [29][30]. Undeniably, SPN based lightweight encryption methods have gained wide-spread attention towards IoT security; however, have always been criticized for limited S-box generation [29][34]. Moreover, their suitability with single key assisted encryption [30] has also been under suspicious lens. It indicates the need of dynamic keying concept with improved S-box function so as to yield better cipher at the end. Furthermore, the efficacy of Feistel Network towards SPN realization [31] too has broadened the horizon for further improvement in lightweight encryption based IoT systems. The hypothesis that "the strategic implementation of SPN with improved FN architecture [31][37] can enable improved lightweight encryption system for IoT" can be considered as the key driving force behind this research.

In sync with above stated problem and allied scopes, in this research paper a state-of-art new and robust Generalized Feistel Network assisted Shannon-Conditioned and Dynamic Keying based SPN Lightweight Encryption System for IoT Security. Our proposed lightweight encryption model, GFS-SSPN intends to improve attack-resilience or cipher generation while retaining lower key size and even significantly small encryption rounds. Architecturally, GFS-SSPN model employed SPN with Generalized Feistel Network (GFN) to perform block-cipher based encryption for IoT device security. Here, being a block-cipher based encryption, our proposed GFS-SSPN model considered 64 bits key size with merely five encryption rounds to perform data encryption. The overall proposed model encompassed dynamic programming based "Dynamic Key Generation and Expansion System", 64-bit block-cipher encryption and decryption. The proposed GFS-SSPN was designed in such manner that it retained higher level of confusion even at the reduced key-size and encryption rounds that enable to resilient to the major IoT attacks like Smart Card Loss Attack, Brute-Force Attack, Impersonation Attack, Linear and Differential attack etc. Unlike classical approaches the use of Shannon criteria for SPN-GFS encryption enabled quality-preserving encryption thus helping GFS-SSPN to have sufficient correlation and possibly minimum bit-error rate at the decryption. It enables it to be used for broad IoT applications including access-control, sensitive data logging and actuation control, telemedicine etc. [38]. The performance assessment in terms of NPCR, UACI, correlation and execution time exhibited that the proposed

GFS-SSPN model exhibits optimal performance even with significantly small key-size and encryption round. Additionally, the use of GFS enabled reduction in decryption programming and hence can be visualized as reduced hardware utilization or allied power exhaustion.

This research paper is organized as follows: Problem formulation is provided in Section II, which is followed by the system model in Section III. The results and discussions are presented in Section IV and conclusion and future scope of this research are presented in Section V.

II. PROBLEM FORMULATION

Most of the classical security systems hypothesize that increasing key-size, encryption rounds and eventual cipher-randomness can help achieving optimal security; however, fail in addressing the resulting complexity and resource (i.e., power, memory, etc.) exhaustion. Additionally, these approaches undergo significantly large delay as well thus making it unsuitable towards realistic IoT-ecosystem. Undeniably, existing IoT-security systems employed cryptosystem(s) whether as standalone architecture or multi-factor authentication [35]. It eventually imposes computational overheads and complexity, thus making it inappropriate for IoT systems. To alleviate such issues, developing a lightweight encryption system has always been the dominant solution. However, retaining optimal key-size, number of rounds etc. has always remained challenge for industries, as any inappropriate design could cause loosely-coupled systems or channel inviting attacker to intrude it. Despite of the fact that the majority of the lightweight encryption methods developed so far are mainly focused towards multimedia data security in IoT systems [6]. On the contrary, a complete IoT ecosystem needs security system to ensure device-level security as well as data-level (within channel) security. To cope up with such demands, designing a "Fit-To-All" solution is must. Considering it as objective, in this paper a state-of-art new and robust Lightweight encryption-based block-cipher technique is proposed for IoT-systems. In GFS-SSPN the key goal is focused on retaining a lightweight encryption solution with smaller key-size, very small encryption and allied computational cost. Moreover, realizing the fact that IoT-systems encompass varied applications including access-control, data-logging, query-based data-retrieval, financial transactions, telemedicine information exchange etc., we focused on retaining optimal data quality along with uncompromising security. This as a result can enable a Fit-To-All solution for both access-control security as well as IoT-data security.

Literatures reveal that SPN based lightweight encryption is more hardware-friendly and computationally efficient towards IoT-security system; however, majority of the related work have applied large key-size with single key-based encryption. Additionally, such approaches have considered higher number of encryption-rounds to introduce encryption. Despite of the lightweight computation ability such approaches are often criticised for higher computational cost and delay that confines its suitability for IoT systems. It indicates the need of a solution employing lower key size with minimum encryption round of computation. Moreover, reduction in programming cost and

allied overhead can help reducing memory consumption or allied gate-element exhaustion. Towards data-quality centric IoT systems, maintaining optimal cipher quality with minimum possible error too is equally significant. Considering these facts as scope or motivation, in this research we designed a novel Generalized Feistel Network (GFN) assisted Shannon-Conditioned and Dynamic Keying based SPN (SSPN) lightweight encryption system for IoT security. As the name indicates, GFS-SSPN model encompasses Shannon-Conditioned SPN encryption system which is employed over GFS encryption structure for plaintext encryption. Architecturally, GFS-SSPN employs 64-bit key which is executed over merely five rounds of encryption to generate the cipher results. Unlike major classical SPN models where authors mainly focus on introducing randomness, without considering its detrimental impact on the cipher generated and eventual decryption results, our proposed SSPN model is designed in reference to the Shannon-condition. Here, Shannon condition states that for an arbitrarily selected input, if one flips the i th bit, then likelihood that the j -th output would be changed must be one half has been taken into consideration, which is also called Strict Avalanche Condition for SPN. This approach can enable sufficiently large confusion in cipher, while maintaining higher association between input plain text and the cipher generated and hence would be advantageous towards error-free decryption. The proposed SSPN model has been applied with GFS architecture that enables its implementation as block-cipher model by partitioning input data into multiple chunks. This method not only helps in enhancing computational efficiency (i.e., time) but also reduces the separate program demands for decryption. It can be vital towards resource efficient and delay-resilient encryption for IoT. Recalling the overall architecture, SSPN model can enable higher confusion with optimal randomness and resource as well as delay efficiency while preserving data-quality. Here, SSPN helps achieving optimal S-box and P-box generation, while GFS would enable dynamic key management of keying. In the proposed model, SPN block cipher has been applied in iterative and alternating rounds of substitution and permutation (say, transposition) while ensuring that it fulfils the demands of Shannon's Confusion and Diffusion characteristics to ensure higher attack-resilience. To achieve it, a novel dynamic key management has been developed to assure that the cipher has been processed in pseudo random manner. To introduce higher level of confusion and security-structure the proposed key-expansion block is implemented over five rounds, where in each round it intends to meet above stated Shannon's Confusion and Diffusion (SCD) conditions. Cumulatively, these approaches can be robust enough to retrieve attack-resilient encryption even at the reduced cost and time. Here, to introduce a non-linear layer for better lightweight encryption, the proposed model applied S-Box. The proposed S-box has been applied in such way that it enables optimal diffusion over each round and thus making it more imperceptible and hence higher attack-resilient. Being a block-cipher concept, it splits 64-bit data into four chunks and performs five rounds of key generation and encryption that introduces sufficiently large randomness in cipher to avoid any attacks including SCLA, Brute-force, impersonation, linear and differential attacks etc.

III. SYSTEM MODEL

As discussed in the previous section, our proposed lightweight encryption-based block-cipher model amalgamated SSPN and GFS optimistically designed to cope-up with high parallelism so as to achieve better time-efficiency. In the proposed model, complete input block is split into four equally partitioned 16-bit blocks. Noticeably, unlike classical SPN based approaches, we improved the overall structure by introducing numerous enhancements such as Shannon-Conditioned SPN, GFN, Dynamic Keying, etc. To enable our proposed GFS-SSPN model for swift computing IoT security system, unlike classical Feistel Network (CFS) which partitions the input into two-equal blocks, we applied GFS which split it into multiple equal blocks for further encryption. This mechanism at first helped SSPN to achieve time as well as computationally efficient parallelized encryption. Moreover, the use of GFS helps reducing additional programming cost that makes it hardware and power efficient when realized in real-world. Before discussing the proposed lightweight encryption model, a snippet of the proposed SPN model, in sync with Shannon criteria is given as follows:

A. Shannon's Confusion and Diffusion Theory

Being an SPN based encryption the proposed GFS-SSPN model applied two consecutive methods, confusion and diffusion for encryption. A snippet of these mechanisms is given as follows:

1) *Confusion*: Typically, in encryption system domain, confusion mechanism states that "each binary bit of the cipher must rely on varied other fractions of the key, representing obscured connection between the two (i.e., cipher and key). Functionally, it hides the associations in between the plaintext and the key applied, and consequently avoids any possible security breaches due to key-loss, such as SCLA or linear and differential attacks. Confusion makes it difficult for attacker to extract key and hence can alleviate any unauthorized retrieval of the original data from cipher. Functionally, for an efficient encryption model with single bit change or manipulation entire cipher bits must be changed. This approach enhances the level of ambiguity of the cipher and hence helps in retaining higher attack-resilience.

2) *Diffusion*: Unlike confusion, diffusion states that in case a single bit is manipulated in input plaintext, then nearly half of the bits in the cipher might be statistically changed. Similarly, if a single bit is manipulated in the cipher then the half of the plaintext bit would change. Since, a single bit may have only two-states, approximately half of the bit can have their state manipulated. In our proposed GFS-SSPN encryption scheme, the predominant concept behind diffusion is to hide the relationship between the input plaintext and the corresponding cipher which make it difficult for attacker(s) to get access of the plaintext. Diffusion is further accomplished because of the increased level of randomness in the plaintext over the different rows and columns. Functionally, it is accomplished by performing transposition. A snippet of the Shannon

Condition of SPN, being applied in the proposed GFS-SSPN lightweight encryption model is given as follows:

Typically, Shannon condition for SPN encryption defines confusion as the process for transforming the associations or the relationship between the key and the cipher as complex as possible. It enables depletion or reduction in the native statistical structure of the input plaintext over cipher. In our proposed SSPN model, such complexity is introduced by perform repetitive substitutions and permutations function, where in substitution, a part of bits in each block is replaced or substituted with other parts following the Shannon criteria of substitution. Permutation on the other hand indicates the mechanism of changing the bit's order as per certain mathematical approach, also called transformation rules. In GFS-SSPN model to achieve high non-linearity, the input bits are distributed across the structure of the cipher text which makes it difficult to detect either key or the plaintext data. In sync with the Shannon condition for SPN, for any randomly selected input, if one changes the i th bit, probability that the j -th output would be changed remains the half. This condition is also known as the Strict Avalanche Condition. It states that it is important to assure that flipping of a definite set of bits must change each output bit in cipher with the probability of minimum one-half. A vital purpose of the use of confusion was to inculcate such randomness that an attacker to identify the key, despite the fact that the one has the significantly large plaintext-cipher pairs generated with the same key. Therefore, each bit of the cipher depends on the entire key in the different ways on the varied bits of the key. In other words, manipulating a single bit can change the entire cipher text. Thus, implementing the Shannon conditioned SPN (say, SSPN) we developed a state of art new Dynamic Keying and Expansion (DKE) mechanism over with GFS over five consecutive rounds. Here, we designed DKE in such way that it ensures the Shannon condition where the cipher gets manipulated in pseudo random manner. The proposed DKE assisted encryption with GFS-SSPN model was applied over five consecutive rounds so as to induce higher confusion and hence attack-resilience, where in each round of computation it follows the Shannon's criteria of confusion and diffusion. The detailed discussion of the overall system implementation is given in the sub-sequent sections.

B. System Implementation

For system implementation, to maintain lower computational cost and allied (possible) hardware resource exhaustion, we designed GFS-SSPN as a 64-bit block cipher model. In this approach we considered 64-bit key to perform encryption of each block encompassing 64-bit input plaintext. Additionally, to perform GFS-SSPN encryption we applied only five round of encryption. In this process, each encryption-round is performed over DKE function which helps in generating confusion and diffusion matrix over different rounds of encryption. Though, maintaining higher encryption round can introduce higher randomness; however, at the cost of increased computational overhead, delay and resource exhaustion. Therefore, to design the proposed encryption model suitable for IoT-applications, we maintained only five round of encryption, while to introduce higher cipher randomness and minimum correlation we introduced a state-of-

art new Dynamic Keying and (key) Expansion concept in which each split of input 16-bit was processed for GFS-SSPN over five rounds, where each round gave rise to a new set of keys for encryption. Thus, performing keying expansion over each round of computation finally we obtained a concatenated 64-bit key to be used for decryption. Architecturally, in GFS-SSPN, at first the input of 64-bit was split into four equally chunks, which was possible by the use of GFS architecture. In other words, by applying the proposed GFS architecture as shown in Fig. 1, we applied SSPN over each chunk of 16-bits. To further enhance randomness in cipher, we split each 16-bit data into four distinct sub-parts each of 4-bits size. To further introduce higher confusion and diffusion over consecutive rounds of encryption, the proposed GFS model applied logical XOR function over each round of confusion and diffusion to result cipher results per block. To achieve the overall implementation our proposed model applies three key functions.

- 1) Multi-round Dynamic Keying, Expansion and Update.
- 2) Block-Cipher Encryption and Cipher Generation, and
- 3) Block-Cipher Decryption.

The detailed discussion of each functional component is given in the subsequent sections.

1) *Multi-round dynamic keying, expansion and update:* In sync with the real-world IoT systems where each node can be assumed to act like a key generator as well as decoder, it becomes vital to minimize cost caused by key generation and update. To enable lightweight encryption, in GFS-SSPN model different mathematical models by using varied logical functions like XOR and XNOR has been developed. Here, our proposed DKE model serves dual purpose; first it acts as a key generation and expansion (KGE) unit while secondly it helps concatenating cipher iteratively. To achieve key generation and expansion, GFS-SSPN applied GFS assisted encryption. In GFS-SSPN, GFS is applied over five encryption rounds, where each round employs distinct key to perform substitution and permutation tasks. Realizing the fact that the proposed model intends to apply minimum (here, only five) round of encryption which is undeniably lower than the classical AES, RSA, ECC, Diffie-Hellman, etc. kind of encryption systems or even many existing SPN based encryption model, our proposed model intended to introduce a concept of dynamic keying. Unlike classical SPN methods or other cryptosystems that apply merely single key to perform encryption, GFS-SSPN to introduce higher randomness or confusion in cipher, we applied dynamic keying concept. In the GFS-SSPN, we estimated a set of keys over each round and thus, we retrieved a final concatenated key of 64-bits. Noticeably, to retain higher attack-resilience a lightweight encryption model requires maintaining sufficiently large key size k_t which can prohibit any intruder to perform 2^{k_t-1} encryption so as to gain key

information for data access or retrieval. Our proposed model applied 64-bit key to perform bit-permutation instruction-based encryption for the 64-bit block sized input data. We applied 64-bit's cipher key k_c as the input of the proposed DKE function to perform a SSPN confusion and diffusion task, as per Fig. 2. In this manner, it provides five distinct keys one for each round of encryption.

In DKE (Fig. 1) we implement the mathematical approach suggested by Barreto et al. [68]. The Khazad cipher model as proposed in [68] applied Broad-Trial-Mechanism (BTM) to perform multiple linear and non-linear transformation to perform encryption. On the contrary, we applied GFS architecture which enabled a definite relationship and inter-dependency between the input bits and the output cipher in a predefined complex manner. As depicted in Fig. 1 the 64-bit input k_c is equally-split into four equal-blocks, each of 16-bits. Moreover, to perform SSPM, we split each 16-bit block into 4-bit chunks, which were later processed for substitution and permutation as depicted in Fig. 2. Thus, performing DKE over each 16-bit input, our proposed model obtained a matrix of 4-bits each. Noticeably, each block of 16-bits generated four 4-bits chunks to be processed for further GFS enabled SSPN. The detail of the proposed SSPN model is given in the subsequent section.

a) Shannon Constrained SPN (SSPN)

As depicted in Fig. 1, once estimating the 16-bits data post DKE, we performed the initial substitution of the chunks of k_c (1).

$$Km_{i \in \{1,2,3,4\}f} = \prod_{j=1}^4 Kn_{4(j-1)+i} \quad (1)$$

The above derived model (1), enabled estimation of a 16-bits output for each DKE block, applied over all four 16-bit blocks. Once estimating the 4×4 matrix from each 16-bit input (i.e., post DKE), we performed circular shifting which helped us to generate four different keys.

Now, estimating all four keys for $Km_{i \in \{1,2,3,4\}f}$ as depicted in Fig. 1, we concatenated these keys and retrieved a set of keys Kc_{if} for round of computation using (2).

$$Kc_{if} = f(Km_{if}) \quad (2)$$

To further introduce higher confusion and diffusion over the consecutive SSPN layers, we applied a state-of-art new Transformation coder containing two different functions, say, "Linear (LF)-Non-Linear Function (NF)", as depicted in Fig. 2.

To enable a computationally efficient SSPN implementation, we applied a predefined transformation coder, encompassing two distinct functions, LF and NLF, as given in Table I. Though, in GFS-SSPN k, we applied LF and NLF transformation coder as predefined value; however, different other transformational values can also be assigned for the different test cases or custom encryption demands.

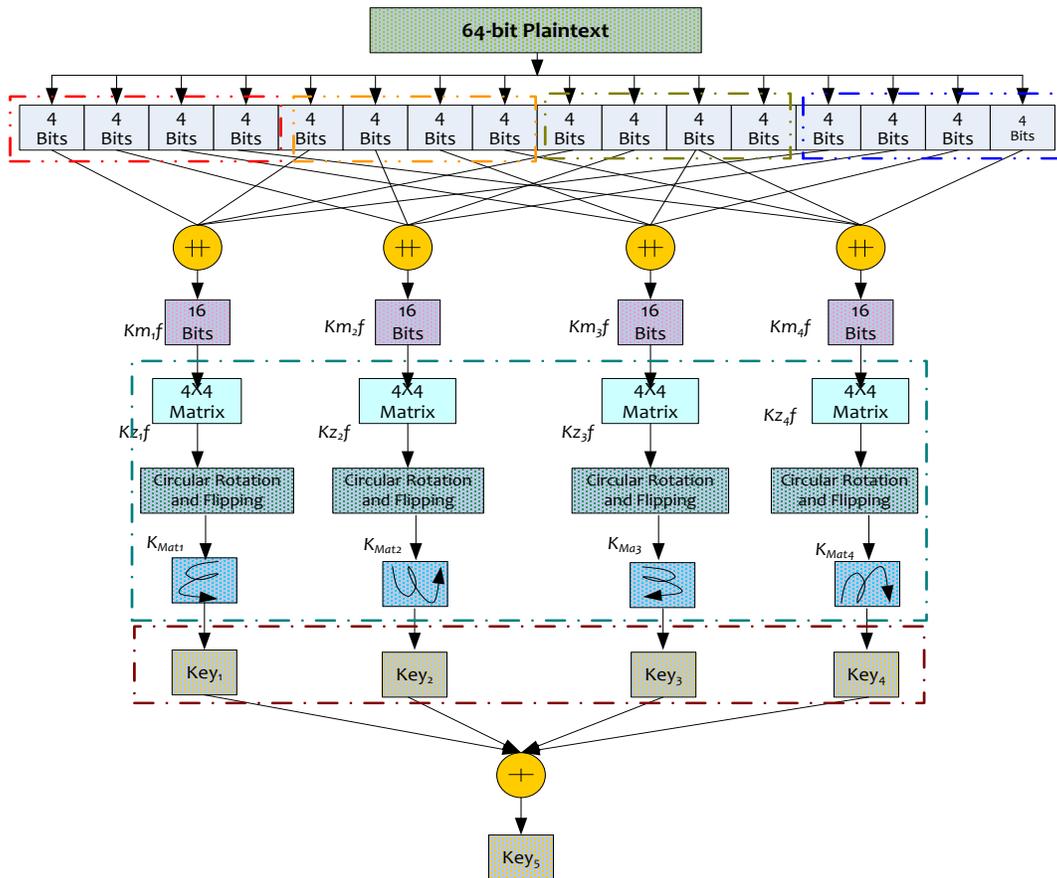


Fig. 1. Proposed Dynamic Key Generation and Expansion.

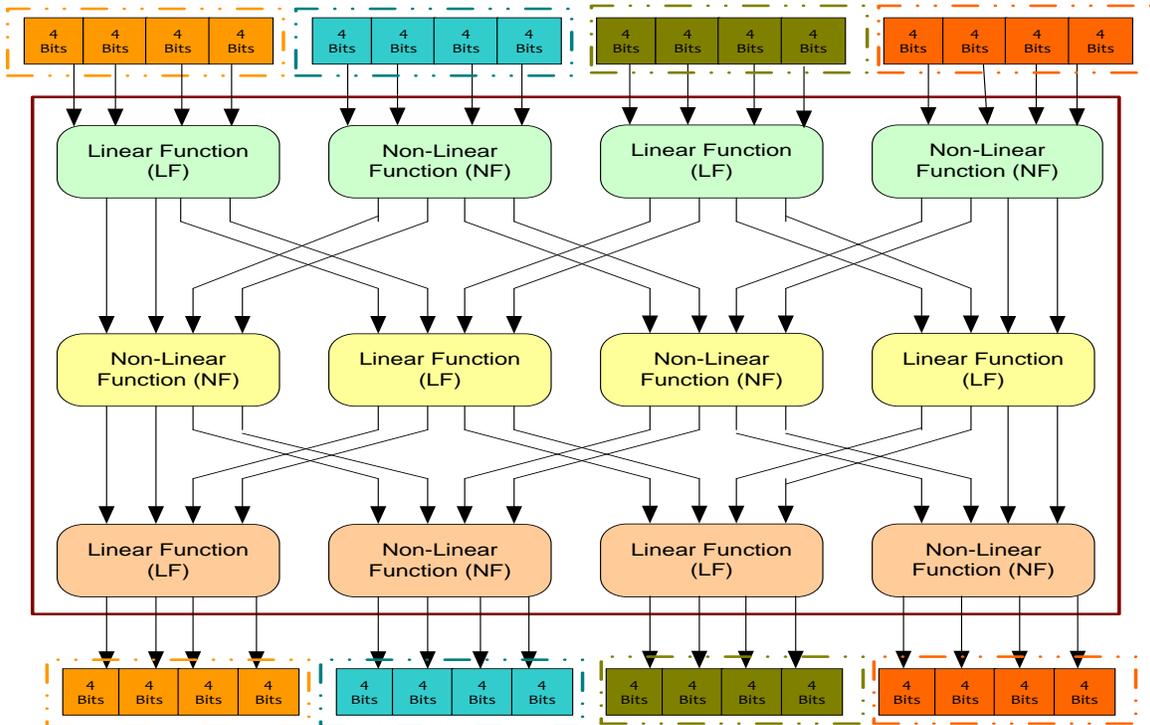


Fig. 2. SSPN Implementation.

TABLE I. TRANSFORMATION CODER

$k_{n,i \in \{1,2,3,4\}}$	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$LFP(k_{n,i \in \{1,2\}})$	3	F	F	0	4	1	E	B	A	F	7	8	6	8	1	1
$NLFP(k_{n,i \in \{1,2,3,4\}})$	9	D	4	4	F	2	3	E	D	2	6	E	7	C	2	7

Thus, with the different inputs, we applied transformation coder (Table I) values. This as a result enabled distinct keys per 16-bit blocks. Estimating the outputs from the DKE model, we obtained 4×4 matrix which was later processed for the ‘‘circular-Shifting’’. This mechanism re sampled the input 4×4 matrix, per 16-bit of input and eventually estimates four distinct keys, as per the equations (3-6). The matrix generated for each 16-bit block is given in equations (3-6).

$$K_{Mat1} = \begin{bmatrix} Kc_1f_1 & Kc_1f_2 & Kc_1f_3 & Kc_1f_4 \\ Kc_1f_5 & Kc_1f_6 & Kc_1f_7 & Kc_1f_8 \\ Kc_1f_9 & Kc_1f_{10} & Kc_1f_{11} & Kc_1f_{12} \\ Kc_1f_{13} & Kc_1f_{14} & Kc_1f_{15} & Kc_1f_{16} \end{bmatrix} \quad (3)$$

$$K_{Mat2} = \begin{bmatrix} Kc_2f_1 & Kc_2f_2 & Kc_2f_3 & Kc_2f_4 \\ Kc_2f_5 & Kc_2f_6 & Kc_2f_7 & Kc_2f_8 \\ Kc_2f_9 & Kc_2f_{10} & Kc_2f_{11} & Kc_2f_{12} \\ Kc_2f_{13} & Kc_2f_{14} & Kc_2f_{15} & Kc_2f_{16} \end{bmatrix} \quad (4)$$

$$K_{Mat3} = \begin{bmatrix} Kc_3f_1 & Kc_3f_2 & Kc_3f_3 & Kc_3f_4 \\ Kc_3f_5 & Kc_3f_6 & Kc_3f_7 & Kc_3f_8 \\ Kc_3f_9 & Kc_3f_{10} & Kc_3f_{11} & Kc_3f_{12} \\ Kc_3f_{13} & Kc_3f_{14} & Kc_3f_{15} & Kc_3f_{16} \end{bmatrix} \quad (5)$$

$$K_{Mat4} = \begin{bmatrix} Kc_4f_1 & Kc_4f_2 & Kc_4f_3 & Kc_4f_4 \\ Kc_4f_5 & Kc_4f_6 & Kc_4f_7 & Kc_4f_8 \\ Kc_4f_9 & Kc_4f_{10} & Kc_4f_{11} & Kc_4f_{12} \\ Kc_4f_{13} & Kc_4f_{14} & Kc_4f_{15} & Kc_4f_{16} \end{bmatrix} \quad (6)$$

Now, once retrieving the values of the key-matrix ($K_{Mat1}, K_{Mat2}, K_{Mat3}$ and K_{Mat4}), to obtain the keys for each 16-bit block, we transformed these metrics into four distinct arrays of 16 bits as derived in (7-10). These 16-bits arrays provided the encryption key per round. Noticeably, in (7-10) the operator $\#$ represents the concatenation function.

$$Key_1 = a_4 \# a_3 \# a_2 \# a_1 \# a_5 \# a_6 \# a_7 \# a_8 \# a_{12} \# a_{11} \# a_{10} \# a_9 \# a_{13} \# a_{14} \# a_{15} \# a_{16} \quad (7)$$

$$Key_2 = b_1 \# b_5 \# b_9 \# b_{13} \# b_{14} \# b_{10} \# b_6 \# b_2 \# b_3 \# b_7 \# b_{11} \# b_{15} \# b_{16} \# b_{12} \# b_8 \# b_4 \quad (8)$$

$$Key_3 = c_1 \# c_2 \# c_3 \# c_{10} \# c_{11} \# c_{12} \# c_{16} \# c_{15} \# c_{14} \# c_{13} \quad (9)$$

$$Key_4 = d_{13} \# d_9 \# d_5 \# d_{11} \# d_7 \# d_3 \# d_4 \# d_8 \# d_{12} \# d_{16} \quad (10)$$

Thus, estimating the four distinct keys; Key_1, Key_2, Key_3 and Key_4 , we performed XOR logical function as per (11) to achieve the final encryption key.

$$Key_{Fused} = Key_1 \oplus Key_2 \oplus Key_3 \oplus Key_4 \quad (11)$$

2) *Lock-Cipher encryption and cipher generation*: Once estimating the complete 64-bit of encryption key (11), we performed plaintext encryption. To perform encryption, we applied the schematic given in Fig. 3. As depicted in Fig. 3, the input 64-bit plaintext was equally split into four 16-bit blocks (i.e., $Px_{0-15}, Px_{16-31}, Px_{32-47}$ and Px_{48-63}). Subsequently, executing bit-permutation instruction over each 16-bit block (each round), GFS-SSPN exhibited bit-wise swapping so to reduce traceability. To achieve it, we performed bit’s order alteration, where the sub-blocks of 16-bit block were changed. Subsequently, we performed bitwise XNOR logical operation between the round key Key_i and the input plaintext Px_{0-15} . This process was repeated four rounds in K_i to Px_{48-63} and eventually the corresponding outputs Ro_{11} and Ro_{14} were obtained. Now, estimating the output from XNOR logical operator (Fig. 3) we fed the results to the DKE component which generated two distinct outputs Ef_{l1} and Ef_{r1} . Now, once estimating the values of Ef_{l1} and Ef_{r1} , we took Ef_{l1} and Px_{32-47} , and performed bitwise-XOR to estimate Ro_{12} . Similarly, the bitwise-XOR between Ef_{r1} and Px_{16-31} gave rise to Ro_{13} . We applied equation (12) to estimate encrypted cipher outputs.

$$Ro_{i,j} = \begin{cases} Px_{i,j} \odot K_i ; & j = 1 \text{ and } 4 \\ Px_{i,j+1} \oplus Ef_{li} ; & j = 2 \\ Px_{i,j-1} \oplus Ef_{ri} ; & j = 3 \end{cases} \quad (12)$$

Thus, the above discussed transformations were performed in such way that for each consecutive round, Ro_{11} turned out to be Px_{16-31} , while Ro_{12} became Px_{0-15} , Ro_{13} as Px_{48-63} . Similarly, Ro_{14} became Px_{32-47} . This process was performed for all rounds by using (12). The outputs of the final round were concatenated as per (13) to generate the cipher output.

$$CT = R_{51} \# R_{52} \# R_{53} \# R_{54} \quad (13)$$

3) *Block-Cipher decryption*: As already stated, unlike classical SPN based encryption or other cryptosystems, our proposed GFS-SSPN lightweight encryption model applied GFS assisted SSPN, and therefore, it avoids any additional programme for decryption. The use of GFS enabled decryption in the same way as performed towards encryption using the different logical functions, as discussed above. This characteristic enabled significantly lower computational overheads and resource exhaustion. Thus, implementing the above discussed methods, we performed block-cipher encryption for the different set of inputs characterizing a small size authentication credential or a large plaintext data to be communicated over the IoT-centric LLNs channels or even to be stored on local memory [70]. The detailed discussion of the overall simulation results and allied inferences is given in the subsequent sections.

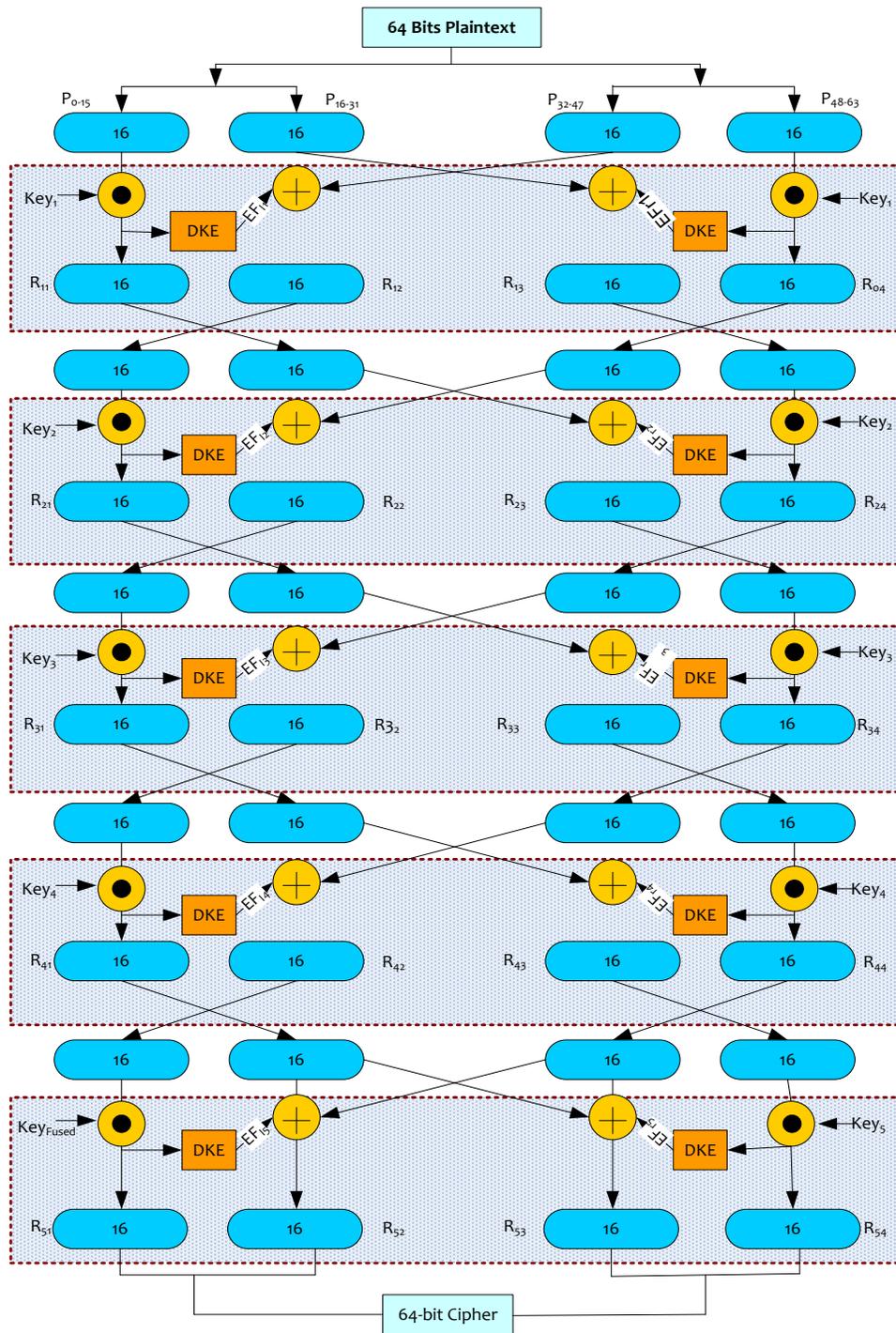


Fig. 3. Proposed Block-Cipher Encryption.

IV. RESULTS AND DISCUSSIONS

As stated, and discussed in details, the predominant emphasis of this research was to contribute a state-of-art new and robust lightweight encryption system or the block-cipher technique for IoT security. Zeroing down the belief that whether a device level or the data level security in IoT, maintaining seamless and attack-resilient cipher is must, the proposed work focused on introducing sufficiently large and optimal randomness (say, confusion) in cipher. On the other

hand, to retain computational efficacy and optimality, we considered block cipher method with 64-bit key and merely five round of encryption. In sync with hypotheses that maintaining or applying higher encryption rounds can enable higher randomness in cipher, unlike classical methods, we applied different state-of-art logical functions implemented over GFS assisted SSPN to perform encryption. In other words, we strategically implemented GFS with Shannon-Conditioned SPN to perform encryption using 64-bit key and only five

round of encryption. Noticeably, the key purpose of using GFS was to assist block-cipher generation, while SSPN was targeted to accomplish encryption with sufficiently large confusion or randomness. Moreover, the use of GFS enabled reduction in programming cost as it avoided any separate decryption program like other classical SPN methods or cryptosystem algorithms. Noticeably, unlike conventional SPN methods, the use of Shannon-conditioned SPN not only intended to achieve optimal confusion and diffusion, but also intends to support optimal decryption at the receiver without bit-loss. This is the matter of fact that not much significant works are done towards lightweight encryption based IoT security, especially by using SPN and/or Feistel network. A few efforts employing technologies have mainly focused on multimedia data security [6]; however, the use of higher encryption key and rounds make them computationally exhaustive. On the other hand, a few methods, targeting hardware efficient lightweight encryption too are found limited and exhaustive due to bit-by-bit encryption and supplementary tool/programme-based randomness insertion (in cipher). In our research we tried to alleviate such limitations and designed GFS-SSPN in such manner that it could retain higher confusion even with minimum encryption round and lower key-size. The use of dynamic keying concept over the multiple rounds helped accomplishing an untraceable keying strategy, where SSPN helped in maintaining sufficiently large randomness as well as associations with the input plaintext. This as a result intended to achieve high attack-resilience.

Repeating the structural details, our proposed GFS-SSPN model encompassed three key components including Dynamic Key Generation, Expansion and Update (DKE), GFS-SSPN encryption and GFS-SSPN decryption. Being a bloc-cipher method, the proposed model at first split 64-bit input into four equal blocks of 16-bit each. Subsequently, applying an initial key value, with input 16-bit plaintext we perform SSPN that eventually performs logical operations and generated each round output. The outputs obtained for each 16-bit blocks were processed for the different logical operations and flipping, circular rotation etc., that eventually generated 64-bit keys (16-bit key per block, which were concatenated together to result a final 64-bit key for encryption) over five round of encryption. Thus, with the final generated key, we performed GFS-SSPN encryption followed by decryption. The overall proposed GFS-SSPN model was developed using C-programming language which was simulated over *DEV C++* compiler. We simulated the proposed lightweight encryption system onto a central processing unit (CPU) with 8 GB RAM, and Intel-i5 processor. To assess performance of the proposed GFS-SSPN model, we examined its efficacy in terms of encryption time, correlation, NPCR and UACI. To be noted, as already stated IoT-ecosystem encompasses the different activities including data logging, device access, live streaming, etc., and hence for both device level security as well as data level or network level, ensuring plaintext (say, data) security is must. Retaining minimum encryption time, negligible or very small correlation between the cipher and key is must. Moreover, there are many IoT applications such as data-streaming, access-credentials, telemedicine in healthcare etc., where maintaining higher NPCR and suitable UACI is must. Considering this fact, we examined the proposed of the proposed GFS-SSPN model in

terms of encryption time, correlation, NPCR and UACI. The performance assessment has been performed under two broad umbrellas; first quantitative assessment and second, the qualitative assessment. Here, quantitative assessment discusses empirical simulation outputs and allied inference, while qualitative assessment discusses different attack-resilience. The detailed discussion is given as follows:

A. Quantitative Assessment

This section primarily discusses some of the key empirical analysis outcomes in terms of correlation, encryption time, etc.

1) *Correlation*: Realizing the fact that to ensure high attack-resilience and imperceptibility in cipher, it is must to retain minimum correlation between cipher generated, original text and the encryption key. There are many IoT-attack conditions such as linear and differential attack analysis, impersonation etc., where the attacker often exploits cipher details to retrieve the key information and allied original data. To avoid such problems, retaining minimum or negligible correlation is must. On the other hand, correlation being the statistical dependency between the two different variables requires to be maintained as minimal as possible so as to alleviate any attack conditions, as stated above. To assess performance of the proposed GFS-SSPN model, we obtained correlation coefficient value γ between the cipher text and the original text using (1). Noticeably, for any encryption-based security system maintaining γ near zero is considered as an ideal condition.

$$\gamma_{x,y} = \frac{cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} \quad (14)$$

In (14), the parameter $cov(x,y)$ signifies the covariance, while the variance for x and y are given as $D(x)$ and $D(y)$, respectively. Typically, the distribution of the variance of any single dimension random variable can be estimated as per (15).

$$D(x) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2 \quad (15)$$

In (15), $D(x)$ states the variance for x . Now, to calculate covariance in between x and y , we used (16) as given below.

$$cov(x,y) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))(y_i - E(y)) \quad (16)$$

In (16), the parameters $E(x)$ and $E(y)$ signifies the targeted values for x and y , respectively. Here, the expectation values for x is obtained as per (17).

$$E(x) = \frac{1}{N} \sum_{i=1}^N x_i \quad (17)$$

In (18), N presents the total number of bits in the data, while x signifies the N -dimensional vector. Noticeably, the component x_i presents the i -th value of the original plaintext. To assess the performance, we simulated proposed model with the input plaintexts of different sizes (here, we call sample(s)). The results obtained towards the correlation outputs between input sample and resulting cipher is given in Table II.

TABLE II. CORRELATION ANALYSIS

Input Samples	Original data	Cipher data
1	0.9919	0.0076
2	0.9971	0.0072
3	0.9829	0.0036
4	0.9801	0.0060
5	1.0000	0.0071
6	0.9672	0.0129
7	0.9721	0.0131
8	0.9408	0.0083
Average Correlation	0.9790	0.0082

The average correlation obtained over a total of 8 different samples encryption reveals that the proposed GFS-SPNN model accomplishes the minimum correlation coefficient of 0.0082. This is significantly small that it can enable seamless encryption for attack-avoidance. In above depicted results (Table III), we considered the different samples of varied sizes for encryption. Table II presents the correlation results between the original test and the cipher data obtained post GFS-SSPN. Observing the results, it can easily be inferred that post-encryption the correlation value significantly decreases. It indicates minimum correlation signifying high attack-resilience.

2) *Number of Changing Pixel Rates (NPCR)*: In addition to the correlation performance, we examined the efficacy of the proposed lightweight encryption system in terms of other statistical performance variables such as the number of pixels changing rate (NPCR) and unified averaged changed intensity (UACI). Typically, those encryption algorithms or methods which hypothesize that merely including randomness in cipher can enable attack-resilience often under a detrimental effect called avalanche effort or plaintext sensitivity [39-41]. Though, the use of Shannon criteria of SPN alleviated such problem, however, maintaining higher changes or sensitivity over encryption is must. In this relation, we measured NPCR performance signifying the total number of characters which are different in cipher than the equivalent plain-text (say, w_1 and w_2). Mathematically, NPCR was estimated as per (18), where its value is often presented in the form of percentile (%).

$$NPCR = \frac{\sum_{i=1}^l W(i)}{l} \times 100 \quad (18)$$

In (18), the variable l states the length of the text to be encrypted, and

$$W(i) = \begin{cases} 0, & \text{if } E_1(i) = E_2(i) \\ 1, & \text{if } E_1(i) \neq E_2(i) \end{cases} \quad (19)$$

In majority of the lightweight encryption models, the ideal NPCR is hypothesized to be near 99.6%. In this reference our accomplished average performance with NPCR=99.63% affirms suitability towards a real-time IoT security system.

TABLE III. NPCR PERFORMANCE OVER DIFFERENT PLAINTEXT SAMPLES

Input Samples	NPCR (%)
1	99.60
2	99.483
3	99.21
4	100.00
5	100.00
6	99.63
7	99.52
8	99.62
Average NPCR	99.63

3) *Unified Averaged Changed Intensity (UACI)*: UACI represents the difference of the mean values between the cipher texts. Typically, the 100% of UACI states that both texts are completely different in amplitude. Mathematically, we used (20) to estimate the UACI performance.

$$UACI = \frac{100}{l \times 95} \sum_{i=1}^l |E_1(i) - E_2(i)| \quad (20)$$

In (20), l represents the text-length, while $E_1(i)$ and $E_2(i)$ signify the symbol value of the cipher texts. For a single text encryption method, (i.e., a unit plaintext is processed at once), $E_i(i)$ can be applied for UACI estimation. The UACI values for the different samples are given in Table IV.

TABLE IV. UACI RANDOMNESS ASSESSMENT

Input Samples	UACI (%)
1	35.77
2	38.32
3	43.81
4	45.00
5	44.00
6	41.06
7	44.87
8	43.31
Average UACI	42.01

4) *Encryption time analysis*: In sync with IoT systems or allied applications, maintaining minimum encryption time is must. As already discussed, IoT system often performs delay-resilient real-time communication and therefore whether to get access of system or data, or even transmits the data, achieving minimum encryption time is inevitable. Taking into consideration of this fact, we examined encryption time (in *ms*) performance over the different samples. Table V presents the encryption time results by the proposed GFS-SSPN lightweight encryption system. Noticeably, we used (21-23) to estimate the encryption and decryption time (ms) [35].

$$Encryption\ Time\ (ms) = \frac{\text{Time required to encrypt i-th data}}{\text{Total number of data}} \quad (21)$$

$$Decryption\ Time\ (ms) = \frac{\text{Time required to Decrypt i-th data}}{\text{Total number of data}} \quad (22)$$

$$Execution\ time = Encryption\ time\ (ms) + Decryption\ time\ (ms) \quad (23)$$

TABLE V. EXECUTION TIME

Plaintext Input Samples	Execution Time (ms)
1	1.67
2	6.02
3	6.31
4	9.82
5	6.98
6	11.41
7	12.14
8	23.28
Average Exec time (ms)	9.70

To be noted, the sample sizes (in bits) considered in this study were varied (here, increasing order) from 1 to 8. The initial sample size we considered for sample-1 was 256 bits, while we kept increasing the size from 1 to 8 samples. In sync with increasing data volume, the encryption and decryption time too increased. However, the average time for execution (including both encryption as well as decryption) was 9.70 ms. This time is sufficiently small to accommodate real-time encryption purposes. To be noted, typically the time efficiency of an algorithm depends on computer's efficacy as well. Since, the proposed model was simulated over Intel i5 processor, execution over superior CPU configuration can yield even more affirmative results.

5) *Inter-Model performance assessment:* This is the matter of fact that the proposed GFS-SSPN lightweight encryption model exhibited significantly well towards delay-resilient and quality-sensitive encryption for IoT. Undeniably, the performance outcomes in terms of correlation, NPCR, UACI and execution time (ms) is encouraging; however, to validate efficacy of the proposed model towards IoT ecosystem we performed an inter-model performance assessment. In this approach we compared the performance of the proposed GFS-SSPN model with the other state-of-art existing methods. Factually, a very few researches have addressed lightweight encryption based IoT security system, and a few researches applying this method are developed for multimedia data security to be used in IoT applications [6]. Researchers [20][21] [40][41] emphasized their lightweight encryption model towards image data security over IoT platforms [6]. Furthermore, these authors have applied different random inputs to assess respective performance; and therefore, comparing performance over the same data is difficult. Considering this fact, we examined relative performance as average performance by the different algorithms.

Authors [21] developed a dynamic structure based lightweight encryption system for image encryption for IoT communication. Unlike our proposed model, authors designed a lightweight encryption concept with forward and backward chaining blocking concept (FBC) that in conjunction with a permutation block performed encryption. Though, similar to our approach authors too suggested applying multi-layered dynamic keying concept to introduce higher randomness and hence high attack-resilience. Noticeably, their proposed cipher

layer comprised a binary diffusion matrix, S-box and P-Box, in sequence. Authors [15][42][43] affirmed their efficacy backed by a binary diffusion matrix, substitution and permutation table. Considering about the NPCR performance by [43], the highest NPCR obtained was 50.10, while UACI for the same algorithm was obtained as 99.64. In sync with the inference in [39][41], for a robust lightweight encryption model retaining maximum possible NPCR (near 100%) is must. Therefore, [43] fails in delivering the expected performance. Murillo-Escobar et al., [39] too designed a lightweight encryption model using Chaotic map based symmetric text-cipher generation [36]. Though, authors applied 128-bits key, the use of two logistic maps with optimized pseudorandom sequences were considered for encryption, made it too computationally exhaustive. Authors employed single round of permutation diffusion to reduce overhead of encryption. A snippet of the recent approaches and their corresponding performance is given in Table VI.

Noticeably, the work proposed by Murillo-Escobar et al. (2014) applied 128-bit key. On the contrary we employed only 64-bit key for encryption. A snippet of the different lightweight encryption models and their configurations is given in Table VII.

Table VI presents some of the key lightweight encryption systems and allied functional architectures like block size, key size and number of encryption rounds. Observing the overall results, it can be found that the proposed GFS-SSPN model employs minimum key size (here, 64) with minimal encryption rounds (here, five) and even retains optimal performance. Though, a few works like KHAZAD [68][29][52], etc. applied low encryption round; however, retained 128 as key size. Similarly, the approaches employing 64-bit of encryption key have applied higher encryption round to retain higher confusion or randomness in cipher. The comparative assessment also reveals (Table VI) that most of the existing lightweight encryption models for IoT have applied SPN as encryption technologies. Though, a few works like [29][45][60] and [67] applied FN assisted SPN for encryption; however, the key size used were 128 or 256 bits. These approaches can easily be visualized to be highly computationally complex and exhaustive, demanding higher gate element (GE) for hardware realization. Another lightweight encryption models such as SPARX [49], LAX [49], Chaskey [52] and LEA [57] employed SPN with ARX-based S-boxes for encryption. However, these approaches demanded higher key-size (128 minimum) that makes them computationally as well as resource exhaustive.

TABLE VI. RELATIVE PERFORMANCE ANALYSIS

Technique	NPCR (%)	UACI (%)	Correlation	Encryption Time (ms)
[43]	50.1029	99.6460	-	-
[39]	98.85	33.31	-	140 ms for 1126 symbols or characters
GFS-SSPN	99.63	43.01	0.0082	6.70 ms

TABLE VII. DIFFERENT LIGHTWEIGHT CRYPTOSYSTEMS FOR IOT SECURITY

Ciphers	Technique	Key Size	Block size	No. of Rounds
Improved Lilliput [44]	EGFN	80	64	30
GIFT [23]	SPN	128	64/128	28/40
SIT [45]	Feistel + SPN	64	64	5
DLBCA [46]	Feistel	80	32	15
LiCi [48]	Feistel	128	64	31
SKINNY [24]	SPN	64-384	64/128	32-56
MANTIS [24]	SPN	128	64	10/12
SPARX [49]	SPN with ARX-based S-boxes	128/256	64/128	24-40
LAX [49]	SPN with ARX-based S-boxes	128/256	64/128	24-40
RoadRunneR [50]	Feistel	80/128	64	10/12
PICO [25]	SPN	128	64	32
RECTANGLE [51]	SPN	80/128	64	25
Chaskey [52]	SPN with ARX-based S-boxes	128	128	8
OLBCA [23]	SPN	80	64	22
ITUBee [53]	Feistel	80	80	20
HISEC [47]	Feistel	80	64	15
LAC [54]	Feistel	80	64	16
SIMON [55]	Feistel	64/ 72/ 96/ 128/ 144/ 192/ 256	32/48/64/96/128	32/36/42/44/52/54/68/69/72
SPECK [55] [69]	Feistel	32/ 64/ 72/ 96/ 128	64/ 72/ 96/ 128/ 144/ 192/ 256	22/ 23/ 26/ 27/ 28/ 29/ 32/ 33/ 34
FeW [56]	Feistel	80/128	64	32
LEA [57]	SPN with ARX-based S-boxes	128/192/256	128	24/28/32
SCREAM [26]	SPN	128	128	10/12
PRINCE [27]	SPN	128	64	12
Hummingbird-2 [29]	SPN + Feistel	128	64	4
TWINE [58]	GFN	80/128	64	36
LED [59]	SPN	64/128	64	32/48
LBlock [60]	Feistel + SPN	80	64	32
PICCOLO [61]	GFN	80/128	64	25/31
KLEIN [28]	SPN	64/80/96	64	12/16/20
CLEFIA [62]	Feistel	128/192/256	128	18/22/26
PRESENT [63]	SPN	80/128	64	31
SEA [64]	Feistel	96	96	93
mCrypton [65]	SPN	64/96/128	64	12
TDEA [66]	Feistel	64	64	48
Camelia [67]	Feistel + SPN	128/192/256	128	18/24/24
KHAZAD [68]	SPN	128	64	3
Proposed GFN-SPNN	GFS+SSPN	64	64	5

B. Qualitative Assessment

In addition to the above discussed performance assessment, to assess attack-resilience of the proposed GFS-SSPN lightweight encryption system we performed qualitative assessment. In this method we examined the robustness and attack-resilience of the proposed lightweight encryption system

in IoT-ecosystem [70]. In real-world IoT communication environment, an attacker can intercept the cipher stored (using SCLA), communicated (Linear and differential attack or impersonation) etc. and can attack the same to retrieve the targeted content or credential. Functionally, a cipher can be stated to be breached in case the attacker gets access or

becomes able to retrieve the secret key. The encryption method applied in the proposed model was robust enough to alleviate the different kinds of attack-vulnerability. In this reference, we examine its robustness theoretically.

1) *Differential and linear cryptanalysis*: The proposed GFS-SSPN block-cipher model employs DKE component that functionally applies different linear and non-linear mathematical functions, substitution a permutation, along with cyclic flipping concepts to introduce higher randomness in cipher without using large key or higher encryption rounds. This approach strengthens it to avoid any cryptanalysis or linear and differential attack probabilities. Moreover, as discussed above, the correlation in between the plaintext and the ciphertext is significantly low exhibiting higher imperceptibility that can be vital to avoid any linear attacks in IoT-ecosystem. Moreover, GFS assisted confusion and diffusion over five consecutive rounds too strengthened the proposed model to achieve higher attack-resilience. Additionally, since the (each) round transformation is maintained uniform which enables treating each bit similar and hence facilitates resilience to the differential attack. The results obtained in terms of NPCR and UACI, as discussed above reveals that the proposed block-cipher model can have sufficient resilience to avoid any kind of differential attacks probability over IoT ecosystem.

2) *Weak key combination*: In major operating conditions, users make a common mistake by keeping poor or weak key combination that helps attacker to get easy access to the ciphers. On contrary, the cipher information where the non-linear operations usually rely on the key value maps the block cipher in such manner that it causes detectable weakness. On the other hand, looking into the proposed security model where it avoids using the same (actual) key in the cipher (due to multiple round key manipulation and/or exchange by XORing the actual key followed by DKE for five rounds). It makes our proposed GFS-SSPN model robust enough to avoid any kind of week-key attack probability.

3) *Related keys combination trial attack*: The attack can be made with the help of certain partially known or unknown keys as well. The related keys primarily depend on either slow diffusion or possessing symmetry in key expansion block, as discussed in the previous section. In our proposed security model, we crafted the key expansion mechanism in such manner that it retains fast computation and non-linear diffusion, especially for the cipher key difference in comparison to the round keys that makes significant confusion to assess related key for credential level or under transit data attack.

4) *Square-attack*: To assess efficacy of a security model, different attack modules are applied to investigate attack-resiliency by the proposed approach. Some of the key approaches applied in cloud-sensitive security models are the RS-Analysis and Square Attack. Considering Square Attack condition, it is capable enough to retrieve one byte of the last key combination and intends to retrieve or recover rest of the

keys by repeating the attack iteratively. Let, such repetition be eight times, then also to achieve above stated information, the attacker needs to identify 28 keys precisely by 28 plaintexts which is equivalent to 216- S-box lookups. This becomes highly complicate and thus the proposed model can avoid such attack efficiently.

5) *Interpolation attack*: In general, such kinds of attacks primarily rely on the generic architecture of the cipher components which could generate certain rational expression with relatively low complexity. However, as already discussed the S-box expression of the proposed security system with diffusion characteristics strengthen it to avoid such limitations and thus makes it impracticable enough to avoid attack.

The above-mentioned relative performance assessment (Table VIII and Table IX), affirms suitability of the proposed GFS-SSPN lightweight security model towards realistic IoT ecosystems.

TABLE VIII. COMPARISON OF TWO FACTOR AUTHENTICATION PERFORMANCES FOR A CONVENTIONAL MODEL WITH OUR PREVIOUS WORKS

Security property	Amin [71]	Tan [72]	Xie [73]	TFA-PUF-IoT [74]	TFA-RPUF-IoT [76]	Proposed GFS-SSPN
Resilience to the impersonation attack	Yes	Yes	Yes	Yes	Yes	Yes
Anonymity and un traceability	Yes	No	Yes	Yes	Yes	Yes
Resilience to the password guessing attack	No	Yes	Yes	Yes	Yes	Yes
Prevents clock synchronization problem	No	Yes	No	Yes	Yes	Yes
Device security	No	No	No	Yes	Yes	Yes
Deployed security algorithm	ECC	ECC	ECC	PUF and FE	RPUF-FE	GFS-SSPN
Random response for every clock cycle	No	No	No	No	Yes	Yes
SCLA	Yes	No	No	No	Yes	Yes
Brute Force	No	No	No	No	No	Yes

TABLE IX. COMPARISON OF SECURITY PERFORMANCES

Comparison matrices	Aman [75]	TFA-PUF-IoT [74]	TFA-RPUF-IoT [76]	GFS-SSPN
Mutual authentication	Yes	Yes	Yes	Yes
Two factor secrecy	No	Yes	Yes	Yes
Privacy of the IoT devices	No	Yes	Yes	Yes
Consideration of noise in the PUF	No	Yes	Yes	Yes
Protection against physical attacks	Yes	Yes	Yes	Yes
Random response for every clock cycle/encryption round	No	No	Yes	Yes
Block cipher	No	No	No	Yes

V. CONCLUSION

This paper primarily focused on designing a state-of-art new and robust dynamic programming assisted lightweight encryption system for IoT security. Unlike classical cryptosystem-based approaches which often undergo increased computational overhead, latency and more importantly resource exhaustion, this research contributed a state-of-art new lightweight encryption system. The proposed model intended to exploit efficacy of both SPN with GFS, where the first enabled optimal confusion and diffusion (in cipher) while the later enabled reduced computation with significantly smaller decryption cost. Additionally, towards "Fit-To-All" IoT security solution for both device-level security as well as data-level security, the proposed model employed Shannon Criteria based SPN (SSPN) which helped in error-free decryption. On the other hand, the use of GFS enabled block-cipher encryption helped in enhancing computation time as well as cost. Architecturally, the proposed GFS-SSPN model employed 64-bit dynamic keying with five rounds of encryption which retained minimal computation and allied cost. It can be vital towards resource efficient hardware realization. A key notable contribution of the proposed model can be the use of dynamic keying concept was derived by processing four 16-bit inputs by means of the SPN followed by cyclic rotation and concatenation. The use of 64-bit key with merely five round of encryption helped maintaining minimal computational overheads and time, which can be vital for IoT security systems. The overall proposed GFS-SSPN model was realized as block-cipher model and hence unlike bit-by-bit encryption it reduced computational overhead and delay significantly. Moreover, the use of predefined transformation coder for SPN (substitution purpose) enabled time-efficient confusion and diffusion process to assure IoT-centric efficacy. Noticeably, being the Feistel Network based Shannon Conditional GFS-SSPN based lightweight encryption model avoided distinct decryption program and therefore can be more resource and computationally efficient. In this reference, the

proposed GFS-SSPN model was examined with the different inputs' credentials, often in plaintext to perform encryption. To assess performance by the proposed GFS-SSPN model, different plaintext samples were taken into consideration where performance in terms of correlation, encryption time, NPCR and UACI confirmed superiority over the existing approaches, including other lightweight cipher models as well as symmetric key cryptographic concepts. The depth analysis revealed that the correlation in between the original plaintext and the encrypted cipher was significantly low 0.006, while it maintained NPCR of 99.6% and UACI of 27%. The encryption time observed was near 2.6 ms which is fairly in sync with major IoT systems and allied real-time encryption demands. The overall results affirmed that the proposed system can be well suited for IoT security system, where it can be applied as a device middleware to safeguard access-level security, while the same can be applied to encrypt the input to ensure attack-resilience during transmission. Qualitative assessment too revealed that the proposed model can be effective enough to alleviate the attacks of MITM type, SCLA, Interpolation as well as linear and differential attacks. Though, GFS-SSPN model exhibited satisfactory performance and efficacy towards IoT security, computationally efficient performance etc. Considering the hardware compatibility, the proposed system is developed using C programming language. However the power and resource profile could not be examined. In feature authors can implement the proposed method in a hardware efficacy so as to examine the power and resource utilization for real word realization.

REFERENCES

- [1] V. A. Thakor, M. A. Razzaque and M. R. A. Khandaker, "Lightweight Cryptography Algorithms for Resource-Constrained IoT Devices: A Review, Comparison and Research Opportunities," in IEEE Access, vol. 9, pp. 28177-28193, 2021.
- [2] Barki A., Bouabdallah A., Gharout S. and Traoré J., "M2M Security: Challenges and Solutions," in IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1241-1254, Secondquarter 2016
- [3] J. Sun, H. Xiong, X. Liu, Y. Zhang, X. Nie and R. H. Deng, "Lightweight and Privacy-Aware Fine-Grained Access Control for IoT-Oriented Smart Health," in IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6566-6575, July 2020.
- [4] B. Aboushousha, R. A. Ramadan, A. D. Dwivedi, A. El-Sayed and M. M. Dessouky, "SLIM: A Lightweight Block Cipher for Internet of Health Things," in IEEE Access, vol. 8, pp. 203747-203757, 2020.
- [5] Philip M. A. and Vaithyanathan, "A survey on lightweight ciphers for IoT devices," 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy), Kollam, 2017, pp. 1-4
- [6] X. Guo, J. Hua, Y. Zhang and D. Wang, "A Complexity-Reduced Block Encryption Algorithm Suitable for Internet of Things," in IEEE Access, vol. 7, pp. 54760-54769, 2019.
- [7] U. Hijawi, D. Unal, R. Hamila, A. Gastli and O. Ellabban, "Lightweight KPABE Architecture Enabled in Mesh Networked Resource-Constrained IoT Devices," in IEEE Access, vol. 9, pp. 5640-5650, 2021.
- [8] S. Banerjee et al., "A Provably Secure and Lightweight Anonymous User Authenticated Session Key Exchange Scheme for Internet of Things Deployment," in IEEE Internet of Things Journal, vol. 6, no. 5, pp. 8739-8752, Oct. 2019.
- [9] E. Uchiteleva, A. R. Hussein and A. Shami, "Lightweight Dynamic Group Rekeying for Low-Power Wireless Networks in IIoT," in IEEE Internet of Things Journal, vol. 7, no. 6, pp. 4972-4986, June 2020.
- [10] V. Odelu, A. K. Das, M. Khurram Khan, K. R. Choo and M. Jo, "Expressive CP-ABE Scheme for Mobile Devices in IoT Satisfying

- Constant-Size Keys and Ciphertexts," in IEEE Access, vol. 5, pp. 3273-3283, 2017
- [11] Hameed A. and Alomary A., "Security Issues in IoT: A Survey," 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakhier, Bahrain, 2019, pp. 1-5.
- [12] M. Ali, M. Sadeghi and X. Liu, "Lightweight Revocable Hierarchical Attribute-Based Encryption for Internet of Things," in IEEE Access, vol. 8, pp. 23951-23964, 2020.
- [13] W. Yu and S. Köse, "A Lightweight Masked AES Implementation for Securing IoT Against CPA Attacks," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 64, no. 11, pp. 2934-2944, Nov. 2017.
- [14] D. Vergnaud, "Comment on "Efficient and Secure Outsourcing Scheme for RSA Decryption in Internet of Things"," in IEEE Internet of Things Journal, vol. 7, no. 11, pp. 11327-11329, Nov. 2020.
- [15] F. Noura; L. Sleem; M. Noura; M. M. Mansour; A. Chehab; R. Couturier; "A new efficient lightweight and secure image cipher scheme", *Multimed Tools Application* (springer), 2017.
- [16] Sehrawat D. and Gill N. S., "Lightweight Block Ciphers for IoT based applications: A Review", *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 13, Number 5 (2018) pp. 2258-2270.
- [17] Salami S. Al, Baek J., Salah K. and E. Damiani, "Lightweight Encryption for Smart Home," 2016 11th International Conference on Availability, Reliability and Security (ARES), Salzburg, 2016, pp. 382-388,
- [18] Leander G., Paar C., Poschmann A., and Schramm K, "New Lightweight DES Variants" *FSE 2007, LNCS*, vol. 4593, pp. 196-210. Springer, 2007.
- [19] Kane L. E., Chen J. J., Thomas R., Liu V. and Mckague M., "Security and Performance in IoT: A Balancing Act," in IEEE Access, vol. 8, pp. 121969-121986, 2020.
- [20] Elhoseny M., Ramírez-González G., O. Abu-Elnasr M., S. A. Shawkat, N. Arunkumar and A. Farouk, "Secure Medical Data Transmission Model for IoT-Based Healthcare Systems," in IEEE Access, vol. 6, pp. 20596-20608, 2018.
- [21] Noura H., Couturier R., C. Pham and Chehab A., "Lightweight Stream Cipher Scheme for Resource-Constrained IoT Devices," 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Barcelona, Spain, 2019, pp. 1-8.
- [22] Standaert F.-X., Piret G., Rouvroy G., Quisquater J.-J., and J.-D. Legat. ICEBERG: an Involutional Cipher Efficient for Block Encryption in Reconfigurable Hardware. In B. Roy and W. Meier, editors, *Fast Software Encryption — FSE 2004*, pages 279–298. Springer-Verlag, 2004.
- [23] Banik, S., Pandey, S.K., Peyrin, T., Sasaki, Y., Sim, S.M. and Todo, Y., 2017, September. GIFT: a small PRESENT. In *International Conference on Cryptographic Hardware and Embedded Systems* (pp. 321-345). Springer, Cham.
- [24] Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P. and Sim, S.M., 2016, August. The SKINNY family of block ciphers and its low-latency variant MANTIS. In *Annual Cryptology Conference* (pp. 123-153). Springer Berlin Heidelberg.
- [25] Bansod, G., Pisharoty, N. and Patil, A., 2016. PICO: An Ultra Lightweight and Low Power Encryption Design for Ubiquitous Computing. *Defence Science Journal*, 66(3).
- [26] Grosso, V., Leurent, G., Standaert, F., Varici, K., Journault, A., Durvaux, F., Gaspar, L. and Kerckhof, S., 2015. SCREAM Side-Channel Resistant Authenticated Encryption with Masking. *CAESAR* submission.
- [27] Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knezevic, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C. and Rombouts, P., 2012, December. PRINCE—a low-latency block cipher for pervasive computing applications. In *International Conference on the Theory and Application of Cryptology and Information Security* (pp. 208-225). Springer, Berlin, Heidelberg.
- [28] Gong, Z., Nikova, S. and Law, Y.W., 2011. KLEIN: A new family of lightweight block ciphers. *RFIDSec*. Springer, 7055, pp.1-18.
- [29] Engels, D.W., Saarinen, M.J.O., Schweitzer, P. and Smith, E.M., 2011. The Hummingbird-2 Lightweight Authenticated Encryption Algorithm. *RFIDSec*, 11, pp.19-31.
- [30] Mohd B. J. and Hayajneh T., "Lightweight Block Ciphers for IoT: Energy Optimization and Survivability Techniques," in IEEE Access, vol. 6, pp. 35966-35978, 2018.
- [31] Mohammed, A.A. and Ibadi, A.O., 2017. A Proposed Non Feistel Block Cipher Algorithm.
- [32] S. Tan, K. Yeow and S. O. Hwang, "Enhancement of a Lightweight Attribute-Based Encryption Scheme for the Internet of Things," in IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6384-6395, Aug. 2019.
- [33] S. Roy, U. Rawat and J. Karjee, "A Lightweight Cellular Automata Based Encryption Technique for IoT Applications," in IEEE Access, vol. 7, pp. 39782-39793, 2019.
- [34] E. Uchiteleva, A. R. Hussein and A. Shami, "Lightweight Dynamic Group Rekeying for Low-Power Wireless Networks in IIoT," in IEEE Internet of Things Journal, vol. 7, no. 6, pp. 4972-4986, June 2020.
- [35] S. Atiewi et al., "Scalable and Secure Big Data IoT System Based on Multifactor Authentication and Lightweight Cryptography," in IEEE Access, vol. 8, pp. 113498-113511, 2020.
- [36] N. Tsafack et al., "A New Chaotic Map With Dynamic Analysis and Encryption Application in Internet of Health Things," in IEEE Access, vol. 8, pp. 137731-137744, 2020.
- [37] B. Aboushousha, R. A. Ramadan, A. D. Dwivedi, A. El-Sayed and M. M. Dessouky, "SLIM: A Lightweight Block Cipher for Internet of Health Things," in IEEE Access, vol. 8, pp. 203747-203757, 2020.
- [38] L. Zhang, X. Gao and Y. Mu, "Secure Data Sharing with Lightweight Computation in E-Health," in IEEE Access, vol. 8, pp. 209630-209643, 2020.
- [39] Murillo-Escobar M. A., Abundiz-Pérez F., Cruz-Hernández C., Lúpez-Gutiérrez R. M., "A novel symmetric text encryption algorithm based on logistic map", *Proceedings of the 2014 International Conference on Communications, Signal Processing and Compute*, pp. 49-53. 2014.
- [40] Gan Z.-H. Chai X.-L., Han D.-J., and Chen Y.-R., "A chaotic image encryption algorithm based on 3-D bit-plane permutation," *Neural Comput. Appl.*, vol. 31, no. 11, pp. 7111-7130, Nov. 2019.
- [41] Zhou J., Li J. and Di X., "A Novel Lossless Medical Image Encryption Scheme Based on Game Theory with Optimized ROI Parameters and Hidden ROI Position," in IEEE Access, vol. 8, pp. 122210-122228, 2020.
- [42] Noura F. Z. H, Mostefaoui A., "An efficient and secure cipher scheme for images confidentiality preservation", *Signal Process Image Communication*, vol 42, pp.90–108, 2016.
- [43] F. Noura; L. Sleem; M. Noura; M. M. Mansour; A. Chehab; R. Couturier; "A new efficient lightweight and secure image cipher scheme", *Multimed Tools Application* (springer), 2017.
- [44] Mumthaz Pookuzhy Ali, Geethu T George, 2017. "Optimised Design of Light Weight Block Cipher Lilliput with Extended Generalised Feistel Network (EGFN)." *International Journal of Innovative Research in Science, Engineering and Tech.*, 2017.
- [45] Usman, M., Ahmed, I., Aslam, M.I., Khan, S. and Shah, U.A., 2017. SIT: A Lightweight Encryption Algorithm for Secure Internet of Things. *arXiv preprint arXiv:1704.08688*.
- [46] AlDabbagh, S.S.M., 2017. Design 32-bit Lightweight Block Cipher Algorithm (DLBCA). *International Journal of Computer Applications*, 166(8).
- [47] AlDabbagh, S.S.M., Shaikhli, A., Taha, I.F. and Alahmad, M.A., 2014, September. HISEC: A new lightweight block cipher algorithm. In *Proceedings of the 7th International Conference on Security of Information and Networks* (p. 151). ACM.
- [48] Patil, J., Bansod, G. and Kant, K.S., 2017, February. LiCi: A new ultra-lightweight block cipher. In *Emerging Trends & Innovation in ICT (ICEI)*, 2017 IEEE International Conference, pp. 40-45.
- [49] Dinu, D., Perrin, L., Udovenko, A., Velichkov, V., Großschädl, J. and Biryukov, A., 2016. Design strategies for ARX with provable bounds: Sparx and LAX. In *Advances in Cryptology—ASIACRYPT 2016: 22nd International Conference on the Theory and Application of Cryptology*

- and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I 22(pp. 484-513). Springer Berlin Heidelberg.
- [50] Baysal, A. and Şahin, S., 2015, September. Roadrunner: A small and fast bitslice block cipher for low cost 8-bit processors. In International Workshop on Lightweight Cryptography for Security and Privacy (pp. 58-76). Springer, Cham.
- [51] Zhang, W., Bao, Z., Lin, D., Rijmen, V., Yang, B. and Verbauwhede, I., 2015. RECTANGLE: a bit-slice lightweight block cipher suitable for multiple platforms. *Science China Information Sciences*, 58(12), pp.1-15.
- [52] Mouha, N., Mennink, B., Van Herrewewege, A., Watanabe, D., Preneel, B. and Verbauwhede, I., 2014, August. Chaskey: an efficient MAC algorithm for 32-bit microcontrollers. In International Workshop on Selected Areas in Cryptography (pp. 306-323). Springer, Cham.
- [53] Karakoç, F., Demirci, H. and Harmanci, A.E., 2013, May. ITUbee: a software oriented lightweight block cipher. In International Workshop on Lightweight Cryptography for Security and Privacy (pp. 16-27). Springer, Berlin, Heidelberg.
- [54] Zhang, L., Wu, W., Wang, Y., Wu, S. and Zhang, J., 2014. LAC: A lightweight authenticated encryption cipher. CAESAR competition.
- [55] Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B. and Wingers, L., 2013. The SIMON and SPECK Families of Lightweight Block Ciphers. Cryptology ePrint Archive, Report 2013/404.
- [56] Kumar, M., Pal, S.K. and Panigrahi, A., 2014. FeW: A Lightweight Block Cipher. IACR Cryptology ePrint Archive, 2014, p.326.
- [57] Hong, D., Lee, J.K., Kim, D.C., Kwon, D., Ryu, K.H. and Lee, D.G., 2013, August. LEA: A 128-bit block cipher for fast encryption on common processors. In International Workshop on Information Security Applications (pp. 3-27). Springer, Cham.
- [58] Suzaki, T., Minematsu, K., Morioka, S. and Kobayashi, E., 2011, November. Twine: A lightweight, versatile block cipher. In ECRYPT Workshop on Lightweight Cryptography.
- [59] Guo J., Peyrin T., Poschmann A., and Matt Robshaw M, Preneel B. and Takagi T., 2011. The LED Block Cipher. CHES 2011, In International Association for Cryptologic Research, LNCS 6917 (pp. 326–341).
- [60] Wu, W. and Zhang, L., 2011. LBlock: a lightweight block cipher. In Applied Cryptography and Network Security (pp. 327-344). Springer Berlin/Heidelberg.
- [61] Shibutani, K., Isobe, T., Hiwatari, H., Mitsuda, A., Akishita, T. and Shirai, T., 2011, September. Piccolo: An ultra-lightweight blockcipher. In CHES (Vol. 6917, pp. 342-357).
- [62] Shirai, T., Shibutani, K., Akishita, T., Moriai, S. and Iwata, T., 2007, March. The 128-bit block cipher CLEFIA. In FSE (Vol. 4593, pp. 181-195).
- [63] Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J., Seurin, Y. and Vikkelsoe, C., 2007, September. PRESENT: An ultra-lightweight block cipher. In CHES (Vol. 4727, pp. 450-466).
- [64] Mace F., Standaert F.-X., and Quisquater J.-J., “ASIC Implementations of the Block Cipher SEA for Constrained Applications”, In RFID Security-RFID sec 2007, Workshop Record, pages 103 – 114, Malaga, Spain, 2007.
- [65] Lim C. and Korkishko T. mCrypton - A Lightweight Block Cipher for Security of Low-cost RFID Tags and Sensors. In J. Song, T. Kwon, and M. Yung, editors, Workshop on Information Security Applications-WISA 2005, volume 3786 of Lecture Notes in Computer Science, pages 243–258. Springer-Verlag, 2005.
- [66] Barker, W.C. and Barker, E., 2012. Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher: NIST Special Publication 800-67, Revision 2.
- [67] Aoki, K., Ichikawa, T., Kanda, M., Matsui, M., Moriai, S., Nakajima, J. and Tokita, T., 2000, August. Camellia: A 128-bit block cipher suitable for multiple platforms-design and analysis. In Selected Areas in Cryptography (Vol. 2012, pp. 39-56).
- [68] Barreto, P.S.L.M. and Rijmen, V., 2000. The Khazad legacy-level block cipher. Primitive submitted to NESSIE, 97.
- [69] Ray B., Douglas S., Jason S., Stefan T., Bryan W., and Louis W., “The simon and speck families of lightweight block ciphers,” Cryptology ePrint Archive, Report. /404, Tech. Rep., 2013.
- [70] Urunov K., Namgung J. and Park S., "Security analysis based on Trusted Environment (TRE) of M2M/IoT," 2015 17th Asia-Pacific Network Operations and Management Symposium, Busan, 2015, pp. 554-557,
- [71] R. Amin, S. K. Islam, M. K. Khan, A. Karati, D. Giri, and S. Kumari, “A two-factor RSA-based robust authentication system for multi-server environments”, Security and Communication Networks, 2017.
- [72] J. Qu, and X. L. Tan, “Two-factor user authentication with key agreement scheme based on elliptic curve cryptosystem”, Journal of Electrical and Computer Engineering, 2014.
- [73] Q. Xie, D. S. Wong, G. Wang, X. Tan, K. Chen, and L. Fang, “Provably secure dynamic ID- based anonymous two-factor authenticated key exchange protocol with extended security model”, IEEE Transactions on Information Forensics and Security, Vol. 12, No. 6, pp. 1382-1392, 2017.
- [74] P. Gope and B. Sikdar, “Lightweight and privacy-preserving two-factor authentication scheme for IoT devices”, IEEE Internet of Things Journal, Vol. 6, No. 1, pp. 580-589, 2018.
- [75] M. N. Aman, K. C. Chua, and B. Sikdar, “Mutual authentication in IoT systems using physical unclonable functions”, IEEE Internet of Things Journal, Vol. 4, No. 5, pp. 1327-1340, 2017.
- [76] K. P. Gurumanapalli, N. Muthuluru, “A Non Linear PUF Circuit Design for Two Factor Authentication in IoT Cryptography”, International Journal of Intelligent Engineering and Systems, Vol.14, No.1, 2021.

Cloud-based Secure Healthcare Framework by using Enhanced Ciphertext Policy Attribute-Based Encryption Scheme

Siti Dhalila Mohd Satar¹
Mohamad Afendee Mohamed²

Faculty of Informatics and Computing
Universiti Sultan Zainal Abidin, Terengganu, Malaysia

Masnida Hussin^{3*}, Zurina Mohd Hanapi⁴
Siti Dhalila Mohd Satar⁵

Faculty of Computer Science and Information Technology
Universiti Putra Malaysia, Selangor, Malaysia

Abstract—Cloud computing is an emerging technology that has been used to provide better healthcare services to the users because of its convenient and economical features. Noted that the healthcare services required fast and reliable data sharing at anytime from anywhere for better monitoring and decision making in medical requirements. However, the privacy and integrity of electronic healthcare record become a significant issue during data sharing and outsourcing in Cloud. The data privacy of clients/patients is important in healthcare services where exposure of the data to unauthorized parties is unexceptional. In order to address this security loophole, this paper presents a Cloud-based Secure Healthcare Framework (SecHS) to offer safe access to healthcare and medical data. Specifically, this paper enhance the Ciphertext Policy Attribute-Based Encryption (CP-ABE) scheme by adding two more modules which aims to provide fine-grained access control and offer privacy and integrity of data. It facilitates encryption and hashing schemes. The proposed framework is compared with existing frameworks that used CP-ABE scheme. It shows the SecHS offers better features towards securing the healthcare services data. Optimistically, data security requirements such as privacy, integrity and fine-grained access control are required to effectively proposed for assuring data sharing in the Cloud environment.

Keywords—Cloud computing; privacy and integrity; fine-grained access control; Ciphertext Policy Attribute-Based Encryption; electronic health record

I. INTRODUCTION

Cloud technology offers an innovative computing method for delivering IT services efficiently. Cloud technology able to enhance the quality of services (QoS) in numerous fields, including healthcare and medical services[1],[2]. Basically, in the healthcare services, electronic health record (EHR) has been widely used to improve storing, accessibility, sharing and realized a collaboration of medical data among medical practitioners. The EHR includes patients' data, laboratory results, medication lists, diagnostic tests, physical assessments, and historical observations. Due to most of these records are crucial and confidential, it is recommended by the Health Insurance Portability and Accountability Act (HIPAA) [3],[4] for ensuring the data and documents are safe and protected.

Cloud adoption in healthcare services has gained many attentions as mentioned in [1], [5]–[7] and it offers superfluous

benefits to the hospital and medical organization. The healthcare services through Cloud computing is expected to reduce execution cost hence might improve the services delivery. Furthermore, the resource management and system administration (infrastructure) can be effectively monitored through Cloud computing makes healthcare service is easy to maintain [8]. However, despite all the advantages, security is one of the most important challenges in Cloud computing. Particularly, the systems that have been used by the healthcare practitioner or end-users are vulnerable to many security issues. It is due to the medical data is confidential and data leaked by the irresponsible entity are unacceptable. Moreover, according to [9], [10], [11] the fact that the data is stored in the Cloud and can be resided anywhere and beyond the geographical boundary might cause the users to lose control over their own data. Other security concern includes the issue of the healthcare and medical organizations that need to have a clear agreement including security concern with the Cloud Service Provider (CSP). This involves security and access control procedures. There is very often where the Cloud users are not given a thorough explanation of their security concerns and needs in renting the Cloud services from CSP. Therefore, the need for data integrity and privacy mechanism which offer fined-grained access control are a necessity in order to provide better security services to both Cloud users and CSP.

This study proposed a Secure Healthcare Framework (SecHC) in Cloud computing using Ciphertext Policy Attribute-Based Encryption (CP-ABE). It aims to provide secure access to healthcare and medical data in the Cloud environment. In this framework, the patient's data is encrypted under Symmetric Encryption Scheme and the access policy in CP-ABE is embedded with the ciphertext. The contribution is summarized as follows:

- Provide a fine-grained access control by implementing the CP-ABE scheme which suited for a Cloud-based electronic health record system.
- Model security analysis related to security requirements in the Cloud environment includes privacy, integrity and fine-grained access control.

The remaining paper organization is as follows. Section 2 discuss on related work. In section 3, a simulation is conducted to proof the weakness of existing work. The proposed secure

*Corresponding Author

healthcare framework is provided in Section 4. Section 5 discussed the feature and security requirement analysis against other frameworks. Lastly, Section 6 provides a conclusion.

II. RELATED WORK

The area of healthcare and medical in Cloud computing is commonly implemented and realized. There are several approaches and techniques are used in securing the Cloud-based healthcare system in the data sharing process.

A. Healthcare and Medical in Cloud

The rapid development of Cloud computing nowadays has transformed the way of healthcare provider and even medical practitioners such as doctors and hospitals, to provide a quality and affordable service to their clients. This transformation is motivated by two major factors which are the business commanding to reduce costs and to convalesce the quality of care [12]-[14]. From the business side of view, the providers can lower their operational expenses to compensate for the increasing costs of infrastructure, administrative and pharmaceutical. Simultaneously, they also must handle requests from the government to increase the quality of healthcare and delivered a common healthcare operating standard [15], [16].

Meanwhile, on the patient side, the provider must offer a service that provides instant and top-quality access because nowadays, a patient is acquainted with the 24/7 accessibility of services especially from the online retailers and financial institutions. Furthermore, currently, the users are interested to involve in managing their own healthcare so the need for a system that could provide diagnosis, information, and treatments is in demand [3], [14]. For example, a patient would like an internet-based service from the healthcare providers that provide a platform for them to converse or consult the healthcare professionals all the time especially before, during and after any health-related procedures. This situation shows that a dire need of healthcare providers to transform themselves from a traditional to a Cloud environment to tackle the business and patient needs of an agile environment and to revolutionize their IT infrastructure.

Although the implementation of Cloud computing in healthcare sectors may seem beneficial and positive, it also fosters many detrimental situations and challenges. Despite the emergent trends of using Cloud computing as the platform to promote healthcare, the anxiety over the security and privacy of confidential information in the Cloud are intensifying over the years [17]. Data leakage and loss, phishing, hijacking of account or service, and unidentified risk profile are an example of the threats that impend the privacy and integrity of Cloud data.

Furthermore, the healthcare organizations have discovered that the existing mechanisms such as Secure Socket Layer protocol (SSL) and Transport Layer Security (TLS) protocol are not sufficient to secure EHR in Cloud because it only protects the privacy during data transmission [3][18]. Apart from it, according to [3], dishonest employees of the CSP can easily overrule a particular role with the authorization to retrieve and read the healthcare data beyond their privileges. Thus, to avoid any infringements of sensitive data by

fraudulent employees and to prevent medical data from other security threats, a secure mechanism must be designed and developed to enhance the data privacy and integrity of EHR in Cloud computing.

Recently, numerous security mechanism has been proposed by researchers to protect medical data in Cloud such as encryption schemes and access control schemes [10][19]. Such schemes permit the data owner to manage the data by restricting access to specific users for a specific file with limited privileges. Fig. 1 shows example of mechanism used to secure EHR in cloud. In this figure, data owner is defined as a patient stored the records in the cloud server and the record can only be accessed and downloaded by authorized physicians. This mechanism helps the healthcare and medical organization to protect the security and privacy of cloud data.

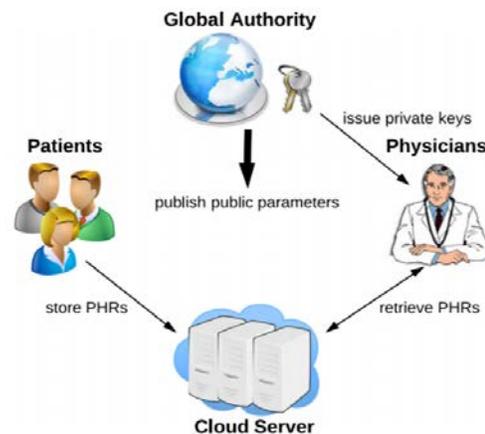


Fig. 1. Securing Electronic Health Record in Cloud [12].

B. Securing Healthcare and Medical Records in Cloud

In medical and health organizations, allowing end-users to control access to the data stored in unreliable cloud servers by using encryption schemes is an effective method to prevent data leakage, ensure safe transmission and secure from any other security threats. According to [20][21], the encryption algorithm permits end-users to encrypt data and ensure only user with the key can decrypt the data. However, this traditional encryption scheme is incapable to satisfy one-to-many encryption which is crucial in the cloud sharing scenario. One-to-many encryption permitting the data owner to encode the data once and it can be decoded or decrypted by several users that have a decryption key.

To deal with this issue, a control access mechanism that supports one-to-many encryption should be implemented to prevent unauthorized access to the Cloud. The author [21] proposes a scheme to support one-to-many encryption using Attribute-Based Encryption (ABE). This scheme is based on a novel patient-centric framework for controlling access to Patient Health Records (PHR) stored in semi-trusted cloud servers.

Meanwhile, authors [13][14][15][16][18][22] proposed another variant of ABE which is CP-ABE scheme wherein the CP-ABE scheme, data owner are allowed to specify the

authorized users to access the data, by enclose the access policy to the ciphertext. In order to generate private keys associated with set of user's attributes, ABE schemes rely on reliable and trusted authorities.

In order to achieve confidentiality and authenticity simultaneously, an efficient and improving CP-ABE scheme has been introduced by [15] and [20]. They used a signature scheme to verify the authenticity of data and the owner of the data. In the meantime, authors [13] and [18] proposed a scheme to enhance the efficiency of the decryption process. They used the Attribute Bloom Filter technique to assess whether an attribute is in the access policy and locate the exact position in the access policy if it is in the access policy. This technique will reduce the decryption time. However, this technique will cause a trade-off in terms of increasing computation overhead.

Besides, authors [21] also construct a scheme used CP-ABE which offers not only fine-grained access control but also providing integrity by checking the access policy of the users before they can access healthcare data. Their works also aimed to offer fast decryption by adding some redundant components to a ciphertext before the decryption stage. However, their works are not sufficient to secure the process of sharing data in the Cloud environment due to the failure to provide a fully hidden policy within the CP-ABE. The same issue applies to the researchers [23] where they proposed a privacy-aware s-health access control system that offers partially policy hiding. In the system, only the attribute value is hidden while the attribute name is sent to the Cloud with ciphertext in a readable form. This will cause data security to be compromised in the event of an attack where the attacker can find out and learn about the policy which leads to privacy leakage.

In addition, Zhang et. al., [17] proposed a privacy-preserving scheme using CP-ABE with efficient authority verification. They preserve the privacy of the data authority identification phase by determining whether the user is authorized or not. Apart from it, a study by [24] proposed a scheme using CP-ABE which successfully achieved high security by hiding the entire access policy of the EHR. However, they overlooked the additional computational cost and decryption time is high due to the newly introduced scheme on verification of the matching process.

Thus, the above literature review indicates that many works have been done to provide additional security towards the security mechanism provided by the CSP. However, according to [17], it still insufficient to protect medical and healthcare records and reduce the risk of threats in the Cloud environment. This is because there are many issues arise such as multi-authority, hidden policy and constant size ciphertext which need to be studied.

C. Framework of Health Data in Cloud

Generally, a framework is built to simplify a complex technological process. It is usually consisting of a few elements, or entities which will be integrated to become a useful process. In medical and healthcare organizations, several frameworks have been proposed to ensure administration of patients, data, staff and operations are running smoothly. For

example, a security framework focusing on data delivery of patient care has been developed by Zhang et al. [25]. In this framework, Zhang has introduced three main components: data collection, secure storage and secure usage model as shown in Fig. 2. The aim of this framework is to provide a safe interaction between healthcare professionals and patients based on security needs and patient privacy in the EHR Cloud.

The first component is collection and integration of various data from various departments in organizations. These components are responsible to make sure all the data is available and safe to be used. Hence, this component needs to verify the data in term of confidentiality, integrity, and ensuring nonrepudiation as well as HIPAA compliance.

The second component is the secure storage that consists of two entities which are secure storage server and access control engine. Data is stored as ciphertext in secure storage server and only authorized user is permitted to access the data. Meanwhile access control engine use role-based and attribute-based method to grant access to the user. As a result, user that does not match the attribute will be denied the access to patient's data in Cloud Storage.

The last component is secure usage model. This component's role is to provide a safe data access by using two methods: signature and verification. Medical practitioners will sign EHR using appropriate signature algorithm before sending it to Cloud storage. Then, user will verify the authenticity of the data by using digital signature verification.

Apart from that, the work proposed by [26] introduces a framework that offers privacy of health data and provides access control to the data in Cloud. There are four entities in this framework are S-Health Authority, S-Health Cloud, Data Owner, and Data User as shown in Fig. 3.

In this framework, S-Health Authority is responsible for system initialization and authorization where it uses attribute-based access to permit authorized user to store or access data in Cloud. Meanwhile, for data owner they manage their EHR by performing encryption process to secure the record. They stored the ciphertext and its access policy in the S-Health Cloud. To enhance the security, they partially hide access policy to prevent untrusted entities take advantages of the access policy.

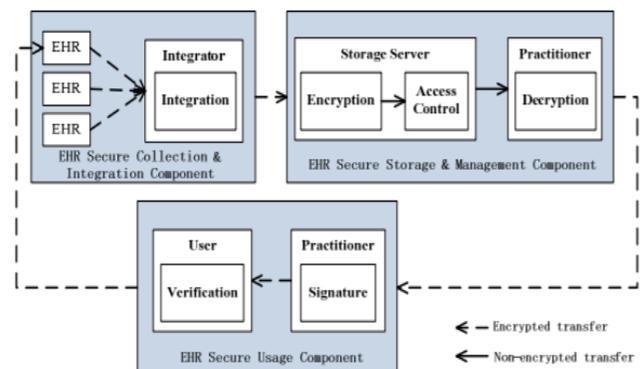


Fig. 2. Security Framework Focusing on Data Delivery of Patient Care [25].

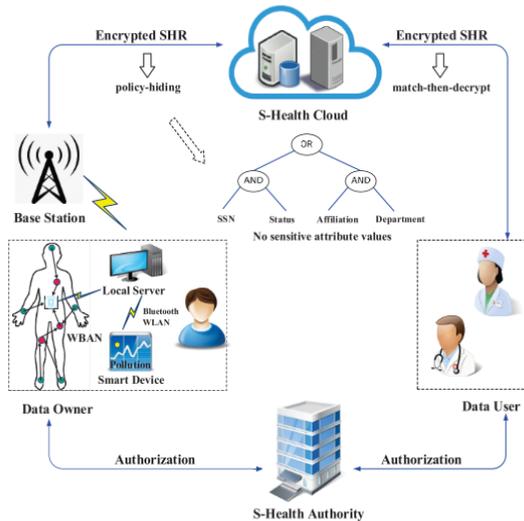


Fig. 3. Framework of the Privacy-Aware S-Health Access Control System [26].

In addition, to make the process efficient, they implement match-then-decrypt procedure in the decryption phase. In this phase, data user will use to his secret key to check whether his attributes match the underlying access policy, and then decrypt the encrypted EHR only if the matching is successful.

Using this framework, the system achieved security requirement by offering fine-grained access control using CP-ABE scheme. They also provide confidentiality using encryption algorithm and offer data privacy using hidden access policy.

However, both frameworks proposed by [25] and [26] can be improved by incorporating other mechanisms such as hashing algorithms which are used to ensure data integrity. In addition, to ensure the privacy of data, fully hidden access policy must be implemented to solve user distrust issues.

III. PRELIMINARIES WORK

This section describes the preliminaries related to the proposed work by [21]. In Zhang's work, they used CP-ABE to provide fine-grained access control to the health data in Cloud storage. The proposed solution was successfully delivered data verifiability and fast decryption by providing the validation of decrypted message. However, the access policy was sent to the Cloud together with the ciphertext in a readable form which led to the data privacy leakage.

In this preliminaries work, a simulation of transferring a file that contained an access policy using FileZilla Tool is conducted. To transfer the file from a client to a server, a File Transfer Protocol (FTP) without encryption is used as shown in Fig. 4.

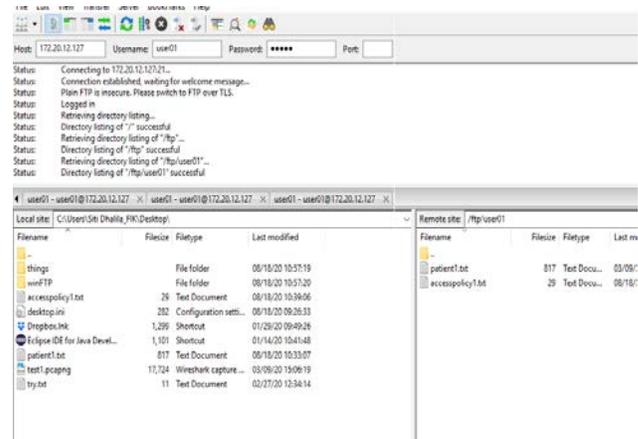


Fig. 4. Transferring a File using FTP in FileZilla Tool.

While transferring a file in FileZilla, Wireshark tool is execute by running a passive attack to sniff the packet as shown in Fig. 5.

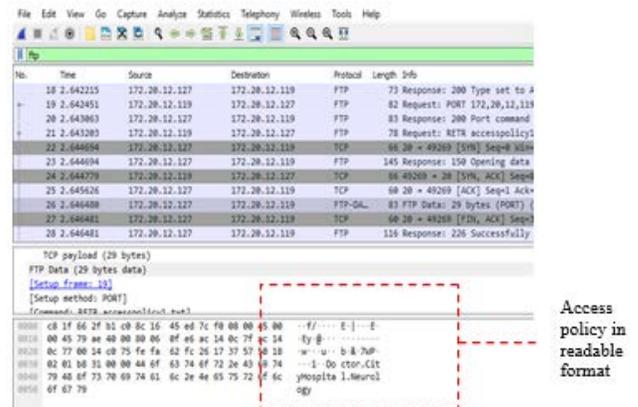


Fig. 5. Packet Sniffing using Wireshark.

In this simulation, because of the access policy was sent in a readable form, the attacker can read packet and gain the information in the access policy. Based on the simulation, it is necessary to provide a framework that can help user to secure their data than preserve the privacy of an access policy.

IV. ENHANCEMENT OF CIPHERTEXT POLICY ATTRIBUTE-BASED ENCRYPTION (CP-ABE) SCHEME

This section describes in detail on how the ciphertext policy attribute-based encryption (CP-ABE) scheme has been utilized to design the proposed framework; called as SecHC. The proposed framework provides security components in order to ensure the healthcare and medical data can securely exchange among healthcare organizations through Cloud environment. The SecHC framework (Fig. 6) have four entities involved are Data Owner, Data User, Attribute Authority and Secure Health Cloud.

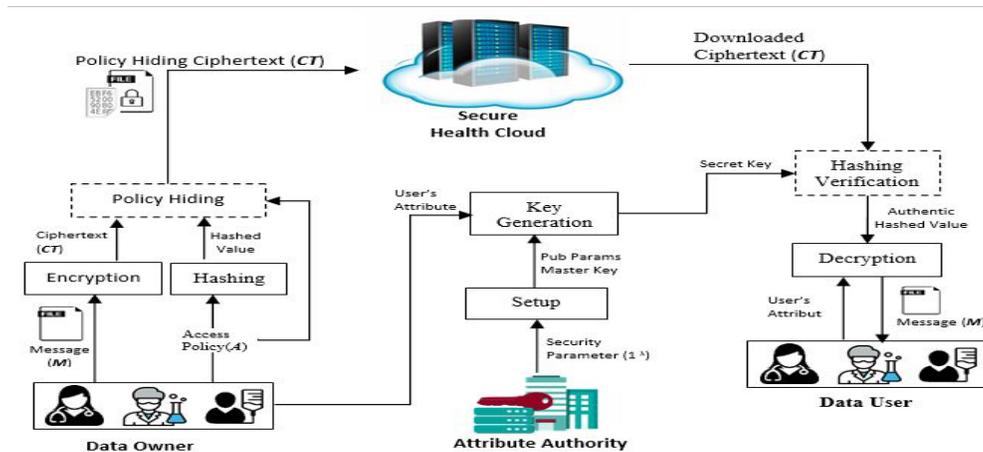


Fig. 6. Cloud-based Secure Healthcare Framework (SecH).

The general framework for CP-ABE scheme has four modules are (i) setup, (ii) key generation, (iii) encrypt and (iv) decryption. The data owner in CP-ABE scheme is been set to some access policy before it been encrypted. Data owners then outsourced the encrypted data along with access policy to the Cloud storage. If the receiver satisfied the defined access policy, the data then will be decrypted. Even though ciphertext attribute-based encryption is a prominent access control scheme however it suffers from privacy preservation of policy. It is because the defined policy is appended to the ciphertext in readable format which can lead to user's information leakage. So, the adversaries can learn confidential information from readable access policy. To address the issue, this work enhance Ciphertext Policy Attribute-Based Encryption (CP-ABE) scheme by adding other components to secure the healthcare data for designing the Secure Healthcare Framework (SecHC). The proposed framework provides security and integrity of data using encryption and hashing scheme. It also provides a policy hiding scheme to ensure access control with privacy-preserving.

Specifically, two more modules have been added in encryption phase namely hashing and policy hiding while for decryption phase, the added module is hashing verification. In the encryption phase, the hashing module is to ensure there is no tampering data while maintaining the data integrity. In the policy hiding module, the access policy will be hidden from the unauthorized users that might be led to privacy leakage. Meanwhile in the decryption phase the hashing verification is introduced which is to verify the authenticity of the access policy. This module aims to accelerate hashing verification thus reducing the time for decryption. Each of the functions is described as follows.

A. Setup Algorithm

The setup algorithm involves a security parameter as an input and produces a public parameter key and a master key as the outputs. Both will be utilized as the input in the key generation algorithm. In this setup algorithm, Bilinear Pairing will be used to create a public key which generated by Attribute Authority.

B. Key Generation

The key generation algorithm is executed by the Attribute Authority. It will take input users' attributes, public parameters, and master key then perform an operation to the inputs to generate a secret key. Later, this secret key will be employed by the data user to decrypt the ciphertext.

C. Hashing

This is the new module introduced in SecHC. The hashing algorithm used a mathematical function to generate a hash value from the input by the user. It used to verify the integrity of a file after it has been transferred from one place to another. In this module, the attributes of access policy will be hashed using the MD5 algorithm. According to [27], MD5 is very efficient and it operates faster than other algorithms. Then, after the hashing process, the hashed value will be sent to the Cloud along with ciphertext.

D. Encryption

In the encryption phase, the algorithm takes the inputs from the Data Owner to be encrypted and produce a ciphertext. Advanced Encryption Standard (AES) is used as an encryption algorithm to converts data into unreadable forms. The data owner will outsource the ciphertext along with the hashed value in the Cloud to be shared with the other users.

E. Access Policy Hiding

In traditional CP-ABE, the access policy is sent along with the ciphertext to the Cloud in readable form. Therefore, anyone who accesses the ciphertext can learn about the access policy which leads to weak policy privacy. Hence, to overcome this issue, an access policy hiding scheme is proposed and developed in SecHC to provide fully hidden access policy. In this module, the Logical Connective Operator is used to hide the attribute names and its values into meaningless values. Then, this meaningless value will be sent to Cloud along with ciphertext and hashed value.

F. Extracting Hidden Access Policy

Data Users need to extract the hidden access policy before he/she can access and download the file from Cloud. In this module, reverse process of hidden access policy from meaningless value to the access policy will be performed.

Then, after extracting the access policy, hashing verification will take place.

G. Hashing Verification

Hashing verification scheme is used to determine the authenticity of ciphertext and access policy. This module is to ensure the transferred ciphertext is not corrupted. In this module, the user needs to perform a calculation (hashing verification algorithm) to access policy. If the hash value of access policy (before and after downloaded from the Cloud) is the same, then the transferred file is an identical copy. So, the user can proceed to the decryption module. But if the values are not the same, the process will be ended.

H. Decryption

The decryption module is taking place after the hashing verification process produced the authentic hashing value. The decryption scheme takes a public parameter, secret key (associated with the user's attributes) and ciphertext as input and produces the message.

V. COMPARISON WITH OTHER FRAMEWORKS

In this section, a comparative analysis of the proposed scheme with some typical CP-ABE schemes has been conducted. By considering the characteristics of the proposed scheme satisfied, only selected schemes [21][26][28][29] that are strongly relevant to the proposed scheme for comparison is discussed.

Based on findings reported in Table I, a comprehensive comparison according to important features is presented including policy hiding, privacy-preserving, data integrity, and fine-grained access control.

By comparing the proposed framework with the schemes by the authors in [21][29] it definitely a different strategy is used for improving CP-ABE. In [21], the enhancement of CP-ABE scheme only achieves partly hiding policy which leads to failure to protect the privacy of the data. However, they accomplish to offer integrity using a hashing scheme. Yet, their scheme can be improved in terms of decryption procedure which focuses on the decryption test.

TABLE I. COMPARATIVE SUMMARY BETWEEN SCHEMES

Technique	Hidden Policy	Privacy-Preserving	Integrity	Fine-grained Access Control
CP-ABE [26]	Partial	√	X	√
Improvised CP-ABE [21]	Partial	X	√	√
CP-ABE [28]	No	√	X	√
Improvised CP-ABE [29]	Partial	X	√	√
SecHC CP-ABE scheme	Fully	√	√	√

Meanwhile, in [29], their CP-ABE scheme focused more on the decryption phase where they introduced KeyGen.out, Decrypt.out, Decrypt.user. Their scheme is designed to broadcast encryption techniques and used outsourcing techniques to realize policy-hide, direct revocation, and secure delegation simultaneously. However, their scheme failed to provide a fully hidden access policy and unable to keep it private.

Meanwhile, in SecHC framework, the proposed CP-ABE scheme capable of supporting fully hidden access policy using data hiding technique, provide integrity using hashing scheme and provision fine-grained access control. Most of the researchers have improvises the CP-ABE scheme except the authors in [26] and [28] which cause them unable to achieved hidden policy and provide data integrity.

VI. DATA SECURITY REQUIREMENT ANALYSIS OF SecHC FRAMEWORK

Particularly, in medical and healthcare organizations, sharing information in the Cloud environment has raised many security problems related to security requirements such as confidentiality, integrity, privacy and so on. Thus, the SecHC framework is developed to meet and satisfy the security requirement for the Cloud environment. The analysis of security requirements is described as follows.

A. Data Privacy

The proposed framework is fully privacy-preserving. It protects users' privacy by using the encryption algorithm. The proposed framework adopted the encryption algorithm in CP-ABE scheme to provide data privacy in a secure health Cloud. The electronic health record's privacy is achieved when the user uploads the encrypted record with hidden policy to the cloud.

B. Data Integrity

In the proposed framework, the hashing algorithm is used to protect the accuracy and consistency of data from any modification, deletion or fabrication. To achieve data accuracy, only correct and trustworthy data must be stored in the Cloud while consistency can be attained when outsourced data is not tampered, changed or maliciously deleted. Based on the framework, the medical and health record will be encrypted, and the access policy will be hashed. This hashing process is to ensure there is no modification that has been made on ciphertext stored in Cloud. If there is any modification on ciphertext, the user could not perform the decryption process. This process proves that the proposed framework offers to protect the integrity of data in a secure health Cloud.

C. Fine-Grained Access Control

In a healthcare Cloud, all users do not have the same privileges to retrieve medical data. This privilege depends on the degree to which a user is involved or specialized in treatment. Therefore, this framework ensures a different user will have different access privileges which defined by access policy imposed by attribute authority. The CP-ABE scheme used in this framework helps us achieve fine-grained access control. This means that all the attributes must be matched with

the user access policy structure to be able to access the required information.

VII. CONCLUSION

In the ever-increasing era of a data breach, Cloud computing is required to provide a security solution to protect sensitive information and transactions. This solution can prevent a third party from eavesdropping or tampering with the data while it is being transmitted. In this work, the Secure Healthcare Framework (SecHC) is designed by enhancing the Ciphertext Policy Attribute-Based Encryption (CP-ABE) scheme. It can further promote the Cloud privacy and integrity to both users and providers in the healthcare and medical organizations. The SecHC framework supports privacy and integrity of healthcare and medical data and offers fine-grained access control strategy. It is provided by using combination of prior and new components in the CP-ABE scheme. Such components offer a fully hidden access policy and provide fast decryption. Analysis of security requirements shows that this framework satisfies the privacy and integrity of healthcare data. In the near future, let the proposed framework to handle the real healthcare and medical data over the Cloud environment.

REFERENCES

- [1] O. Ali, A. Shrestha, and S. Fosso, International Journal of Information Management, "Cloud computing-enabled healthcare opportunities, issues , and applications : A systematic review," vol. 43, no. April, pp. 146–158, 2018.
- [2] F. Shiferaw and M. Zolfo, "The role of information communication technology (ICT) towards universal health coverage: the first steps of a telemedicine project in Ethiopia," Global health action, 5(1), 15638, no. June 2014, pp. 0–8, 2012.
- [3] J. J. Yang, J. Q. Li, and Y. Niu, "A hybrid solution for privacy preserving medical data sharing in the cloud environment," Futur. Gener. Comput. Syst., vol. 43–44, pp. 74–86, 2015.
- [4] A. Jemai, R. Attia, N. Kaaniche, S. Belguith, and M. Laurent, "PHOABE: Securely outsourcing multi-authority attribute-based encryption with policy hidden for cloud assisted IoT," Comput. Networks, vol. 133, pp. 141–156, 2018.
- [5] N. Sultan, "International Journal of Information Management Making use of cloud computing for healthcare provision: Opportunities and challenges," Int. J. Inf. Manage., vol. 34, no. 2, pp. 177–184, 2014.
- [6] N. Y. Lee and B. H. Wu, "Privacy Protection Technology and Access Control Mechanism for Medical Big Data," Proc. - 2017 6th IIAI Int. Congr. Adv. Appl. Informatics, IIAI-AAI 2017, pp. 424–429, 2017.
- [7] L. Ibraimi, M. Asim, and M. Petko, "Secure Management of Personal Health Records by Applying Attribute-Based Encryption," Proc. 6th Int. Work. Wearable, Micro, Nano Technol. Pers. Heal., pp. 71–74.
- [8] H. A. Al Hamid, S. M. M. Rahman, M. S. Hossain, A. Almogren, and A. Alamri, "A Security Model for Preserving the Privacy of Medical Big Data in a Healthcare Cloud Using a Fog Computing Facility with Pairing-Based Cryptography," IEEE Access, pp. 22313–22328, 2017.
- [9] T. Kajiyama, M. Jennex, and T. Addo, "To cloud or not to cloud: how risks and threats are affecting cloud adoption decisions," Inf. Comput. Secur., pp. 00–00, 2017.
- [10] M. Sookhak, F. R. Yu, M. K. Khan, Y. Xiang, and R. Buyya, "Attribute-based data access control in mobile cloud computing: Taxonomy and open issues," Futur. Gener. Comput. Syst., vol. 72, pp. 273–287, 2017.
- [11] Satar, S.D., Hussin, M., Hanapi, Z., & Mohamed, M.A. (2018). Data Privacy and Integrity Issues Scheme in Cloud Computing: A Survey. International journal of engineering and technology, 7, 102.
- [12] F. Xhafa, J. Li, G. Zhao, J. Li, X. Chen, and D. S. Wong, "Designing cloud-based electronic health record system with attribute-based encryption," Multimedia Tools and Applications, 74(10), 3441–3458.2014.
- [13] K. Yang, Q. Han, H. Li, K. Zheng, Z. Su, and X. Shen, "An Efficient and Fine-Grained Big Data Access Control Scheme with Privacy-Preserving Policy," IEEE Internet Things J., vol. 4, no. 2, pp. 563–571, 2017.
- [14] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 1, pp. 131–143, 2013.
- [15] S. Sabitha and M. S. Rajasree, "Access control based privacy preserving secure data sharing with hidden access policies in cloud," J. Syst. Archit., vol. 75, pp. 50–58, 2017.
- [16] H. Wang, X. Dong, and Z. Cao, "Multi-value-Independent Ciphertext-Policy Attribute Based Encryption with Fast Keyword Search," IEEE Transactions on Services Computing vol. 1374, no. c, 2017.
- [17] L. Zhang, Y. Cui, Y. Mu, and S. Member, "Improving Security and Privacy Attribute Based Data Sharing in Cloud Computing," IEEE Systems Journa, pp. 1–11, 2019.
- [18] Q. Han, Y. Zhang, and H. Li, "Efficient and robust attribute-based encryption supporting access policy hiding in Internet of Things," Futur. Gener. Comput. Syst., vol. 83, pp. 269–277, 2018.
- [19] Abd Hamid, N., Ahmad, R. and Selamat, S.R., 2017. Recent Trends in Role Mining Algorithms for Role-Based Access Control: A Systematic Review. World Applied Sciences Journal, 35(7), pp.1054-1058.
- [20] F. Deng, Y. Wang, L. I. Peng, H. U. Xiong, and Z. Qin, "Ciphertext-Policy Attribute-Based Signcryption With Verifiable Outsourced Designcryption for Sharing Personal Health Records," IEEE Access, vol. 6, pp. 39473–39486, 2018.
- [21] L. Zhang, G. Hu, Y. Mu, and F. Rezaeibagha, "Hidden Ciphertext Policy Attribute-Based Encryption with Fast Decryption for Personal Health Record System," IEEE Access, vol. 3536, no. c, pp. 1–1, 2019.
- [22] D. Slamanig and C. Stingsl, "Privacy Aspects of eHealth," pp. 1228–1235, 2008.
- [23] Z. Ying, L. U. Wei, Q. I. Li, X. Liu, and J. I. E. Cui, "A Lightweight Policy Preserving EHR Sharing Scheme in the Cloud," vol. 6, 2018.
- [24] N. Muhammad, J. M. Zain, and M. Mohamad, "Current Issues in Ciphertext Policy-Attribute Based Scheme for Cloud Computing : A Survey," International Journal of Engineering & Technology, vol. 7, pp. 64–67, 2018.
- [25] R. Zhang and L. Liu, "Security models and requirements for healthcare application clouds," Proc. - 2010 IEEE 3rd Int. Conf. Cloud Comput. CLOUD 2010, pp. 268–275, 2010.
- [26] Y. Zhang, D. Zheng, and R. H. Deng, "Security and Privacy in Smart Health : Efficient Access Control," IEEE Internet Things J., vol. 5, no. 3, pp. 2130–2145, 2018.
- [27] Rachmawati, D., Tarigan, J. T., & Ginting, A. B. C. (2018, March). A comparative study of Message Digest 5 (MD5) and SHA256 algorithm. In Journal of Physics: Conference Series (Vol. 978, No. 1, p. 012116).
- [28] S. Sharaf and N. F. Shilbayeh, "A Secure G-Cloud-Based Framework for Government Healthcare Services," IEEE Access, vol. 7, pp. 37876–37882, 2019.
- [29] H. Xiong, Y. Zhao, L. Peng, H. Zhang, and K. Yeh, "Partially policy-hidden attribute-based broadcast encryption with secure delegation in edge computing," Futur. Gener. Comput. Syst., vol. 97, pp. 453–461, 2019.

Determining Optimal Number of K for e-Learning Groups Clustered using K-Medoid

S. Anthony Philomen Raj¹
Research Scholar
Department of Computer Science
Periyar University
Tamil Nadu
India

Vidyaathulasiraman²
Assistant Professor
Department of Computer Science
Government Arts and Science College for women
Tamil Nadu
India

Abstract—e-Learning is appropriate when the learners are grouped and facilitated to learn according to their learning style and at their own pace. Elaborate researches have been proposed to categorize learners based on various e-learning parameters. Most of these researches have deployed the clustering principles for grouping eLearners, and in particular, they have utilized K-Medoid principle for better clustering. In the classical K-Medoid algorithm, predicting or determining the value of K is critical, two methods namely the Elbow and Silhouette methods are widely applied. In this paper, we experiment with the application of both these methods to determine the value of K for clustering eLearners in K-Medoid and prove that Silhouette method best predicts the value of K.

Keywords—Clustering; e-learning; elbow method; k-means; k-medoid; machine learning; silhouette method

I. INTRODUCTION

The educational systems nowadays are slightly moved from Traditional Teaching Method to Electronic Teaching Method. There are a variety of tasks that can be performed in e-learning, such as; assignments, quizzes, and so on. These activities are used to assess the learner's performance. To facilitate appropriate e-learning activities by grouping users into a possible number of groups, Clustering is the Machine Learning (ML) technique that is used to group the related objects. There are numerous existing methods available for cluster analysis in the field of Data Analytics. Determining the optimal number of clusters in a data set is a fundamental problem in partitioning clustering, such as K-means clustering, which allows the user to define the number of clusters K to be generated. The possible number of clusters is rather arbitrary and is determined by the method used for measuring similarities and the parameters used for partitioning [1]. There are many clustering algorithms used for the group the similar objects in many domains such as the Medical domain, Education domain, Governance domain, etc. The clustering algorithm is the most suitable one to group users based on the learners preferred learning activities in e-learning.

The existing methods mainly focus on the majority of learning activities based on the learner's style. This could be improved further by grouping the users based on their learning activities. The main objective of this paper is to identify the optimal number of groups by using cluster validation methods. If we identify the possible number of groups of learners, we can easily enhance their learning abilities according to their preferred learning activities.

The flow of organization of work is as follows: This paper introduces a different method for identifying the value of K. Then it elaborates two major method such as Elbow and Silhouette method. The paper experiments with data using both methods. It further denotes the best method for identifying K values in K-means along with the eLearners.

A. Choosing the Optimum Number of Cluster

The optimum number of clusters obtained by using the following two methods such as *Elbow* and *Silhouette methods* [2][3].

1) *Elbow method*: The number of clusters (K) in the Elbow method ranges from 1 to n. We calculate WCSS (Within-Cluster Sum of Square) for each value of K. In a cluster, WCSS is the number of squared distances between each point and the centroid. The plot looks like an Elbow when we plot the WCSS with the K meaning. The WCSS value will begin to decrease as the number of clusters grows. When K = 1, the WCSS value is the highest. When we examine the graph, we can see that it will shift rapidly at a point, forming an elbow shape. The graph begins to travel almost in the same direction as the X-axis [2][3]. The optimal K value or the optimum number of clusters corresponds to this point.

Algorithm

Step 1: Compute clustering algorithm (e.g., K-means clustering) for different values of K. For instance, by varying K from 1 to N clusters.

Step 2: For each K, calculate the total Within-Cluster Sum of Square (WCSS).

$$\sum_{k=1}^K \sum_{i \in S_k} \sum_{j=1}^p (x_{ij} - \bar{x}_{kj})^2 \quad (1)$$

Where S_k is the set of observations in the K^{th} cluster and \bar{x}_{kj} is the j^{th} variable of the cluster center for the K^{th} cluster.

Step 3: Plot the curve of WCSS according to the number of clusters K.

Step 4: The location of a bend (knee) in the plot is generally considered as an indicator of the appropriate number of clusters.

2) *Silhouette method:* This method calculates the similarity of an object to its own cluster called *cohesion*, when compared to other clusters is called *separation*. The Silhouette value, which is a value in the range [-1, 1], is the comparison's means; a value close to 1 indicates a close relationship with objects in its cluster, while a value close to -1 indicates the opposite [2][3]. A model that produces mostly high Silhouette values from a clustered collection of data is most likely acceptable and reasonable.

Algorithm

Step 1: Choose the K from 1 to n clusters.

Step 2: For each k, calculate the Silhouette value.

Let $C(i)$, be the cluster to which the i^{th} data point has been allocated.

Let $|C(i)|$, be the number of data points allocated to the i^{th} data point in the cluster.

Let $a(i)$, indicates how well the i^{th} data point is allocated to its cluster.

$$a(i) = \frac{1}{|C(i)|-1} \sum_{C(i), i \neq j} d(i, j) \quad (2)$$

Let $b(i)$, be the average dissimilarity to the cluster nearest to it, but not its own

$$b(i) = \min_{i \neq j} \left(\frac{1}{|C(i)|} \sum_{j \in C(i)} d(i, j) \right) \quad (3)$$

The Silhouette Coefficient $S(i)$ is given by:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4)$$

Step 3: Plot the curve of Silhouette value according to the number of clusters K.

Step 4: The location of the highest point is taken as the suitable number of clusters.

B. Flow Chart of Elbow and Silhouette Method

Fig. 1 portrays the flow chart of *Elbow* and *Silhouette* method to identify the optimum number of clusters.

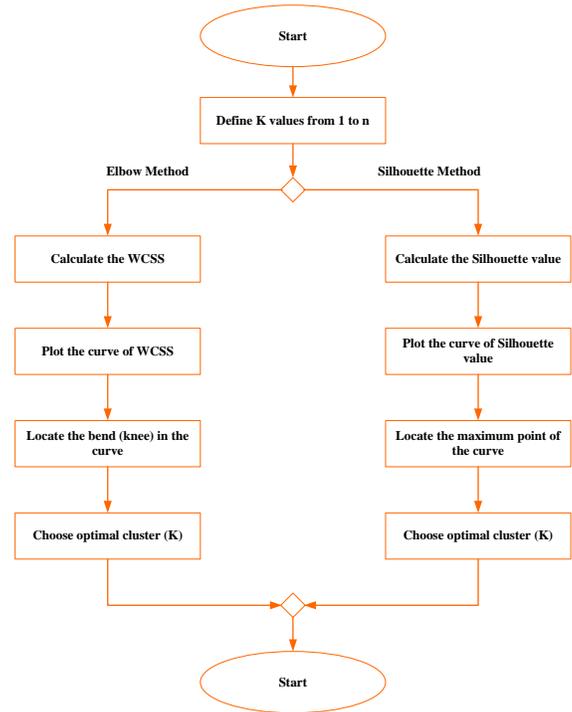


Fig. 1. Flow Chart of Elbow and Silhouette Method.

C. Comparison of Elbow and Silhouette Method

Table I contrasts the comparison between *Elbow* and *Silhouette* methods.

TABLE I. COMPARISON BETWEEN ELBOW AND SILHOUETTE METHOD

Elbow method	Silhouette method
It is more of a criterion for making decisions.	The Silhouette is a validation metric used in clustering.
The WCSS is a metric for clustering compactness, and it should be as low as possible.	This approach is useful for determining the consistency of clustering, or how well an object fits into its cluster.
The Elbow method is not computationally demanding.	The Silhouette method is the most computationally demanding.
Sometimes we don't get a clear elbow point on the plot, in such cases it's very hard to finalize the size of the cluster.	Based on the Silhouette we can identify the cluster size.

II. RELATED WORKS

In cluster analysis, especially in the field of Data Analytics determining the optimum cluster number is a major challenge. Many methods are used to find optimal clusters among Elbow and Silhouette methods which are frequently used.

H Humaira et al. [4] proposed the method of the identifying size of the clusters using the Elbow method for the K-Means Algorithm. However, this approach lacks comparison with another method called Silhouette.

Integration K-Means Clustering method and Elbow method for identification of the best Customer profile cluster were suggested by M A Syakur et al. [5]. This approach is used to determine the best number of clusters with the elbow method and it will be the default for characteristic process based on the case study.

Mohammad Khalil et al. [6] applied a Clustering pattern of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to identify student groups. This research is used to predict the learners' engagement in MOOCs by choosing optimal clusters.

Brahim Hmedna et al. [7] proposed a method, how does a learner prefer to process information in MOOCs? Using the K-Means clustering algorithm, this study discovered that the majority of learners favour active learning styles.

Towards an optimal personalization strategy, MOOCs were suggested by Alaa A.Qaffas et al. [8]. Using K-Means clustering, this approach was used to increase the retention rate and quality of learning in MOOCs.

MOOC Video Personalized Classification Based on Cluster Analysis and Process Mining was suggested by Feng Zhang et al. [9]. They suggest a process model for a group of students that represent the students' overall video-watching behaviour. Then, based on the video watching data of the students involved, it suggested using the process mining technique to mine the process model of each student cluster. Finally, the method is used to measure the difficulty and importance of a video based on a process model.

Mohammad KHALIL et al. [10] introduced a Portraying MOOCs Learners: a Clustering Experience Using Learning Analytics. The study used Clustering Analysis to group the students into suitable profiles based on their participation in a university-mandated MOOC that was also accessible to the public.

Delali Kwasi Dake et al. [11] applied a K-means clustering algorithm to analyze students' clusters for centered project-based learning. K clusters of 20 are used in this study. The findings show that the K-means clustering algorithm is good at grouping learners based on similar performance characteristics.

Analysis of University Students' Behavior Based on a Fusion K-Means Clustering Algorithm was proposed by Wenbing Chang et al. [12]. They proposed a new algorithm based on K-means and clustering by quick search and find the density peaks (K-CFSFDP), which improves data point distance and density.

Abdallah Moubayed et al. [13] proposed a model for Student Engagement Level in an e-learning Environment: Clustering Using K-means. This study recommends that students be clustered using the K-means algorithm based on 12 engagement metrics divided into two categories: interaction-related and effort-related.

Using Self-Organizing Map and Clustering to Investigate Problem-Solving Patterns in the Massive Open Online Course: An Exploratory Study proposed by Youngjin Lee et al. [14]. This study suggests that combining self-organizing map and

hierarchical clustering algorithms in a clustering technique can be a useful exploratory data analysis method for MOOC instructors to classify related students based on a large number of variables and analyse their characteristics from multiple perspectives.

Prerna Joshi et al. [15] proposed a model for Prediction of Students Academic Performance Using K-Means and K-Medoids Unsupervised Machine Learning Clustering Technique. The K-mean and K-Medoids grouping algorithms were used in this study to examine students' consequence information.

Yaminee S. Patil et al. [16] suggested a technique K-means Clustering with Map Reduce Technique. This research article identified the implementation of the K-Means Clustering Algorithm over a distributed environment using Apache Hadoop.

Xin Lu et al. [17] proposed a method Improved K-means Distributed Clustering Algorithm based on Spark Parallel Computing Framework. This research identified, a density based initial clustering center selection method proposed to improve the K-means distributed clustering algorithm.

Literature review reveals that the authors have mostly focused on evaluating learner's performance by clustering techniques based on their learning styles, learning activities, and e-learning tools. The existing approaches lack in choosing an optimum number of cluster size K to group the learners. Hence, this research proposes an algorithm to identify optimum numbers of cluster size K to group the learners based on their preferred e-learning activities.

III. RESEARCH OBJECTIVES

- 1) To identify learners preferred e-learning activities using PCA.
- 2) To find correlation coefficient for selected e-learning activities using Pearson Correlation.
- 3) To identify an optimal number of clusters using Elbow and Silhouette method.
- 4) To select best method for choosing optimal cluster size.
- 5) To group the learners based on their preferred e-learning activities with optimal cluster size.

IV. PROPOSED ARCHITECTURE OF CHOOSING OPTIMAL CLUSTER

The purpose of finding optimal number cluster is to extract the most possible number of groups with learner's preferred e-learning activities. The course teacher can implement those selected activities to learners based on the clusters, which helps the learners to enhance their learning abilities. Fig. 2 portrays the process to find an optimal number of groups with the help of *Elbow* and *Silhouette* method. It has been classified in the following stages:

Stage 1: Identify learners preferred e-learning activities.

Stage 2: Apply cluster validation to fix the cluster size.

Stage 3: Identify the possible number of clusters using Elbow and Silhouette method.

Stage 4: Select the method which is suitable to group the learners based on their preferred learning activities.

Stage 5: List the optimal cluster to group the learners.

Stage 1: Identify learner's preferred e-learning activities: In the first stage, identify the learners' preferred e-learning activities by using Principal Component Analysis (PCA) and compute the correlation coefficient using Pearson Correlation.

Stage 2: Apply cluster validation to fix the cluster size: In the second stage, validate cluster size by using appropriate cluster validation methods. Identifying possible number cluster size is a big challenge.

Stage 3: Identify the possible number of clusters using Elbow and Silhouette method: These two methods are used to identify the size of the cluster known as Elbow and Silhouette method. In the Elbow method by calculation value of WCSS the cluster size is fixed. Similarly, in the Silhouette by calculation of Silhouette Value the cluster size is fixed.

Stage 4: Select the method which is suitable to group the learners based on their preferred learning activities: Apply the data set in both methods and list the possible number of clusters. Let K1 be the number of cluster sizes identified by the Elbow method. Let K2 be the number of cluster sizes identified by the Silhouette method. Choose a suitable method by comparing both the cluster size (K1 and K2) and give the highest priority by choosing cluster size which one is bigger (either K1 or K2).

Stage 5: List the optimal cluster to groups the learners: Finally, the most possible cluster size (Kn) is identified by using Stage 4. Use this cluster size to identify the possible number of learners groups according to their preferred e-learning activities.

V. PROPOSED ALGORITHM OF IDENTIFICATION OF CLUSTER SIZE USING ELBOW AND SILHOUETTE METHOD

Input: S : Data Set

Output: K : Number of cluster size

1: Let S be the given set of preferred e-learning activities.

$$S = A_{ij} \quad (5)$$

2: Compute the sum of activities based on learner wise.

$$S = \sum_{i=1}^n \sum_{j=1}^m A_{ij} \quad (6)$$

3: Apply the cluster validation methods

Elbow method:

$$\sum_{k=1}^K \sum_{i \in S_k} \sum_{j=1}^p (x_{ij} - \bar{x}_{kj})^2 \quad (7)$$

Where S_k is the set of observations in the K^{th} cluster and \bar{x}_{kj} is the j^{th} variable of the cluster center for the K^{th} cluster

Silhouette method:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (8)$$

4: Choose an appropriate method to fix cluster size

$$K = \begin{cases} \text{Choose Elbow Method if } K1 > K2 \\ \text{Choose Silhouette Method if } K2 > K1 \end{cases} \quad (9)$$

5: List the optimal cluster size

6: End

VI. RESULTS AND DISCUSSION

The optimal number of cluster size was implemented with the following e-learning activities such as Continuous Assessment (CA), Assignment, Test, Practical, Seminar, and Course Work.

Step 1: Preferred e-learning activities

Table II listed the preferred e-learning activities of 70 users and their performances. These preferred e-learning activities are identified through PCA and compared with the Pearson Correlation to find a correlation coefficient with each attribute.

Step 2: Compute the sum of activities based on learner wise.

Compute the sum of selected activities for each learner. Table III listed the sum of e-learning activities of 70 users and their performances.

Step 3: Apply the cluster validation methods

Graph 1 portrays dataset of 70 users and their performance with their preferred e-learning activities.

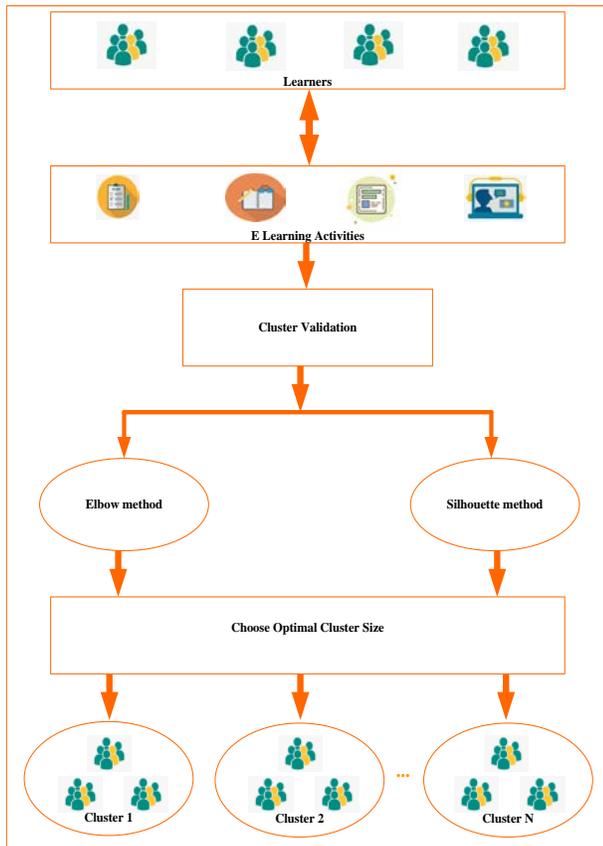


Fig. 2. Architecture of Choosing Optimal Cluster.

TABLE II. PREFERRED E-LEARNING ACTIVITIES OF 70 USERS. A1- CONTINUOUS ASSESSMENT (CA), A2 - ASSIGNMENT, A3 - TEST, A4 - PRACTICAL, A5 - SEMINAR, A6 - COURSE WORK.

Users/ Activities	A1 (20)	A2 (5)	A3 (10)	A4 (20)	A5 (10)	A6 (10)
BP191001L	14.5	4.5	7.8	13	8	8
BP191002L	9	3.5	8.6	12	7	7
BP191003L	3.1	2.5	4.9	10	10	8
BP191004L	4.5	3.3	6.2	8	8	8
BP191005L	7.5	4	6.4	13	8	8
BP191006L	7.5	3.8	5.9	11	8	8
BP191007L	9.3	4.2	7.3	20	10	8
BP191008L	12.3	4.5	7.4	20	10	8
BP191009L	13.2	4.3	8.6	14	7	7
BP191010L	8.3	3.3	5.8	8	10	8
BP191011L	13	3.5	4.4	20	10	8
BP191012L	11.1	3.4	6.4	20	10	8
BP191013L	12.5	3.9	7.9	14	7	7
BP191014L	11.1	3.4	9.2	11	7	7
BP191015L	8.9	4	8.4	20	10	8
BP191016L	8	3	7.6	15	7	8
BP191017L	12	4	8.2	15	7	8
BP191018L	6.5	2.9	7.4	5	7	7
BP191019L	13.2	3.6	8.9	20	10	8
BP191020L	5.8	3.2	7.1	15	8	8
BP191021L	6.5	3.8	7.3	20	10	8
BP191023L	12	4	7.4	10	10	8
BP191024L	9	4.3	6.4	10	10	8
BP191025L	13.5	4.4	7.9	15	8	8
BP191026L	9.6	3.4	6.9	15	8	8
BP191027L	10.5	3.5	7.9	20	10	8
BP191028L	7	3	8.3	20	10	8
BP191029L	12.5	4	7.6	15	8	8
BP191030L	10.2	3	6.6	8	7	7
BP191031L	8.8	2.8	6.9	10	10	8
BP191032L	6.8	2.8	6.8	8	8	8
BP191033L	6.5	2.5	8.6	13	7	7
BP191034L	13	3.3	7.1	14	7	7
BP191035L	4.2	2	6.4	11	7	7
BP191036L	11.3	4	7.8	14	7	7
BP191037L	5.3	2.5	4.6	14	7	7
BP191038L	14	3.5	8.6	8	7	7
BP191039L	10.8	3	6.6	14	7	7
BP191040L	8.8	2.9	8.2	20	10	8
BP191041L	6.5	2.5	4.9	8	7	7
BP191042L	10.9	3.3	6.7	15	7	7

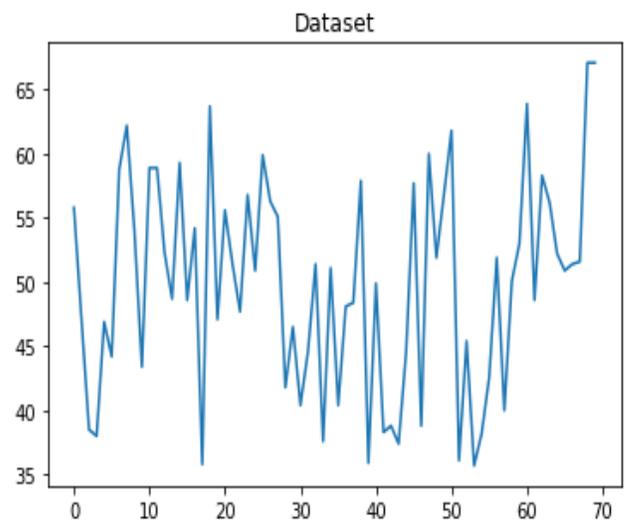
BP191043L	4.5	2.5	6.3	11	7	7
BP191044L	5.5	2.6	6.7	10	7	7
BP191045L	7.5	3	4.9	8	7	7
BP191046L	8.3	3.4	7.9	11	7	7
BP191047L	7.8	3.5	8.4	20	10	8
BP191049L	6.6	2.5	5.7	10	7	7
BP191050L	10.1	3.3	8.6	20	10	8
BP191051L	11.8	3.5	9.6	13	7	7
BP191052L	8	3.5	7.3	20	10	8
BP191053L	11	4	8.8	20	10	8
BP191054L	6	2.5	5.6	8	7	7
BP191055L	9.1	3.1	6.2	13	7	7
BP191056L	6.5	2.4	4.8	8	7	7
BP191057L	6.5	2.5	7.2	8	7	7
BP191058L	9.5	3.5	7.6	8	7	7
BP191059L	12.5	4	7.4	14	7	7
BP191060L	7.3	2.3	5.4	11	7	7
BP191061L	7.8	2.5	5.8	18	8	8
BP191001	8	5	8	16	8	8
BP191002	14.9	5	8	18	9	9
BP191003	5.6	5	6	16	8	8
BP191004	9.3	5	8	18	9	9
BP191005	7.2	5	8	18	9	9
BP191006	7.2	5	8	16	8	8
BP191007	5.9	5	8	16	8	8
BP191008	6.4	5	8	16	8	8
BP191009	6.6	5	8	16	8	8
BP191010	12.1	5	10	20	10	10
BP191011	12.1	5	10	20	10	10

TABLE III. SUM OF E-LEARNING ACTIVITIES OF 70 USERS

Users	Total Score
BP191001L	55.8
BP191002L	47.1
BP191003L	38.5
BP191004L	38
BP191005L	46.9
BP191006L	44.2
BP191007L	58.8
BP191008L	62.2
BP191009L	54.1
BP191010L	43.4
BP191011L	58.9
BP191012L	58.9

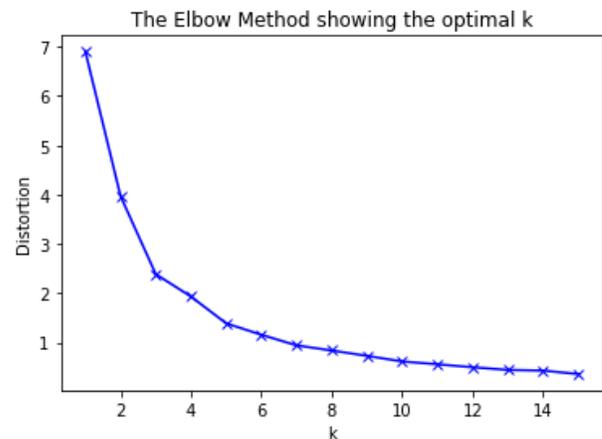
BP191013L	52.3
BP191014L	48.7
BP191015L	59.3
BP191016L	48.6
BP191017L	54.2
BP191018L	35.8
BP191019L	63.7
BP191020L	47.1
BP191021L	55.6
BP191023L	51.4
BP191024L	47.7
BP191025L	56.8
BP191026L	50.9
BP191027L	59.9
BP191028L	56.3
BP191029L	55.1
BP191030L	41.8
BP191031L	46.5
BP191032L	40.4
BP191033L	44.6
BP191034L	51.4
BP191035L	37.6
BP191036L	51.1
BP191037L	40.4
BP191038L	48.1
BP191039L	48.4
BP191040L	57.9
BP191041L	35.9
BP191042L	49.9
BP191043L	38.3
BP191044L	38.8
BP191045L	37.4
BP191046L	44.6
BP191047L	57.7
BP191049L	38.8
BP191050L	60
BP191051L	51.9
BP191052L	56.8
BP191053L	61.8
BP191054L	36.1
BP191055L	45.4
BP191056L	35.7
BP191057L	38.2
BP191058L	42.6
BP191059L	51.9

BP191060L	40
BP191061L	50.1
BP191001	53
BP191002	63.9
BP191003	48.6
BP191004	58.3
BP191005	56.2
BP191006	52.2
BP191007	50.9
BP191008	51.4
BP191009	51.6
BP191010	67.1
BP191011	67.1



Graph 1. Dataset of 70 users and their Performance.

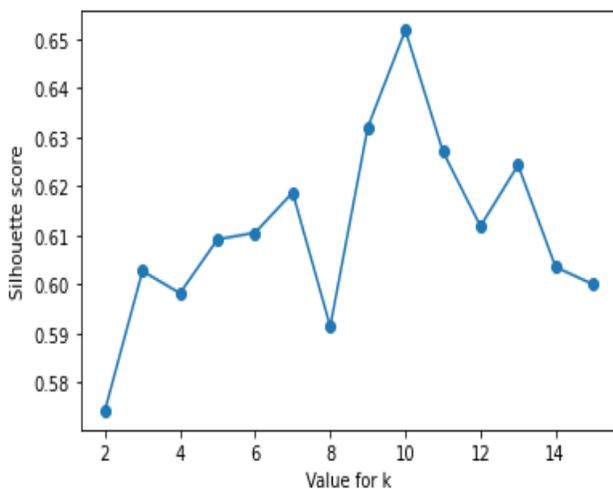
Elbow method – Graph 2 portrays optimal numbers of cluster size identified by using Elbow method. The graph shows the possible optimal number of cluster $K = 5$.



Graph 2. Elbow Method: Optimal Number of Cluster.

Silhouette method – Graph 3 portrays optimal numbers of cluster size identified by using Silhouette method. The graph shows the possible optimal number of cluster $K = 10$. The Silhouette Score for cluster size $K = 2$ to 15 are as follows:

- Silhouette Score for 2 Clusters: 0.5742
- Silhouette Score for 3 Clusters: 0.6004
- Silhouette Score for 4 Clusters: 0.5982
- Silhouette Score for 5 Clusters: 0.6091
- Silhouette Score for 6 Clusters: 0.6034
- Silhouette Score for 7 Clusters: 0.5960
- Silhouette Score for 8 Clusters: 0.6094
- Silhouette Score for 9 Clusters: 0.6185
- Silhouette Score for 10 Clusters: 0.6518
- Silhouette Score for 11 Clusters: 0.6248
- Silhouette Score for 12 Clusters: 0.5981
- Silhouette Score for 13 Clusters: 0.5995
- Silhouette Score for 14 Clusters: 0.6061
- Silhouette Score for 15 Clusters: 0.6047



Graph 3. Silhouette Method: Optimal Number of Cluster.

Step 4: Choose appropriate method to fix cluster size

From Step 3, the appropriate method for validating cluster size for given data set is Silhouette method. This can be achieved by comparing cluster size of both the methods.

Step 5: List the optimal cluster size

The most possible number of cluster size is $K = 10$. The possible number of learners' groups is also 10, according to their preferred e-learning activities.

The paper experiments data with two cluster validation methods such as Elbow and Silhouette method. These two methods are frequently used to validate cluster size. The cluster size return by Elbow method for the given data set is 5 ($K=5$). The cluster size return by Silhouette method for the given data set is 10 ($K=10$). Naturally when the cluster size increases, the learning abilities of the each learner is identified in depth. Based on their learning abilities, we can optimize and predict e-learning activities for each user. Finally this paper concludes that Silhouette method is the optimal method for validating cluster size for the given data set.

VII. CONCLUSION AND FUTURE WORK

In this paper, we have illustrated that determining optimal value for K to cluster eLearners using K-Medoid algorithm is essential. Further, we have experimented the Elbow and Silhouette methods to compute the optimal value for K . To decide the suitable method among these two, sufficient experiments were conducted, and the results of the experiments were investigated carefully. The results were indicative that the Silhouette method best suits to fix the optimal value for K . This paper is focused in estimating K value for K-Medoid based clustering of eLearners, computing the optimum value for k , number of clusters to be formed can also be done in other clustering methods which are applied in eLearners groupification, which is suggested by the authors as a future enhancement.

REFERENCES

- [1] Channamma Patil and Ishwar Baidari, "Estimating the Optimal Number of Clusters k in a Dataset Using Data Depth", Data Science and Engineering, vol 4, pp. 132–140, June 2019.
- [2] Sukavanan Nanjundan, Shreevinkesh Sankaran, C.R. Arjun and G. Paavai Anand, "Identifying the number of clusters for K-Means: A hypersphere density based approach", International Conference on Computers, Communication and Signal Processing, December 2019.
- [3] Chunhui Yuan and Haitao Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm", Multidisciplinary Scientific Journal, vol 2, pp. 226-235, June 2019.
- [4] H Humaira and R Rasyidah, "Determining The Appropriate Cluster Number Using Elbow Method for K-Means Algorithm", WMA-2, January 2020.
- [5] M A Syakur, B K Khotimah, E M S Rochman and B D Satoto, "Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster". IOP Conference Series: Materials Science and Engineering, vol 336, 2018.
- [6] Mohammad Khalil and Martin Ebner, "Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories", Journal of Computing in Higher Education, vol 29, pp.114–132, October 2016.
- [7] Brahim Hmedna, Ali El Mezouary and OmarBaz, "How Does Learners' Prefer to Process Information in MOOCs? A Data-driven Study", Procedia Computer Science, vol 148, pp. 371-379, 2019.
- [8] Alaa A.Qaffas, Kaouther Kaabi, Rustam Shadiey and Fathi Essalmi, "Towards an optimal personalization strategy in MOOCs", Smart Learning Environments, pp. 7-14, April 2020.
- [9] Feng Zhang, Di Liu and Cong Liu, "MOOC Video Personalized Classification Based on Cluster Analysis and Process Mining", Sustainability, vol 12, 2020.
- [10] Mohammad KHALIL, Christian KASTL and Martin EBNER, "Portraying MOOCs Learners: a Clustering Experience Using Learning Analytics", Proceedings of the European MOOC Stakeholder Summit, 2016.
- [11] Delali Kwasi Dake and Esther Gyimah, "Using K-Means to Determine Learner Typologies for Project-based Learning: A Case Study of the University of Education, Winneba", International Journal of Computer Applications (0975 – 8887), vol 178, pp. 29-34, August 2019.
- [12] Wenbing Chang, Xinpeng Ji, Yinglai Liu, Yiyong Xiao, Bang Chen, Houxiang Liu and Shenghan Zhou, "Analysis of University Students' Behavior Based on a Fusion K-Means Clustering Algorithm", Applied Sciences, vol 10, September 2020.
- [13] Abdallah Moubayed, Mohammadnoor Injadat, Abdallah Shami and Hanan Lutfiyya, "Student Engagement Level in an eLearning Environment: Clustering Using K-means", American Journal of Distance Education, vol 34, pp. 137-156, 2020.
- [14] Youngjin Lee, "Using Self-Organizing Map and Clustering to Investigate Problem-Solving Patterns in the Massive Open Online

- Course: An Exploratory Study”, *Journal of Educational Computing Research*, vol 57, pp. 471-490, January 2018.
- [15] Prerna Joshi and Pritesh Jain, “Prediction of Students Academic Performance Using K-Means and K-Medoids Unsupervised Machine Learning Clustering Technique”, *International Journal of Scientific Development and Research (IJS DR)*, vol 3, pp. 162-171, June 2018.
- [16] Yaminee S. Patil and M. B. Vaidya, “K-means Clustering with MapReduce Technique”, *International Journal of Advanced Research in Computer and Communication Engineering*, vol 4, pp. 349-352, November 2015.
- [17] Xin Lu, Huanghuang Lu, Jiao Yuan and Xun Wang, “An Improved K-means Distributed Clustering Algorithm Based on Spark Parallel Computing Framework”, *Journal of Physics: Conference Series*, vol 1616, 2020.

Abnormal Pulmonary Sounds Classification Algorithm using Convolutional Networks

Alva Mantari Alicia¹, Arancibia-Garcia Alexander²
Ramos-Cosi Sebastian⁶

Image Processing Research Laboratory (INTI-Lab)
Universidad de Ciencias y Humanidades, Lima, Perú

Chávez Frías William³, Cieza-Terrones Michael⁴
Herrera-Arana Víctor⁵

Hospital Nacional Cayetano Heredia
Lima, Perú

Abstract—In the world and in Peru, Acute Respiratory Infections are the main cause of death, especially in the most vulnerable population, children under 5 years of age and older adults. Pneumonia is the leading cause of death of children in the world. 60.2% of pneumonia cases affect children under 5 years of age. Thus, prevention and timely treatment of lung diseases are crucial to reduce infant mortality in Peru. Among the main problems associated with this high is percentage the lack of medical professionals and resources, especially in remote areas, such as Puno, Huancavelica and Arequipa, which experience temperatures as low as -20°C during the cold season. This study develops an algorithm based on computational neural networks to differentiate between normal and abnormal lung sounds. The initial base of 917 sounds was used, through a process of data augmentation, this base was increased to 8253 sounds in total, and this process was carried out due to the need of a large number of data for the use of computational neural networks. From each signal, features were extracted using three methods: MFCC, Melspectrogram and STFT. Three models were generated, the first one to classify normal and abnormal, which obtained a training Accuracy of 1 and a testing accuracy of 0.998. The second one classifies normal sound, pneumonia and other abnormalities and obtained training Accuracy values of 0.9959 and a testing accuracy of 0.9885. Finally, we classified by specific ailment where we obtained a training Accuracy of 0.9967 and a testing accuracy of 0.9909. This research provides interesting findings about the diagnosis and classification of lung sounds automatically using convolutional neural networks, which is the beginning for the development of a platform to assess the risk of pneumonia in the first moment, thus allowing rapid care and referral that seeks to reduce mortality associated mainly with pneumonia.

Keywords—Algorithm; classification; computational neural networks; lung sounds; mortality; pneumonia

I. INTRODUCTION

Pulmonary diseases in the new pandemic context are a public health concern. Pneumonia, for example, is evolving rapidly and its complications threaten the lives of the population, especially in low-income countries such as Peru[1],[2]. The Ministry of Health recognizes Acute Respiratory Infections (ARI) as a constant concern in Peru, caused by viruses, bacteria and fungi. Among ARI, pneumonia is consistently the leading cause of death in children and older adults in the world. In Peru, 60.2% of pneumonia cases occur in the 0-5 age group, and it is the leading cause of death[3].

Prevention of pneumonia complications is a priority to reduce mortality in children, especially in the most remote areas of the country.

Peru is a country with a diversity of weather and regions. There are more than 30 cities, with altitudes ranging from 2000 to 5100 meters above sea level[4]. Populated centers that endure extreme temperatures, even below -20 degrees Celsius, in departments such as Puno, Huancavelica, Arequipa, Junin, Pasco and Cuzco, among others[5]. In cold weather seasons, vulnerability, remoteness and limited access to medical services cause ARI, especially pneumonia, to be the main cause of death for many years, being an unfortunate constant[6].

Pneumonia, when detected in early stages, responds to antibiotic treatment with good results, when there are no medical or viral complications. Mortality associated with this disease is due to late diagnosis and complications associated with viruses, bacteria or comorbidities([7],[8],[9],[10]). For this reason, WHO created a program for the control of respiratory infections aimed at community workers, which did not have the desired impact due to the imprecision in evaluating and calculating some clinical data, such as respiratory frequency for example [11], [12].

Accurate diagnosis of pneumonia depends on medical expertise and requires evaluation of various clinical features; shortage of expertise and appropriate diagnostic tools hinders timely treatment[13].

The diagnosis of pneumonia provided by a health specialist is made with the chest X-ray and clinical data, but in rural areas of Peru these resources are not available, and even more health professionals to perform them, which is why 2 out of 3 deaths from pneumonia are out-of-hospital because of the mobilization of the patient for a diagnosis from one population center to another, which brings complications that increase the possibility of death in the patient [14].

One of the most important clinical data to be considered by the health specialist is the result of pulmonary auscultation. However, this is subjective and depends on the training and skill of the health personnel in charge. The health personnel must have the ability to recognize the different sound patterns, our research quantifies these values eliminating subjectivity to have a reliable classification of abnormal sounds to support the health personnel or community agent at the first level of care

that allows a diagnosis at the first moment in the most vulnerable, distant and in need areas[15].

The most common pulmonary sounds are crackles for pneumonias and wheezing in case of bronchitis. According to the sound analysis, the sounds are located between 100-1000 Hz, which is the frequency range we will cover for this study [16]. On the other hand, the acoustic stethoscope has a frequency response that attenuates the frequency components of the sound signal that the human ear is not particularly sensitive to, thus the analysis of sound in a wider frequency range can benefit the follow-up and evolution of cases[16].

This study details the creation of an initial algorithm for the classification of lung sounds from the characteristics of the audio signals, as a first step in the development of an automatic model that will further aim to classify them automatically and continuous learning, being this the first approach to algorithms of such kind by means of the definitions and characteristics of the sounds in general[15].

This research develops, by means of convolutional networks, the evaluation of lung sound using an automatic algorithm that with clinical data will allow assessing the risk of pneumonia at an early stage.

This research aims to contribute to this problem, proposing a mobile tool that through the use of Deep learning techniques can classify whether a patient is going through a pulmonary process or pneumonia from clinical data and the capture of lung sounds classifying them with convolutional neural networks. This project is based on computational neural network models for the classification of lung sounds, and identification of wheezing, hoarseness, or crackles for an effective analysis of support in those places where there is no medical personnel for timely diagnosis. In this way, artificial intelligence provides by means of learning the capacity of automatic classification based on artificial neural networks through the use of computational servers for the respective calculation[17]. The system could help non-expert technicians, such as health assistants or community workers, to detect pneumonia in low-resource settings where specialized personnel are not available to interpret ultrasound images.

II. POPULATION AND SAMPLE

A. Sample

In this investigation a public sound base was used, 2 research groups from Portugal and Europe, which has classifications and clinical data of 126 patients from which a total of 920 sounds were collected. The equipments used for sound collection are electronic stethoscopes:

- AKG C417L Microphone (AKGC417L).
- 3M Littmann Classic II SE Stethoscope (LittC2SE).
- 3M Litmann 3200 Electronic Stethoscope (Litt3200).
- WelchAllyn Meditron Master Elite Electronic Stethoscope (Meditron).

It is worth mentioning that due to the audio recording by the aforementioned equipment directly on the patients, the sounds captured have those emitted by the lungs and have

some noise coming from the heart or any other sound that exists in the body or external at the moment of the capture.

The sounds have a duration from 10 seconds to 90 seconds, captured by specialists and classified into five main ailments:

- Bronchiectasis
- Bronchiolitis
- COPD: Chronic Obstructive Pulmonary Disease
- Healthy
- Pneumonia
- URTI: Upper Respiratory Tract Infection

The use of convolutional neural networks for this research requires a large amount of data. The database used had a limited amount of them in some cases of classification because of this the process of data augmentation was used to obtain a base with better characteristics that support a better classification.

The Data Augmentation process has been carried out using three sequential processes. Fig. 1 presents the spectrogram of a normal sound at the beginning of the process.

Noise Addition: This process involves the addition of noise which means white noise to the sample. White noises are random samples distributed at regular intervals with a mean of 0 and a standard deviation of 1. In order to achieve this, we will use numpy's normal method, generate the above distribution and add it to our original sample. The following factors were used.

- Noise addition amplitude 0.005.
- Noise addition amplitude 0.007.

In Fig. 2 we can observe the effect of the Noise Addition on the normal sound shown in Fig. 1.

Time Shifting: This is the process of moving the acoustic wave to the right by a given factor along the time axis.

To accomplish this, we used numpy's roll function to generate time shifts. The following factors were used:

- Time shifting de factor $\text{sample_rate}/10$
- Time shifting de factor $\text{sample_rate}/5$.

In Fig. 3 we can observe the effect of the Time Shifting on the normal sound shown in Fig. 1.

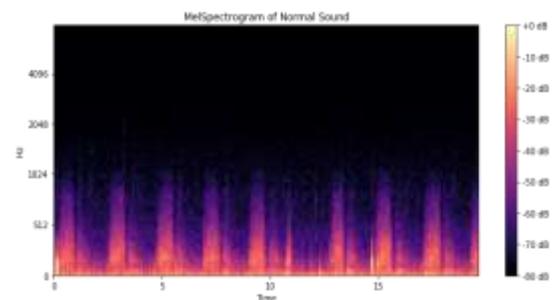


Fig. 1. Mel Spectrogram of Normal Sound.

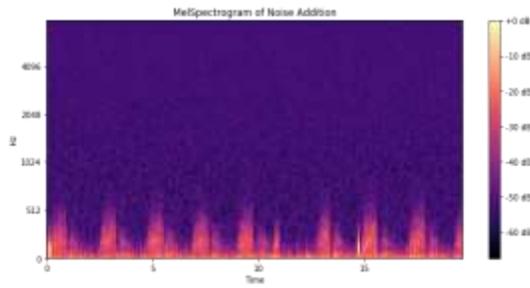


Fig. 2. Mel Spectrogram of Noise Addition with Amplitude of 0.005 for Normal Sound.

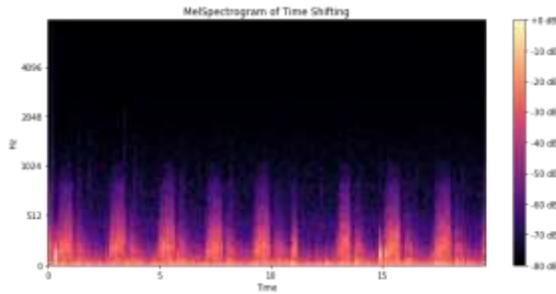


Fig. 3. Mel Spectrogram of Time Shifting on a Normal Sound with a Factor Sample_rate/5.

Pitch Shifting: An implementation of the pitch scale used in musical instruments. It is a process of changing the pitch of the sound without affecting its speed. We will use the `pitch_shift` function of `librosa` again (example in the Fig. 4). It samples the waveform, sampling frequency and number of steps through which the pitch should be shifted. The following factors are used:

- Pitch shifting de factor +2.
- Pitch shifting de factor +1.
- Pitch shifting de factor -1.
- Pitch shifting de factor -2.

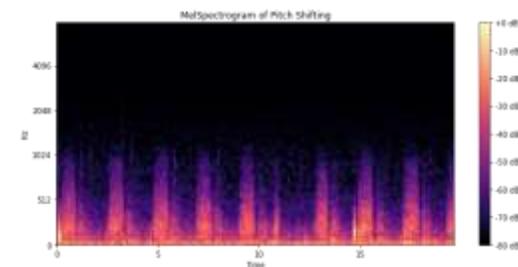


Fig. 4. Mel Spectrogram of Pitch Shifting de Factor -2, on a Normal Sound.

With this process, the data ended up increased by a factor of 9, as follows Table I shows the original data and the final data after the complete process.

In Table I, we can see in the first column, the initial classification followed in the second column by the number of data originally in the database. In the third column, the amount of data obtained after the data augmentation process. In the last two columns we see the classification that was performed, with

the classification between healthy and sick and in the last one the identification of healthy, pneumonia, and some other disease. Each base and classification has its respective model, based on the characteristics of the pulmonary sounds.

TABLE I. TABLE OF DATABASE ANALYSIS AND ITS CLASSIFICATION FOR EACH MODEL DEVELOPED

Pulmonary Sounds				
Classification by disease	Original number	Data Aumentation	Dichotomic classification	Classification Pneumonia and others
Healthy	35	315	Normal	Normal
Pneumonia	37	333	Abnormal	Pneumonia
Bronchiectasis	16	144	Abnormal	another disease
Bronchiolitis	13	117	Abnormal	another disease
COPD	793	7137	Abnormal	another disease
URTI	23	207	Abnormal	another disease

III. METHODOLOGY

The objective of this study is to model by means of convolutional neural networks, the classification of normal and abnormal sounds. To this end, the general process we used for the design of the study was based on a sequential process consisting of three main parts schematized in Fig. 5.

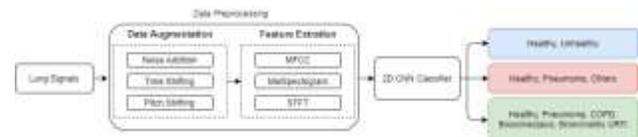


Fig. 5. Classification Process Flowchart.

The procedures shown are:

- Data Augmentation process, to achieve a more solid base, our original base had 917 sounds and our final base of the process has 8253 sounds, which we finally worked with, thus allowing an adequate classification and model. This process was explained in the population and sample section [17].
- Feature extraction process, for this process 3 methods were used: MelSpectrogram, STFT and MFCC, the sound features were combined and utilized to generate 3 models[7].
- Modeling process, after the features were extracted they were classified in three ways, the first to classify Normal and Abnormal sound, the second to classify Healthy, Pneumonia or other diseases, and the last one to allow classification by disease.

The feature extraction process is performed by 3 methods:

- Melspectrogram: is a nonlinear transformation of the Short Time Fourier Transform (STFT) on the Mel scale, a scale based on the human perception of tones.
- STFT: is a sequence of Fourier transforms of a windowed signal that provides time-localized frequency information for situations where the frequency components of a signal vary over time. As a general

rule, a narrow window width generates better resolution in the time domain, but generates poor resolution in the frequency domain and vice versa. Visualization of STFT is often done through its spectrogram, which is an intensity plot of the STFT magnitude over time. In the present work, a window length of 2048 and an offset length of 512 were used. Once the STFT is obtained, it is transformed to the Mel scale by applying a bank composed of multiple triangular band-pass filters, for this work 128 filters were used.

- MFCC: Concerning the field of sound signal processing, the analysis is usually completed with the calculation of the logarithm of the energy of each frequency band and with the calculation of the Direct Cosine Transform, obtaining the Mel Frequency Cepstral Coefficients (MFCC).

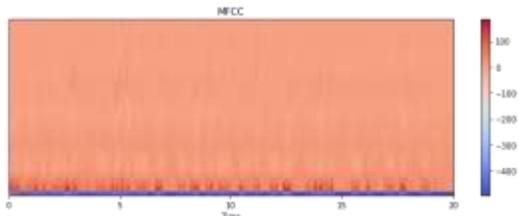


Fig. 6. Mel Frequency Cepstral Coefficients (MFCC).

The anomaly classifier using the convolutional neural networks that we have used is summarized in Fig. 6. For each classification, different variations of the convolutional neural network (CNN) architecture were performed, but all share certain similarities. For all models, the structure of the convolutional network was as follows: an input layer with dimensions of 256 x 256, followed by 7 convolutional layers with a 3x3 kernel and RELU activation function whose filters had dimensions of 32,64,128,256,512 and 1028 respectively, 7 layers of Max Pooling, 7 layers of dropout of 20%.

Next follows the final convolutional layer which has a GlobalAveragePooling2D type with softmax activation function followed by a fully connected 4 neuron and softmax activation function. Finally, an output layer that varies depending on the classification required, 2 dimensions for dichotomous classification, 3 dimensions for pneumonia classification and others, and 6 dimensions for complete classification. The Process is in the Fig. 7.

The Modeling process starts after the extraction of features given by the previous methods, with this data 3 classifications are modeled which are:

- Model 1: classification of normal and abnormal sound.
- Model 2: classification between normal sound, pneumonia and other diseases.
- Model 3: classification by disease and wellness, which would allow a finer classification.

Statistical Analysis

To evaluate the classification models of each of the 3 models selected for analysis in this paper, we have considered

some measures to validate their prediction or effectiveness[18], [19].

Among the measures to evaluate each model are:

Training Accuracy, this measure assesses the accuracy of the model that has been designed on the training set.

Testing Accuracy, this value represents the accuracy of the model performed on the test set, which is different from the one used for training.

Accuracy is the measure that quantifies the number of positive class predictions that actually belong to the positive class.

To remember, this measure quantifies the number of positive class predictions made from all positives in the data set.

F1-score, combines the measures of accuracy and completeness to return a more general measure of model quality.

Support, is the amount of data you have to evaluate the model.

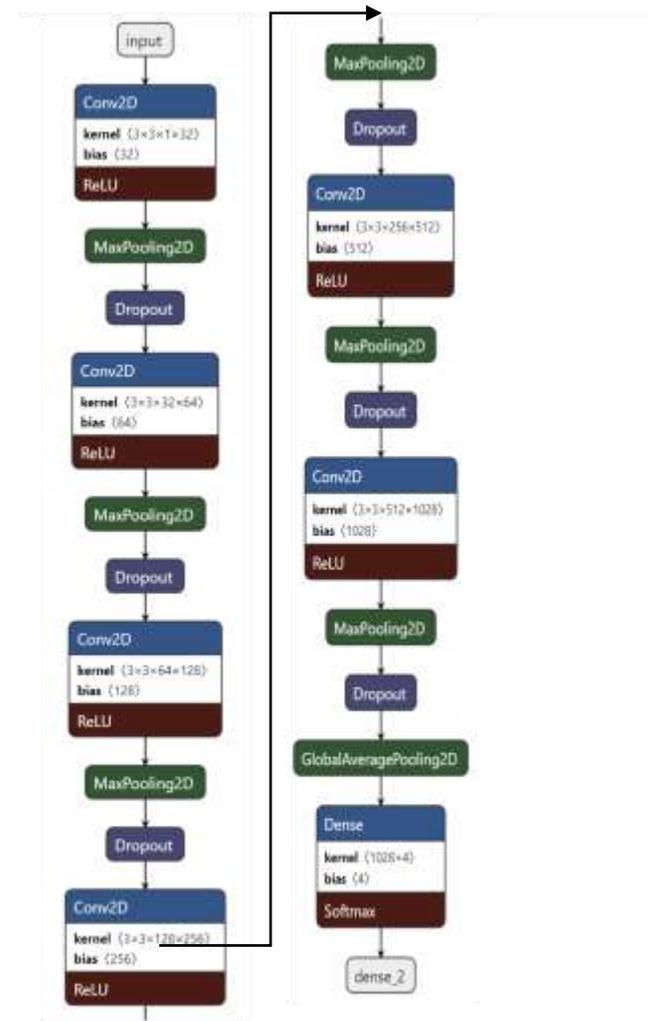


Fig. 7. Convolutional Neural Network Structure for Classification.

Confusion matrix, within the field of artificial intelligence, it is important to measure the power of statistical classification. The confusion matrix allows the visualization of the performance of an algorithm used in supervised learning, observing clearly in absolute values or percentages the way in which the model has classified the subgroups that we seek to model.

Receiver Operating Characteristic or ROC curve is a graphical representation of sensitivity against specificity for a binary classifier system according to varying discrimination threshold.

These measures will help us to assess and estimate the predictive ability of the classification on the set selected by the convolutional neural networks.

IV. RESULTS

This study was formulated to demonstrate the feasibility of an automatic design software based on convolutional neural networks in order to categorize pulmonary sounds depending on the extraction of the aforementioned features. We have detailed the variables that will support us in the evaluation of each model or classification set. In the following, we will show the results of the best model for each subgroup.

Model 1: classification of normal and abnormal sound.

For this model we work with a classified database based on the output of the data augmentation process, the details of the database conformation are shown in Table II.

For the MFCC feature extraction method. A training accuracy of 1 and a testing accuracy of 0.998 were obtained.

The details of the classification report are shown in Table III.

The confusion matrix based on percentages is detailed in the graph (Fig. 8).

For this model, we obtained an ROC curve of 1 for each classification.

Model 2: Classification between normal sound, pneumonia and other diseases.

For this model we work with a classified database based on the output of the data augmentation process, the details of the database conformation are shown in Table IV.

TABLE II. REGRESSION MODEL 1 RESULTS

Description	Number of samples	Percentage
Healthy	315	3.82
UnHealthy	7938	96.18
Total	8253	100

TABLE III. NORMAL (HEALTHY) AND ABNORMAL (NONHEALTHY) MODEL 1 CLASSIFICATION REPORTS

Description	Precision	Recall	f1-score	Support
Healthy	0.9981	1.0000	0.9991	1588
UnHealthy	1.0000	0.9524	0.9756	63

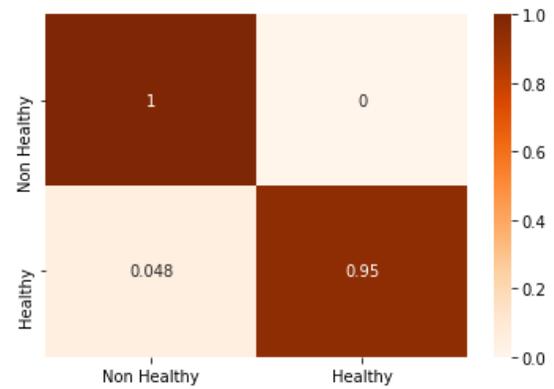


Fig. 8. Confusion Matrix in Percentages for Healthy or unhealthy Classification.

TABLE IV. REGRESSION MODEL 2 RESULTS

Description	Number of samples	Percentage
Healthy	315	3.82
Pneumonia	333	4.35
Others	7605	91.83
Total	8253	100

For the MEL feature extraction method. A training Accuracy of 0.9959 and a testing accuracy of 0.9885 were obtained.

Classification report of wellness, pneumonia, and other illnesses is described in Table V.

TABLE V. WELLNESS, PNEUMONIA, AND OTHER ILLNESSES CLASSIFICATION REPORT OF MODEL 2

Description	Precision	Recall	f1-score	Support
Healthy	0.9492	0.8889	0.9180	63
Pneumonia	0.9683	0.91044	0.9385	67
Others	0.9915	0.9967	0.9941	1521

The confusion matrix based on percentages is detailed in the Fig. 9.

Model 3: Classification by disease and wellness, which would allow a finer classification

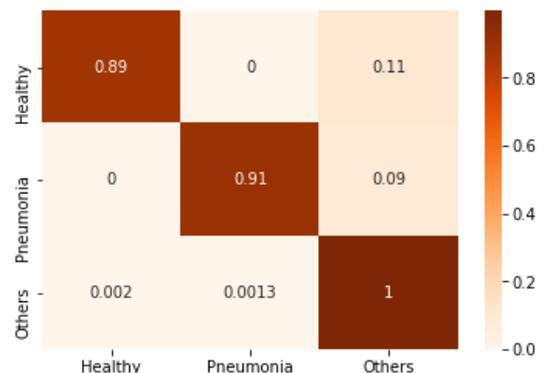


Fig. 9. Confusion Matrix in Percentages for the Classification of Wellness, Pneumonia, and other Ailments.

For this model we work with a database classified from the output of the data augmentation process, the details of the conformation of the database are shown in Table VI.

TABLE VI. REGRESSION MODEL 3 RESULTS

Description	Number of samples	Percentage
Healthy	315	3.82
Pneumonia	333	4.35
URTI	207	2.51
COPD	7137	86.48
Bronchiolitis	117	1.42
Bronchiectasis	144	1.74
Total	8253	100

For the MFCC feature extraction method. A training Accuracy of 0.9967 and a testing accuracy of 0.9909 were obtained.

The details of the classification report are shown in Table VII.

TABLE VII. NORMAL (HEALTHY) AND ABNORMAL (NONHEALTHY) MODEL 3 CLASSIFICATION REPORTS

Description	Precision	Recall	f1-score	Support
Healthy	1	0.9524	0.9756	63
Pneumonia	0.9692	0.9423	0.9545	67
URTI	0.8511	0.9756	0.9091	41
COPD	0.9958	0.9951	0.9954	1428
Bronchiolitis	1	1	1	23
Bronchiectasis	1	1	1	29

The confusion matrix based on percentages is detailed in Fig. 10 as follows.

The models presented in the results are based on the extraction of features provided by the methods explained, as part of the convolutional neural network that we have applied. These results have been applied taking as a sample the database after the data augmentation process, and the random division of the training and testing set.

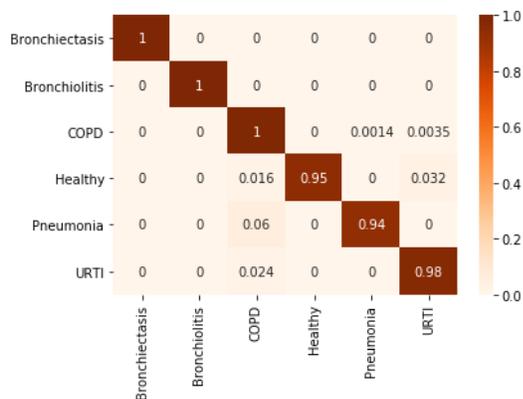


Fig. 10. Confusion Matrix in Percentages for the Classification by Health and each Disease.

For each case, we have provided tables showing the distribution of the sample in the respective categories within the models and results, where the values for each sub-classification are also described.

V. DISCUSSION

In this research, model 1, which classifies normal and abnormal sounds by means of convolutional neural networks, obtained precision values for normal sounds of 0.998 and precision values for abnormal sounds of 100.

Other studies in the literature have pursued similar objectives, some of which are discussed below.

In 2017, Aykanat [20], their research aims to improve audio data analysis through machine learning in order to classify respiratory sounds. They are based on auscultation, a non-invasive process for the diagnosis of pulmonary anomalies. In terms of materials, the Thinklabs One electronic stethoscope was used. They classify their information into 4 sets: (1) healthy versus pathological classification [17 930 clips], (2) rale, rhonchus and normal sound classification [15 328 clips], (3) singular respiratory sound type classification [14 453 clips] and (4) audio type classification with all sound types [17930 clips]. The acquired database is utilized in the highest quality format and in a real environment. In addition, they process their raw data, without filtering, amplification and with a single instrument and software adapted to the hospital environment. The spectrograms used are in 28x28 grayscale (to reduce storage and post-processing memory). They have a classification of healthy and pathological (1) with an average accuracy of 86% using Convolutional Neural Networks. Our research uses similar techniques, with better results, also using sufficient amount of data, but our approximations are higher than those detailed in this analysis.

Tracey [21] in 2011 researches on a low-cost cough monitoring system for drug-resistant tuberculosis (TB) patients (MDR), which could be used in places with poor access to specialized laboratories, the validation of the system algorithm, based on other research, is presented. The cough detection algorithm implements an image analysis of the MFCC through machine learning algorithms and compares the performance of manual diagnosis of cough audios and diagnosis by Neural Networks such as: Multilayer perceptron (MLP), Support Vector Machine (SVM) and Sequential Minimal Optimization (SMO). Its samples consist of 30 audio files corresponding to 10 random patients and a recording time of 30 minutes. In addition, a total of 13 429 cough frames and 43 925 non-cough frames were used to train the Neural Network. The algorithm designed with SMO classifies events as cough and non-cough. Moreover, it presents a sensitivity of 81%, which in the validation with 28 patients with MDR TB yields a promising recovery due to the indication of cough in the patient. The designed algorithm proposes an improvement in cough detection by estimating an empirical filter for environmental noise. Each recording file contains a number of frames used to train the neural network. It consists of two known and manually diagnosed patients. To reduce the CPU load in the training, the "divide and conquer" clustering method was used. If we compare with the present research we focus on pulmonary sound, the amount of sound samples they contain

are not comparable with those used in our study, it is worth mentioning that our analysis also has higher accuracy results than those indicated in the aforementioned research.

In 2019, Tariq[17] used the same database as the one presented in this paper and a similar methodology, but the results we obtained and the data augmentation processes are different. As their research in medicine and advanced technologies is associated with more accurate care and treatment, so the research aims to evaluate the degree of accuracy of Deep Learning and 2D Convolutional Neural Networks (CNN) using lung spectrogram data. The methodology is based on three steps: (1) normalization for information cleaning; (2) data augmentation, for training; and (3) CNN model, as identifier and classifier through libraries on MEL spectrograms. As a result of the 2D CNN (3) we obtain a model called lung disease classification (LDC) with 7 classifications: Asthma (0), bronchiectasis (377), COPD (10 205), Healthy (455), LRTI (26), Pneumonia (481) and URTI (403). Additionally, an accuracy of approximately 83% for original data and 97% for augmented and normalized data. The data augmentation methods expanded the original data by approximately 95%, which serves as a useful baseline for training. They do not specify the database format and do not prioritize the use of resources such as GPU or data storage. Accuracy increases significantly according to the normalization and data augmentation technique to the point of reaching 97% of their previous research. Our research uses the same data, without data normalization but uses data augmentation at 900%. This makes the results work better for model building, achieving a positive accuracy of 0.998 and a negative accuracy of 1.0.

VI. CONCLUSIONS

We can conclude that the highest classification found is with respect to the classification of normal and abnormal, with classification values of normal of 0.998 and abnormal of 100, which gives a high probability of success in implementing this technique in an automatic classification. The second finding is the one shown in the second model, where we reached model accuracy values of 0.9959 and 0.9885, this gives us evidence of an acceptable classification of the recognition of pneumonia over other ailments, since this is the primary interest of the study.

It is important to mention that in this work every model has been provided with a training and testing set. Also, the analysis of our experts showed that the sounds of the mentioned base are normal heterogeneous sounds of child and adult, trachea and lung that are captured by the digital stethoscope, commercial, captured in normal conditions of attention which leads to environmental noise, which could negatively influence the design of each model, but even so, the models are acceptable and we obtain from the analysis an adequate and significant value that encourages us to continue analyzing this method but now for sounds of patients with ailments such as COVID, during the pandemic in order to develop a tool for the first level of care in vulnerable areas with scarce resources.

This research is the proof of concept for the development of a complete system for the first level of care that supports the diagnosis through the automatic classification of sounds

between normal and abnormal, in conjunction with clinical data, is the first step as part of a process that we are starting in favor of reducing the gaps in health. Although there are recent studies that work with neural networks for classification, our study aims to achieve a better understanding of the differences between these sounds to achieve with an adequate result the discrimination of these, as part of a study to implement an automatic platform for remote and remote locations to assess the risk of pneumonia at the point of care so that we can work on low-cost tools to reduce child mortality in times of cold and highland areas of Peru.

ACKNOWLEDGMENT

We want to thank the Image Processing Research Laboratory.

(INTI-Lab) and the Universidad de Ciencias y Humanidades.

(UCH) for their support in this research, the National Fund for.

Scientific, Technological and Technological Innovation (FONDECYT), according to the research: "SAMAYCOV: "Desarrollo de un dispositivo electrónico portátil a bajo costo para evaluar riesgo de neumonía basado en sonido pulmonar anormal en pacientes con sospecha de COVID-19 en zonas vulnerables". CONVENIO 054-2020-FONDECYT"; for the financing of this research and the Electronics Laboratory of the UCH for assigning us their facilities and being able to carry out the respective tests.

REFERENCES

- [1] Vigilancia de las neumonías y meningitis bacterianas en menores de 5 años: guía práctica. Accessed: May 30, 2021. [Online]. Available: <https://iris.paho.org/handle/10665.2/49175>.
- [2] V. A. Laguna-Torres et al., "Vigilancia, prevención y control del virus de la influenza en Perú," *Rev. Peru. Med. Exp. Salud Pública*, vol. 36, pp. 511–514, Dec. 2019, doi: 10.17843/rpmesp.2019.363.4481.
- [3] P. y C. de E. Equipo Técnico de Vigilancia de las Enfermedades Crónicas Centro Nacional de Epidemiología, "2020 Boletín Epidemiológico del Perú," MINSa, Lima, 2020. [Online]. Available: <moz-extension://b645d046-b790-41d3-b658-78f8a97cc1ce/enhanced-reader.html?openApp&pdf=https%3A%2F%2Fwww.dge.gob.pe%2Fportal%2Fdocs%2Fvigilancia%2Fboletines%2F2020%2F02.pdf>.
- [4] ALEXIS PEREA FLORES, "El Frijaje y las heladas: diagnósticos de la problemática en el Perú y legislación comparada," Congreso de la República del Perú, Lima, 2018. [Online]. Available: moz-extension://b645d046-b790-41d3-b658-78f8a97cc1ce/enhanced-reader.html?openApp&pdf=https%3A%2F%2Fwww2.congreso.gob.pe%2Fsicr%2Fencodocbib%2Fcon5_uibd.nsf%2F322B8DD56F06DD6E0525832700584752%2F%24FILE%2F6-Frijaje_Heladas.pdf.
- [5] O. Feo et al., "Cambio climático y salud en la región andina," *Rev. Peru. Med. Exp. Salud Pública*, vol. 26, no. 1, pp. 83–92, Jan. 2009.
- [6] O. P. de la Salud, "Infecciones respiratorias agudas en el Perú: Experiencia frente la temporada de bajas temperaturas," OPS, 2014. [Online]. Available: <https://iris.paho.org/handle/10665.2/28549>.
- [7] P. Cisneros-Velarde et al., "Automatic pneumonia detection based on ultrasound video analysis," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2016, pp. 4117–4120, Aug. 2016, doi: 10.1109/EMBC.2016.7591632.
- [8] F. Cortellaro, S. Colombo, D. Coen, and P. G. Duca, "Lung ultrasound is an accurate diagnostic tool for the diagnosis of pneumonia in the emergency department," *Emerg. Med. J. EMJ*, vol. 29, no. 1, pp. 19–23, Jan. 2012, doi: 10.1136/emj.2010.101584.

- [9] C. Cillóniz, D. Rodríguez-Hurtado, A. Nicolini, and A. Torres, "Clinical Approach to Community-acquired Pneumonia," *J. Thorac. Imaging*, vol. 33, no. 5, pp. 273–281, Sep. 2018, doi: 10.1097/RTL.0000000000000343.
- [10] M. Fiszman, W. W. Chapman, D. Aronsky, R. S. Evans, and P. J. Haug, "Automatic detection of acute bacterial pneumonia from chest X-ray reports," *J. Am. Med. Inform. Assoc. JAMIA*, vol. 7, no. 6, pp. 593–604, Dec. 2000.
- [11] A. Hernandez, Y. Yamiléx, S. Valdes, M. Angel, and G. Cordero López, "Neumonía adquirida en la comunidad: aspectos clínicos y valoración del riesgo en ancianos hospitalizados," *Rev. Cuba. Salud Pública*, vol. 41, no. 3, pp. 413–426, Sep. 2015.
- [12] G. R. Gupta, "Tackling pneumonia and diarrhoea: the deadliest diseases for the world's poorest children," *Lancet Lond. Engl.*, vol. 379, no. 9832, pp. 2123–2124, Jun. 2012, doi: 10.1016/S0140-6736(12)60907-6.
- [13] Á. P. and A. María, "Neumonía adquirida en la comunidad en niños: Aplicabilidad de las guías clínicas," *Rev. Chil. Infectol.*, vol. 20, pp. 59–62, 2003, doi: 10.4067/S0716-10182003020100010.
- [14] T. Hazir et al., "Chest radiography in children aged 2-59 months diagnosed with non-severe pneumonia as defined by World Health Organization: descriptive multicentre study in Pakistan," *BMJ*, vol. 333, no. 7569, p. 629, Sep. 2006, doi: 10.1136/bmj.38915.673322.80.
- [15] A. Alva Mantari, B. Meneses-Claudio, and A. Roman-Gonzalez, "Automatic Algorithm for identifying Abnormal Lung Sounds through the Recognizing of Sound Patterns," in *2019 IEEE XXVI International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, Aug. 2019, pp. 1–4. doi: 10.1109/INTERCON.2019.8853612.
- [16] L. E. Ellington et al., "Computerised lung sound analysis to improve the specificity of paediatric pneumonia diagnosis in resource-poor settings: protocol and methods for an observational study," *BMJ Open*, vol. 2, no. 1, p. e000506, 2012, doi: 10.1136/bmjopen-2011-000506.
- [17] Z. Tariq, S. K. Shah, and Y. Lee, "Lung Disease Classification using Deep Convolutional Neural Network," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2019, pp. 732–735. doi: 10.1109/BIBM47256.2019.8983071.
- [18] X. Yu, S.-H. Wang, and Y.-D. Zhang, "CGNet: A graph-knowledge embedded convolutional neural network for detection of pneumonia," *Inf. Process. Manag.*, vol. 58, no. 1, p. 102411, Jan. 2021, doi: 10.1016/j.ipm.2020.102411.
- [19] "Morphological characterization of Mycobacterium tuberculosis in a MODS culture for an automatic diagnostics through pattern recognition. - PubMed - NCBI." <https://www.ncbi.nlm.nih.gov/pubmed/?term=24358227> (accessed Sep. 09, 2019).
- [20] M. Aykanat, Ö. Kılıç, B. Kurt, and S. Saryal, "Classification of lung sounds using convolutional neural networks," *EURASIP J. Image Video Process.*, vol. 2017, no. 1, p. 65, Sep. 2017, doi: 10.1186/s13640-017-0213-2.
- [21] B. H. Tracey, G. Comina, S. Larson, M. Bravard, J. W. López, and R. H. Gilman, "Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2011, pp. 6017–6020, 2011, doi: 10.1109/IEMBS.2011.6091487.

Cost Effective Hybrid Fault Tolerant Scheduling Model for Cloud Computing Environment

Hybrid Fault Tolerant Scheduling

Annabathula.Phani Sheetal¹, K.Ravindranath²

Department of Computer Science and Engineering

Koneru Lakshmaiah Education Foundation

Green Fields, Vaddeswaram, Guntur-522502, Andhra Pradesh, India

Abstract—Cloud computing provides flexible and cost effective way for end users to access data from multiplatform environment. Despite of support by the features of cloud computing, there are also chances of resource failure. Hence there is a need of a fault tolerant mechanism to achieve undisrupted performance of cloud services. The task reallocation and duplication are the two commonly used fault tolerant mechanisms. But task replication method results in huge storage and computational overhead, when the number of tasks is increasing gradually. If the number of faults is high, it incurs more storage overhead and time complexity based on task criticality. In order to solve these issues, we propose to develop a Cost Effective Hybrid Fault Tolerant Scheduling (CEHFTS) Model for cloud computing. In this model, the Failure Occurrence Probability (FoP) of each VM is estimated by finding the previous failures and successful executions. Then an adaptive fault recovery timer is maintained during a fault, which is adjusted based on the type of faults. Experimental results have shown that CEHFTS model achieves 43% reduced storage cost and 13% reduced response delay for critical tasks, when compared to existing technique.

Keywords—Cloud computing; failures; fault tolerant; critical tasks; scheduling; fault recovery; overhead

I. INTRODUCTION

Cloud computing is one of the new areas of technology that has emerged recently. On-demand resource sharing can be accomplished by utilizing internet and cloud services [1]. The primary goal is to offer services or distribute resources at the request of the customer with the lowest possible cost. There is a huge amount of heterogeneous resources [2], a vast user base, and a great range of application tasks in the cloud computing system. They have to deal with numerous individual activities, as well as big data. In terms of finance and science, cloud computing is a very valuable service [3]. A difficult challenge arises when deadlines are imposed on the jobs, since the required resources may be unavailable due to failures [4].

Along with missing project deadlines, fault tolerance is an important cloud computing aspect. It is designed to handle real-time tasks even if a failure occurs. Due to the higher storage overhead of taking backups, it is vital to use alternative approaches that have a high resource utilization ratio. Even though cloud features are attractive and a fault

tolerant technique is required to guarantee uninterrupted cloud service functioning [5].

Both internal and external factors contribute to some of the flaws. The two most often utilized fault-tolerant strategies in cloud computing are task reallocation and duplication. In task reallocation, a task is reassigned after a fault occurs. This method boosts the overall system resource usage. However, failing to meet the deadline constraints [6] of jobs could cause the response time to be prolonged. The rise in the amount of electricity consumed while making resource allocations causes resource requests to fail in data centers.

Both traditional application components and new traditional application components may be heterogeneous in their component composition, composed of scientific techniques. With virtualization technology, these components have their own specialized execution environment. Workflows can also be implemented and run in cloud environments that offer an ever-growing resource pool [7] that you pay just for what you use. On-demand cloud resources can thus be acquired and shared on workflows. Clouds serve as a favored scientific process execution environment due to these characteristics.

However, processes that operate in the clouds face obstacles, because the complexity and dynamics of processes and cloud breakdowns occur more frequently. Failures tend to stop continuous implementation and have considerable impacts on the performance of processes, particularly for big, long-run workflows. It is crucial that scientific workflow scheduling fault tolerants be achieved in order to provide consumers with flawless experience.

Due to the time-consciousness of many scientific procedures, their successful completion does not depend exclusively upon the correct results but also on the time when those results are available. For a workflow that must complete the working phase and generate the computer results, a deadline is specified for the workflow. Failures lead to deadline conscious workflows that cannot be concluded on time without an appropriate faults-tolerant schedule. In that circumstance, QoS is badly damaged, although beyond the deadline[8] the results may be acquired. Therefore, it is required and crucial to plan a defect tolerant workflow since processes can be successfully completed before the deadline.

This paper is organized as follows. Section 2 presents the motivation and objectives. Section 3 presents the proposed methodology. Section 4 presents the simulation results. Section 5 presents the conclusion.

II. MOTIVATION AND OBJECTIVES

In fault-tolerant workflow scheduling [6], task replication method is used. Here duplicate multiple copies of tasks [9] can execute simultaneously. But it results in huge storage and computational overhead, when the number of tasks is growing. Moreover, it did not present the strategy for finding the exact number of duplicate copies.

The Power, memory, and other network elements were addressed in order to enhance the reliability of resources. In task scheduling, the VM with the greatest level of reliability is picked. Hussein El Ghor et al. [8] has presented two scheduling algorithms. The first approach uses energy-efficient fault-free scheduling and the second approach provides essential slack time for fault recovery. The proposed model is limited in handling small group of tasks that cannot improve the performance levels.

Redundancy is a common way to protect against errors that can be done quickly or in space. Any additional resources required to perform the identical duplication or resilience replication process are designated as spatial redundancy. When considered in depth, an example of spatial redundancy may be parallel execution. There are considerable space expenses when you simultaneously run many jobs on different resources. In the case of redundancy, the task that was missed with respect to initial resources does not have to be redone because the clock was reset.

With Fault Tolerance Algorithm utilizing Selective Mirrored Tasks Approach (FAUSIT), they proposed a selective mirrored task method to ensure the correct balance between parallelism and topology for applications. DAG-based application fault tolerance is dealt with by limiting the number of make spans and minimizing computation costs.

For important jobs or tasks with permanent failures, the dynamic fault-tolerant workflow scheduling [10] uses spatial re-execution (SRE) and temporal re-execution (TRE) schemes in unison. However, as the number of issues detected for important operations is greater, the SRE strategy consumes more disc space. If the number of faults for non-critical tasks is excessive, then it has an adverse effect on time complexity. Additionally, the backups' odds of failure are not examined.

In FESTAL [11], for each primary task, a backup task is created and VMs are allocated for both. When the primary task is completed successfully within the execution time, then the execution of back task is terminated. But in this approach also, it results in huge storage and computational overhead [12], when the number of tasks becomes high. The performance levels are not satisfactory as the model causes overhead in the system.

In [13], a fault tolerant technique to ensure task completion deadline, has been proposed. But it does not check the types of tasks and storage overhead also is more that improves the delay levels in the system.

In [15], a trust based scheduling technique is proposed. For selecting the trusted computation service, set-based particle swarm optimization (S-PSO) and for selecting the storage service, ser covering problem (SCP) tree search are applied. The task scheduling concept in [16] is built on trust mechanisms. This model uses the Bayesian cognitive technique to determine the trustworthiness of computer nodes by examining the trust relationships that exist between them.

A research paper [17] outlines a trust-based Meta heuristic workflow scheduling technique called TMWS. The manual offers strategies and procedures to avoid the hazards associated with workflow scheduling.

Hence the fault tolerant algorithm should meet the following objectives:

- To reduce storage overhead.
- To reduce time complexity.
- To reduce the chances of backup VM failures.
- To meet deadline of critical tasks.
- To reduce the response delay.
- To reduce the fault recovery time.
- To increase the CPU utilization.

In order to solve these issues, we propose to develop a cost effective hybrid fault tolerant scheduling scheme for cloud computing.

III. COST-EFFECTIVE HYBRID FAULT TOLERANT SCHEDULING (CHFTS) MODEL

A. Overview

In this paper, Cost-effective Hybrid Fault Tolerant Scheduling (CHFTS) model is proposed. Figure 1 shows the block diagram of the CHFTS model.

The CHFTS model uses previous failures and successful executions to calculate the VM failure occurrence probability (FoP). The anticipated execution time (EET) is estimated for each job T_i . At this point, the projects with deadlines of less than one week are classed as high-priority activities. The designation of some VMs as fault-tolerant based on FoP. A fixed number of primary and backup VMs are assigned to each task. A recovery timer is initiated if a failure occurs in any VM. A notification will be delivered to the failed VM, and the VM will continue execution from the point in time that the failure occurred.

B. System Model

Virtualization divides each server in the Cloud system into a series of heterogeneous virtual machines (VMs). The fact is that a VM is an essential component of a cloud system.

Suppose that a cloud system offers a set of VM resources, as given by.

$$VM = \{VM(1), \dots, VM(k), \dots, VM(k)\}$$

to users.

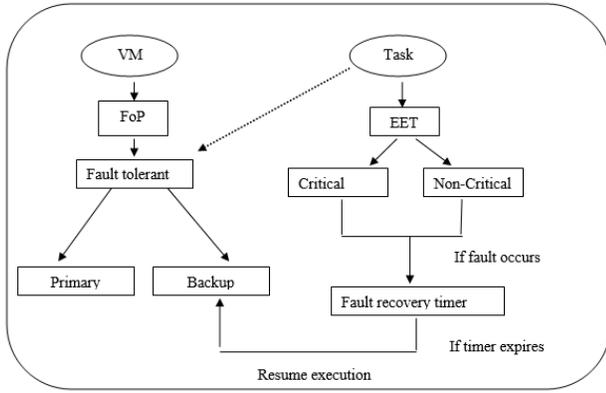


Fig. 1. Block Diagram of CHFTS Model.

The model represents the system structure, as seen in Figure 2. Processing capacity and cost per hour are particularly important for selecting VM(k) because of processing capacity and cost per hour, respectively (k). Users can access an endless number of VMs on cloud computing platforms that utilize virtualization. In addition, all VMs are hosted in one cloud data center to provide consistent bandwidth for all the VMs.

C. Fault Model

If a task is incomplete because of internal or external issues, it may be considered as a task failure.

Faults may occur during the execution of scientific work flows in cloud computing environments. There are various reasons behind these faults. The most common reason for task failure is the failure of corresponding VM. The other reasons for faults include unavailability of enough resources, resource overloading, delayed execution time etc.

There are two cases of faults that may occur in cloud, during execution. The first one is a permanent fault and other one is a temporary fault. In temporary faults, the faults may be recovered within short span of time whereas in permanent faults, the faults cannot be rectified until the failed element is repaired or replaced. Fail-signal or acceptance test is a fault detection mechanism, commonly used.

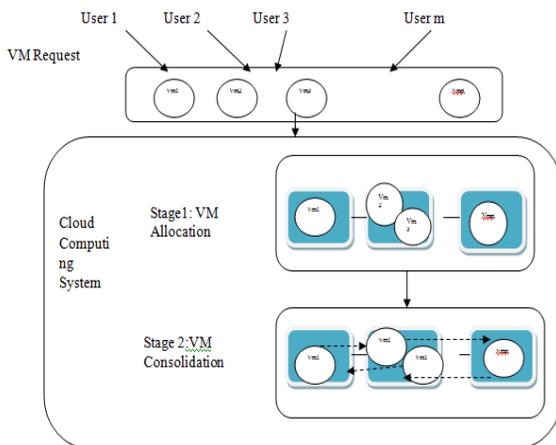


Fig. 2. System Model.

Due to task failures, the make span time of work flows will be increased and the service-level agreement (SLA) will be violated. Re-execution is one of the commonly used less expensive fault tolerant technique, which improves the reliability of workflows. Re-execution can be implemented in two methods: Spatial re-execution on other resources (SRE) and Temporal re-execution on the same resources, after fault recovery (TRE).

1) *Estimation of FoP*: The exponential probability density function for a specified mean time (M) between failures, is given by.

$$f(t) = \frac{1}{M} e^{-t/M} \quad (1)$$

Then, the probability of a failures, that occurs before time Δt for a VM is given by.

$$P(t \leq \Delta t) = \int_0^{\Delta t} \frac{1}{M} e^{-t/M} dt = 1 - e^{-\Delta t/M} \quad (2)$$

Hence the probability of successful completion during the time Δ , is given by.

$$P(t > \Delta t) = 1 - P(t \leq \Delta t) = e^{-\Delta t/M} \quad (3)$$

Let T_c be the computation time for a subtask executing on a VM and ϕ be the computation interval between two check points.

Then, the average number of attempts (NoA) required to finish T_c is given by.

$$NoA = \frac{T_c/\phi}{P(t > \Delta t)} = \frac{T_c e^{\Delta t/M}}{\phi} \quad (4)$$

The number of success is given by

$$NoS = \frac{T_c}{\phi} \quad (5)$$

Hence the number of failures NoF, during Δt is given by

$$(ie) NoF = NoA - NoS \quad (6)$$

$$\frac{T_c e^{\Delta t/M}}{\phi} - \frac{T_c}{\phi} = \frac{T_c}{\phi} (e^{\Delta t/M} - 1) \quad (7)$$

Hence

$$FoP = NoF / T_c, = \frac{1}{\phi} (e^{\Delta t/M} - 1) \quad (8)$$

Note: It is assumed that only single failure occurs during one computation interval.

D. Estimation of Expected End Time (EET)

Let $P(k)$ is considered as processing capacity of VM_k , $k=,1,2...K$.

Let $S(t_i)$ and $W(t_i)$ be the workload size of given input task t_i .

If and only if all the input data has been received from all prior tasks, then the task is started (t_i).

start time of t_i is signified by Eq. (9).

$$T_{start}(t_i) = \max_{t_j \in pre(t_i)} \{T_{end}(t_j)\} \quad (9)$$

Note that if $t_i=t_{entry}$, then $T_{start}(t_i)=0$. Then, the transmission time is computed by,

$$T_{trans}(t_i) = S(t_i) / BW \quad (10)$$

where BW is the bandwidth between two VMs.

Then, the execution time is given by,

$$T_{exec}(t_i, VM(k)) = W(t_i) / P(k) \quad (11)$$

Thus, the end time of task t_i is given by.

$$T_{end}(t_i) = T_{start}(t_i) + T_{trans}(t_i) + T_{exec}(t_i, VM(k)) \quad (12)$$

Hence the make span of workflow is given by.

$$makespan = T_{end}(t_{exit}) \quad (13)$$

E. Fault Tolerant Scheduling

Let {FVM} be the set of fault tolerant VMs with maximum FoP.

The type of each task is categorized into CRITICAL and NON-CRITICAL based on the values of EET. (ie) The tasks with shorter EET are considered as CRITICAL and others are considered as NON-CRITICAL.

For each task, a pair of VMs $\langle VM_p, VM_b \rangle$ from the k VMs, are allocated, where VM_p and VM_b are primary and backup VMs. If there is a fault occurs in VM_p , then a fault recovery timer (T_{FR}) will be started.

Case-1: If the task is a critical task, then T_{FR} is fixed such that.

$T_{FR} < EET-FOT$, where FOT is the time of fault occurrence.

Case-2: If the task is a non-critical task, then T_{FR} is fixed such that.

$T_{FR} < (EET-FOT) * \mathcal{D}$, where \mathcal{D} is a scaling factor ranges between (2, 4).

The total cost of the Hybrid Fault Tolerant Scheduling Model is calculated as.

$$CC = NoS * \min(NoA) + \frac{T_{End}-T_{exec}}{\text{count}(VM)} + \Delta t * \varphi \quad (14)$$

If the fault cannot be recovered within T_{FR} , (ie) If fault recovery time (FRT) is more than T_{FR} , then immediately, a notification will be sent to VM_b , which will resume the execution from that point.

Algorithm - Fault Tolerant Scheduling

Input : List of VMs and Tasks

Output: fault tolerant scheduling

1. **foreach** VMj
 2. Estimate FoP_j using (5)
 3. If FoP_j >= Max(FoP), then
 4. Include FoP_j in {FVM}
 5. End if
 6. **End for**
 7. **foreach** task t_i
 8. Estimate EET_i using (11)
 9. **if** EET_i <= EET_{min}, **then**
 10. Type(t_i) = CRITICAL
 11. **else**
 12. Type(t_i) = NON-CRITICAL
 13. **end if**
 14. **end For**
 15. **foreach** incoming task t_j
 16. Allocate (VM_p, VM_b) ∈ {FVM}
 17. **if** fault occurs at time FOT, **then**
 18. **if** Type(t_i) = CRITICAL, **then**
 19. Set $T_{FR} < EET-FOT$
 20. **else**
 21. Set $T_{FR} < (EET-FOT) * \mathcal{D}$
 22. **end if**
 23. **if** FRT < T_{FR} , **then**
 24. Resume t_j at FRT
 25. **else**
 26. Send notification to VM_b
 27. Resume t_j at VM_b
 28. **end if**
 29. **end if**
 30. **end For**
-

Special outcomes of this approach

1) Since fault free VMs are selected, the chances of faults are less.

2) It does not result in huge storage since the VMs are allocated only when the fault is not recoverable (permanent fault) within a time span.

3) The time critical tasks will be executed within their deadlines since the tasks are migrated to backup VM, within the deadline.

4) The non critical tasks can be executed in the same VMs if the fault is recovered within tolerable time span (which is an even multiple of its deadline).

5) If the task could not be recovered beyond the tolerable time span, then only the non critical task are migrated to backup VMs, thereby reducing the storage overhead significantly.

6) Since the execution of tasks starts in the backup VM only from the time of fault, the time complexity will be less.

IV. EXPERIMENTAL RESULTS

The proposed CEHFTS model is compared with the Spatial-Temporal Re-Execution method (SRE) [10]. The NASA workload [14] has been used as the emulator of Web users requests to the Access Point (AP). This workload represents realistic load deviations over a period time. It comprises 100960 user requests sent to the Web servers during a day. Table 1 shows the experimental parameters assigned in this work.

TABLE I. EXPERIMENTAL PARAMETERS

Parameter	Value
Work load	NASA traces
Resource Utilization Thresholds	$U^{low-thr} = 20\%$ and $U^{high-thr} = 80\%$
Response Time Thresholds	$RT^{low-thr} = 200ms$ and $RT^{high-thr} = 1000ms$
Scaling Intervals	$\Delta t = 10min$
Desired Response Time	$DRT = 1000ms = 1s$
Fault rate	1 to 2
Configuration of VMs	Medium and Large
Maximum On-demand VM Limitation	$MaxVM = 10VM$
Task and Resources Scheduling Policy	Time-Shared

A. Results for Critical Tasks

In this section, work flows with strict deadlines are considered (critical tasks) and allocated to VMs. The FoP of each task is increased in terms of fault rate. The performance metrics response delay, percentage of missed deadlines, CPU utilization and storage overhead of server are considered for evaluation.

Table 2 Performance Results of CEHFTS and SRE schemes for fault rate of critical tasks.

Figure 3 shows the response delay measured for the critical tasks when the fault rate is increased from 1 to 2. Since in the case of time bound critical tasks, backup VM migration is mostly supported, the response delay will be high. Hence, CEHFTS has the delay in the range of 1.4 to 2.8 seconds whereas SRE has delay in the range of 1.7 to 3.1 seconds. However, in case of CEHFTS, the failed tasks are resumed when the recovery time is less than EET. Hence it has 13% reduced response delay, when compared to SRE scheme.

TABLE II. SHOWS THE PERFORMANCE RESULTS OF BOTH CEHFTS AND SRE SCHEMES FOR VARYING THE FAULT RATE OF CRITICAL TASKS

Fault rate (Critical)	Response Delay (sec)		% of Missed Deadlines		CPU Utilization (%)		Storage cost of Server	
	CEHFTS	SRE	CEHFTS	SRE	CEHFTS	SRE	CEHFTS	SRE
25	1.472	1.782	18	25	57	47	150	270
50	1.573	1.893	26	32	52	43.1	178	356
75	2.069	2.245	31	38	45.3	38.2	287	452
100	2.275	2.628	35	44	44.2	34	368	628

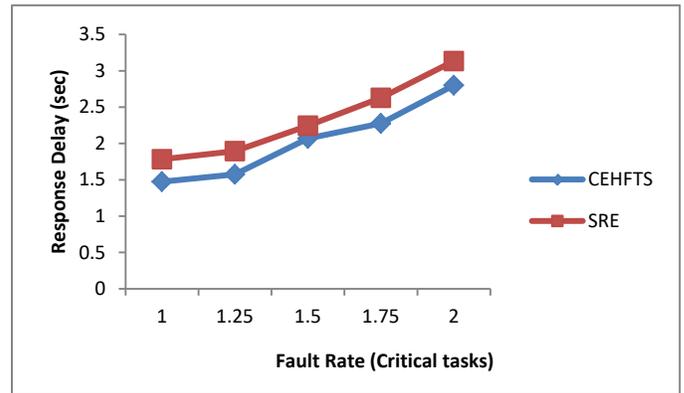


Fig. 3. Response Delay for Critical Tasks.

Figure 4 shows the missed deadlines for the critical tasks when the fault rate is increased from 1 to 2. Since in the case of time bound critical tasks, if the recovery time is higher than EET or if the fault is permanent fault, the chances of deadline miss are high. As depicted from the figure, CEHFTS has the missed deadlines in the range of 18% to 41% whereas SRE has missed deadlines in the range of 25% to 47%. However, in case of CEHFTS, the failed tasks are resumed when the recovery time is less than EET. Hence it has 19% reduced missed deadlines, when compared to SRE scheme.

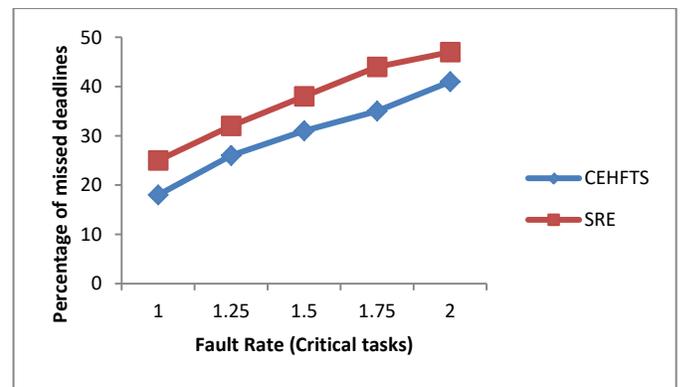


Fig. 4. Missed Deadlines for Critical Tasks.

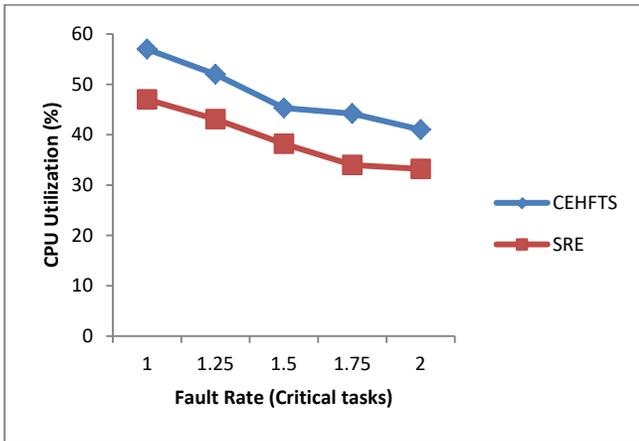


Fig. 5. CPU Utilization (%) for Critical Tasks.

Figure 5 shows the CPU utilization for the critical tasks when the fault rate is increased from 1 to 2. In case of permanent failures, the CPU utilization of primary VMs will be affected for the critical tasks. As it can be seen, from the figure, the utilization of CEHFTS reduces from 57% to 41% and the utilization of SRE reduces from 47% to 33%. However, CEHFTS has 18% higher CPU utilization than SRE scheme, since it minimizes the use of backup VMs.

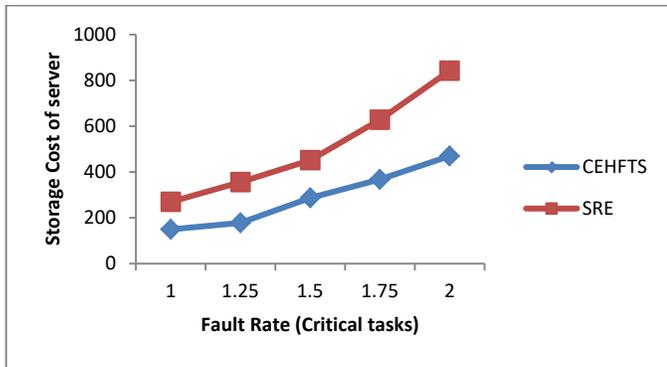


Fig. 6. Storage Cost for Critical Tasks.

Figure 6 shows the storage cost of servers in case of critical tasks. Since in case of time bound critical tasks, if the recovery time is higher than EET or if the fault is permanent fault, the tasks are migrated to backup VMs, leading to increased storage cost. As depicted from the figure, CEHFTS has cost in the range of 150 to 470 whereas SRE has cost in the range of 270 to 840. However, in case of CEHFTS, the failed critical tasks are resumed when the recovery time is less than EET. Hence it has 43% reduced storage cost, when compared to SRE scheme.

B. Results for Non-Critical Tasks

In this section, work flows with elastic (flexible) deadlines are considered (critical tasks) and allocated to VMs. The FoP of each task is increased in terms of fault rate. The performance metrics response delay, CPU utilization and storage overhead of server are considered for evaluation.

Table 3 Performance Results of CEHFTS and SRE schemes for fault rate of non-critical tasks

TABLE III. SHOWS THE PERFORMANCE RESULTS OF BOTH CEHFTS AND SRE SCHEMES FOR VARYING THE FAULT RATE OF CRITICAL TASKS

Fault rate (Non-critical)	Response Delay (sec)		CPU Utilization (%)		Storage cost of Server	
	CEHFTS	SRE	CEHFTS	SRE	CEHFTS	SRE
25	2.672	3.582	52	45	120	165
50	3.573	4.813	46	42.7	138	228
75	5.069	7.145	44.2	36.4	187	280
100	5.755	7.628	41.5	33.1	228	355

Figure 7 shows the response delay measured for the non-critical tasks when the fault rate is increased from 1 to 2. Since in case of non-critical tasks, the fault response timer is set as high, the response delay will be higher than that of critical tasks. Hence, CEHFTS has the delay in the range of 2.6 to 6.4 seconds whereas SRE has delay in the range of 3.5 to 8.3 seconds. However, in case of CEHFTS, the failed tasks are resumed when the recovery time is less than EET. Hence it has 25% reduced response delay, when compared to SRE scheme.

Figure 8 shows the CPU utilization for the non-critical tasks when the fault rate is increased from 1 to 2. In case of permanent failures, the CPU utilization of primary VMs will be affected for the non-critical tasks also. As it can be seen, from the figure, the utilization of CEHFTS reduces from 52% to 40% and the utilization of SRE reduces from 45% to 30%. However, CEHFTS has 16% higher CPU utilization than SRE scheme, since it minimizes the use of backup VMs.

Figure 9 shows the storage cost of servers in case of non-critical tasks. Since, if the fault is permanent fault, the tasks are migrated to backup VMs, leading to increased storage overhead. As depicted from the figure, CEHFTS has cost in the range of 120 to 270 whereas SRE has cost in the range of 165 to 472. However, in case of CEHFTS, the failed non-critical tasks are resumed when the recovery time is less than EET. Hence it has 35% reduced storage cost, when compared to SRE scheme.

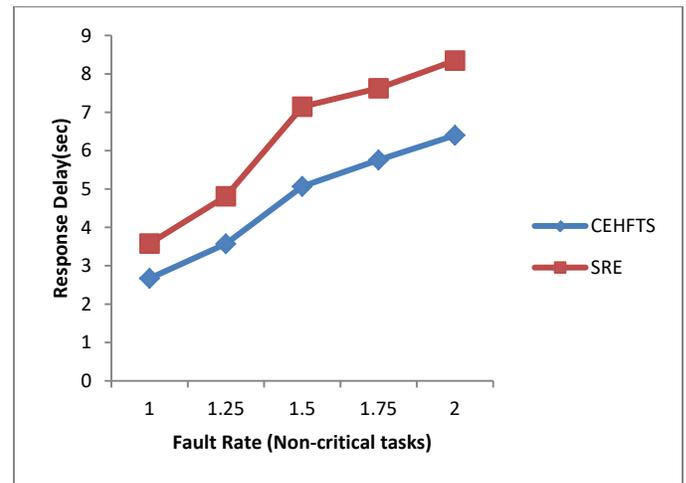


Fig. 7. Response Delay for Non-critical Tasks.

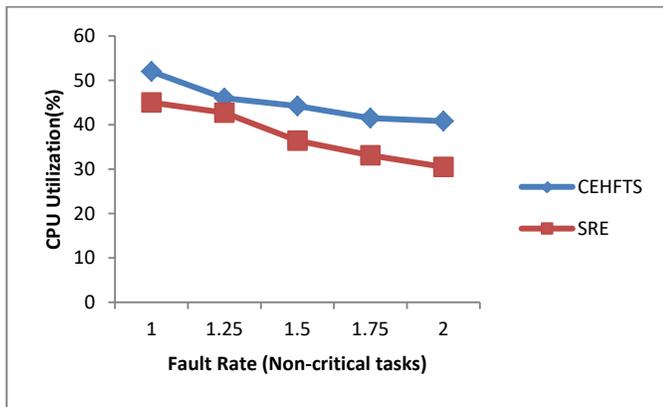


Fig. 8. CPU Utilization (%) for Non-critical Tasks.

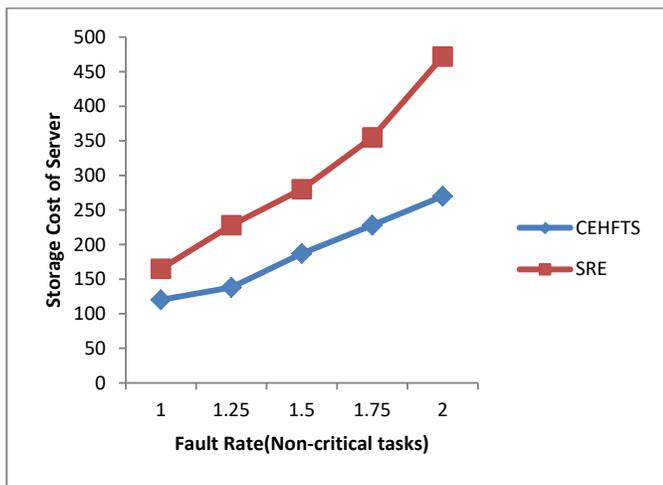


Fig. 9. Storage Overhead for Non-critical Tasks.

V. CONCLUSION

In this research, a new hybrid fault tolerant scheduling approach, called Cost Effective Hybrid Fault Tolerant Scheduling (CEHFTS), is proposed for cloud computing. For this model, the prior failures and successful executions are used to estimate the VM failure probability (FoP). Estimated anticipated execution time (EET) is calculated for each task. At this point, the projects with deadlines of less than one week are classed as high-priority activities. An adaptive fault recovery timer is kept running until a fault occurs, which is then increased as a function of the various fault types. Experiments are conducted for critical and non-critical tasks by varying the fault rate. Experimental results have shown that CEHFTS model achieves 43% reduced storage cost and 13% response delay for critical tasks, when compared to existing technique. Future work focus on grouping the tasks based on its type and requirements so that the CPU utilization and battery power can be further reduced. The task deadline and time scheduling can still be improved to enhance the performance levels.

REFERENCES

- [1] Priti Kumari and Parmeet Kaur, "A survey of fault tolerance in cloud computing", Journal of King Saud University – Computer and Information Sciences, Elsevier, 2018.
- [2] Vinay K and S M Dilip Kumar, "Fault-Tolerant Scheduling for Scientific Workflows in Cloud Environments", IEEE 7th International Advance Computing Conference (IACC), 2017.
- [3] J.Soniya,J.Angela Jennifa Sujana and T.Revathi, "Dynamic Fault Tolerant Scheduling Mechanism for Real Time Tasks in Cloud Computing", IEEE International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT),2016.
- [4] Pengze Guo and Zhi Xue, "Real-Time Fault-Tolerant Scheduling Algorithm with Rearrangement in Cloud Systems", IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC),2017.
- [5] Sreelekshmi S and K R Remesh Babu, "Synchronized Multi-Load Balancer with Fault Tolerance in Cloud", International Journal of Computer Information Systems and Industrial Management Applications, Vol-10,pp:107-114,2018.
- [6] Zhongjin Li, Jiacheng Yu, Haiyang Hu, Jie Chen, Hua Hu, Jidong Ge and Victor Chang, "Fault-Tolerant Scheduling for Scientific Workflow with Task Replication Method in Cloud", IEEE 3rd International Conference on Internet of Things, Big Data and Security,2018.
- [7] Diptee H. Devmurari and Kashyap Raiyani, "Resource Reliability using Fault Tolerance aware Scheduling in Cloud", International Journal of Innovative Research in Computer and Communication Engineering, Vol-5,No-4,2017.
- [8] Hussein El Ghor, Julia Hage, Nizar Hamadeh And Rafic Hage Chehade, "Energy-Efficient Real-Time Scheduling Algorithm For Fault-Tolerant Autonomous Systems", Scalable Computing: Practice and Experience, Volume 19, Number 4, pp. 387–400,2018.
- [9] Hao Wu , Qinggeng Jin ,Chenghua Zhang and He Guo, "A Selective Mirrored Task Based Fault Tolerance Mechanism for Big Data Application Using Cloud", Wireless Communications and Mobile Computing, Hindawi, Volume 2019,pp:1-12,2019.
- [10] Na Wu , Decheng Zuo and Zhan Zhang, "Dynamic Fault-Tolerant Workflow Scheduling with Hybrid Spatial-Temporal Re-Execution in Clouds", Information, MDPI, Vol-10,No-169,2019.
- [11] Ji Wang, Weidong Bao, Xiaomin Zhu, Laurence T. Yang and Yang Xiang, "FESTAL: Fault-Tolerant Elastic Scheduling Algorithm for Real-Time Tasks in Virtualized Clouds", IEEE Transactions on Computers, Volume: 64 , Issue: 9, Sept. 1 2015.
- [12] Punit Gupta and S. P. Ghrera, "Power and Fault Aware Reliable Resource Allocation for Cloud Infrastructure", International Conference on Information Security & Privacy (ICISP2015), Elsevier, December 2015, Nagpur, INDIA.
- [13] Sheng Di, Cho-Li Wang, "Error-Tolerant Resource Allocation and Payment Minimization for Cloud System", IEEE Transactions On Parallel And Distributed Systems, Vol. 24, No. 6, June 2013, pp 1097-1106, DOI 10.1109/TPDS.2012.309.
- [14] M. F. Arlitt and C. L. Williamson, "Internet web servers: Workload characterization and performance implications," IEEE/ACM Transactions on Networking (ToN), vol. 5, pp. 631-645, 1997.
- [15] Y.L. YANG, X.G. PENG, J.F. CAO, "Trust-Based Scheduling Strategy for Cloud Workflow Applications", INFORMATICA, 2015, Vol. 26, No. 1, 159–180.
- [16] Wei Wang, Guosun Zeng, Junqi Zhang and Daizhong Tang, "Dynamic trust evaluation and scheduling framework for cloud computing", Security and Communication Networks,2012, 5,311–318.
- [17] G. Jeeva Rathanam and A. Rajaram, "Trust Based Meta-Heuristics Workflow Scheduling in Cloud Service Environment", Circuits and Systems, 2016, 7, 520-531.

An Efficient Image Clustering Technique based on Fuzzy C-means and Cuckoo Search Algorithm

Lahbib KHRISSI¹, Hassan SATORI³
Khalid SATORI⁴

LISAC, Department of Computer Science
FSDM of Fez, Sidi Mohamed Ben Abdellah University
Fez Morocco

Nabil EL AKKAD²

Laboratory of Engineering
Systems and Applications, ENSA of Fez.
Sidi Mohamed Ben Abdellah University Fez Morocco

Abstract—Clustering is a predominant technique used in image segmentation due to its simple, easy and efficient approach. It is very important for the analysis, extraction and interpretation of images; which makes it used in multiple applications and in various fields. In this article, we propose a different image segmentation technique based on the cooperation between an optimization algorithm which is the Cuckoo Search Algorithm (CSA) and a clustering technique which is the Fuzzy C-means (FCM). The clustering method we propose goes through two major steps. In the first step, CSA explores the entire search space of the specified data to find the optimal clustering centers. Subsequently, these centers are evaluated using a new objective function. The result of the first step is used to initialize the FCM algorithm in the second step. The efficiency of the suggested method is measured on several images selected from the BSD300 database and we compare it with other algorithms such as FCM optimized by genetic algorithms (FCM-GA) and FCM optimized by particle swarm optimization (FCM-PSO). The experimental results on the different algorithms used in this paper show that the proposed method improves the segmentation results, based on the analysis of the best values of fitness, MSE, PSNR, CC, RI, GCE, BDE and VOI.

Keywords—Clustering; classification; image segmentation; fuzzy c-means; cuckoo search algorithm

I. INTRODUCTION

Segmentation is an important step in extracting qualitative information from the image. It is done via dividing the image in question into regions with homogeneity according to a predefined criterion (gray level, color, intensity, texture, etc.). Several segmentation approaches have appeared in recent years. Some of them seek to delimit homogeneous regions by their contours (contour approach) while others seek to find homogeneous regions (region approach).

The segmentation process represents a crucial step in computer vision systems, as features and decisions are extracted and made from its output. The first image segmentation algorithms were developed in the 1970s. Since then, many techniques and methods of segmentation have been experimented to try to improve the results. Nevertheless, until today, no classical image segmentation algorithm can provide perfect results on a wide variety of images.

Several segmentation techniques exist and four main categories can be distinguished: segmentation by classification [1- 3], by regions [4], by contours [5], and finally segmentation

by region-contour cooperation [6]. Clustering is considered among the most used image segmentation algorithms. The latter is a field of machine learning belonging to unsupervised learning. Clustering is mainly used to group populations into communities with similar common criteria. It is a data mining task that aims at dividing the elements of a set into groups, i.e. to establish a partition of this set. Each group must be as homogeneous as possible, and the groups must be as heterogeneous as possible. However, classical clustering methods converge to the local optimum and require a prior initialization of cluster centers. Therefore, unsupervised classification is studied as an optimization domain, i.e., finding a partition of the data that optimizes a given feature.

The image processing system presented has a multidisciplinary aspect. Its applications can be found in various fields like medical imaging [7], video analysis [8], and remote sensing [9]. In the literature, we do not find a technique to generalize it to image segmentation. Each method is used for a given type of image and in a well-defined computing context to know its performance and efficiency. Therefore, the different techniques proposed for image segmentation have asserted their defects and limitations. Researchers then found new, more flexible and efficient strategies to solve the segmentation problem, using metaheuristic approaches that now occupy an increasing place in the clustering framework for image segmentation. Metaheuristics are a set of algorithms that allow finding the fastest and most efficient solution for several optimization problems for which no more efficient classical method is known. They are iterative, i.e. starting from a single solution considered as a starting point, the search consists in moving from one solution to a neighboring solution by successive moves in a neighborhood constituted by the set of solutions by examining the fitness function.

The growing interest in metaheuristics is justified by the development of machines with enormous computational capacities, which has allowed the design of more and more metaheuristics that have proven to be quite efficient in addressing many problems like image segmentation. Thus, metaheuristics is a generic approach which operating principle is based on general mechanisms independent of any problem. Metaheuristics are stochastic and can therefore avoid being trapped in local minima. They are mainly guided by chance; however they are often combined with other algorithms in order to accelerate their convergence.

Metaheuristics are classified into two broad classes: single solution metaheuristics and solution population metaheuristics. Population-based optimization methods improve a population of solutions over time. The advantage of these methods is to use the population as a factor of diversity. Furthermore, single-solution optimization methods are called trajectory methods, i.e., they allow a trajectory to be described during the search process. In the literature, there are several algorithms based on a population of solutions employed to increase the quality of segmented images, like: evolutionary algorithms [10], genetic algorithms [11], PSO algorithm [12], ABC (Bee Colony Algorithm) algorithm [13], the SCA algorithm (Sine Cosine Algorithm) [14], the cuckoo search algorithm [15] and others. The operation of metaheuristics is progressive and iterative. The initial step is often chosen randomly and the stopping step is often fixed using a stopping criterion. All metaheuristics rely on the balance between search intensification and diversification. Otherwise, we will see a convergence towards local minima through too long an exploration due to lack of intensification or lack of diversification.

Several different methods have been devoted to unsupervised automatic classification. However, the evolution towards metaheuristics has given, in some difficult cases, very good results. In order to help improve the efficiency and performance of clustering-based image segmentation methods, we used a metaheuristic called "Cuckoo Search Algorithm" which was described by authors in [16]. This metaheuristic is an iterative stochastic method for solving many optimization problems. This method has been very successful in the optimization community; its good performance in different applications and the possibility of hybridization with other metaheuristics have contributed to this craze. In particular, this algorithm is based on the Cuckoo Search, which is inspired by the fascinating life style, habitat and reproduction of a bird species called cuckoo. It is also based on the parasitic behavior of this species combined with levitating flight-like movement logic specific to certain bird and fly species.

In the literature, there are several metaheuristic algorithms that have been used in the field of image segmentation, but the reasons for the choice of the cuckoo search are due to the use of two fundamental mechanisms:

- Intensification, which refers to the exploitation close to the optimal solution found.
- Diversification, which refers to the efficient exploration of the totality of the research field.

To solve the problem of image segmentation and improve the quality of segmented images for use in various applications. In this paper, we propose a new image segmentation method based on the hybridization of FCM clustering and CSA algorithm, which focuses on the issue of finding the optimal cluster centers in the first step and starting the FCM clustering operation in the second step. Our method can not only search for the optimal solution in the global range, but it can also exercise the accuracy of the local optimization ability of FCM algorithm. We also compared the proposed technique with other existing clustering-based segmentation algorithms, such as FCM-GA and PSO-GA. The experimental results showed the efficiency of our hybrid algorithm on the

different types of images used in our work and proved its performance by making a visual and statistical analysis of the different results obtained.

The fundamental contributions of this document can be mentioned by the following points:

- The creation of a clustering approach based on the cooperation between the FCM and CSA algorithms.
- The fixation of the number of clusters k and consequently of the center of each cluster.
- The use of CSA operators to generate the initial centers and then the FCM starts with the generated centroids.
- The proposed method has been tested by several evaluation criteria well recognized in this field.
- The results obtained confirmed the robustness and performance of the proposed method compared to other algorithms.

In the next sections, we will first present a description of the workings of the CSA and FCM algorithms, in order to exploit the advantages of CSA to optimize the FCM clustering problem. In addition, the results achieved by our approach will be compared to other well-known image segmentation methods in the literature.

The structure of this document is as follows: Section 2 lists the related work. Section 3 presents the background used. The description of our approach is provided in Section 4, while the discussion of the obtained results is presented in Section 5. Last, the conclusion is displayed in Section 6.

II. RELATED WORK

Data mining refers to the set of algorithms and methods used to explore and analyze large computer databases in order to detect in these data: unknown rules, associations and trends (not fixed a priori) and particular structures, restoring in a concise way, the essential information useful for decision support. It uses advanced statistical methods such as data partitioning (gathering data in homogeneous packets), and regularly employs artificial intelligence mechanisms or neural networks. In other words, clustering allows to group objects with similar properties into several homogeneous classes so that the intersection of the formed classes in pairs gives an empty set and the union of all classes gives the initial data set. Note that the degree of overlap between classes and the multidimensional nature are the most important difficulties in solving a classification problem. Knowing that, data elements from different clusters have minimal similarity [17]. Clustering is classified as the main unsupervised learning problem; thus, the clustering process can be hard or fuzzy. The hard method assigns each object a single label, e.g., K-means is the most popular classification technique [18] for hard clustering; while in fuzzy classification, an object can simultaneously belong to several classes [19], e.g., FCM which is widely used for image segmentation by fuzzy classification [20]. Fuzzy methods can be easily converted into hard methods.

The use of FCM standard for image segmentation has limited performance because the result is strongly dependent

on the initial cluster centers. As a result, the algorithm quite often falls into locally optimal solutions and misses global solutions. Another disadvantage of FCM is its high sensitivity to image artifacts, such as noise and intensity inhomogeneity. In the literature, many bio-inspired techniques, such as Algorithm (GA), Whale Optimization Algorithm (WOA), Ant Colony Optimization (ACO), Differential Evolution and also PSO, were proposed in addition to FCM to reduce its weaknesses [21].

Recently, other metaheuristic approaches have been employed to address several optimization problems and can open new perspectives and improve image segmentation. Some of these approaches are:

In [22], the author described an algorithm for fireflies based on fuzzy classification. This algorithm has two phases. In the first phase, an optimal value is identified of the number of predetermined clusters, and then the result of the first step is input to the FCM algorithm to perform the cluster segmentation operation. The results obtained show promising results compared to the traditional FCM algorithm.

In [23], the author introduced a new method for liver segmentation using the whale optimization algorithm (WOA). The proposed technique starts by dividing the image into a predefined number of classes. The clustering process of this method converts the prepared image into a binary image and after multiplication by the WOA segmented image. This technique is tested using a database of MRI images. The results demonstrate the robustness of the technique suggested by the authors.

Based on a metaheuristic algorithm called Grey Wolf Optimizer, the authors in [24], proposed a new algorithm for satellite image segmentation. This algorithm has been modified to work as an automatic clustering algorithm. This technique has been evaluated on satellite images and shows an efficient accuracy with a shorter computation time.

In [25], the author used a clustering strategy for fish image segmentation using the Salp Swarm Algorithm (SSA). This method is used to cluster the image pixels to produce compact and quasi uniform super pixels. The results of the experiments conducted by the proposed model show the performance and efficiency for different cases compared to the work.

In this work, we introduce a new clustering-based approach to image segmentation. This technique is performed by hybridizing the FCM algorithm and the CSA algorithm which was proposed by the researchers in [16]. The CSA is a recent optimization algorithm based on artificial intelligence, which has shown its robustness and efficiency on a large number of optimization problems. Many researchers have proven the efficiency of Yang's CS algorithm in different applications, such as face recognition [26], neural network training [27] and engineering design [28]. The CSA has also been used in clustering problems and as examples we can cite [29]. Although the CS algorithm is simple and very efficient and also has few parameters, it sometimes falls into the local optimum during the search. Therefore, many researchers have been working to improve the performance of this algorithm, and thus they have proposed improved versions of the CS

algorithm [30, 31]. In our paper, a hybrid algorithm between CSA and FCM is introduced for image segmentation using clustering technique. On the proposed method, the initial step size is randomly calculated without being designed in advance. In addition, to reduce local extremes and improve the variety of cuckoos, the pa value changes non-linearly with iterations. To judge the efficiency and the performances of our proposal, we have tested it on several images of different types and compared it by several classical clustering algorithms or metaheuristics.

III. BACKGROUND

A. Fuzzy C-Means Algorithm

FCM is a clustering technique, developed by Bezdek in 1981. In image processing, FCM consists in finding the exact membership of a pixel to a cluster. Each pixel is initially assigned a value that corresponds to its degree of membership in each cluster. This degree varies between 0 and 1: this is the fuzzification. We apply the chosen fuzzy rule; this rule manages the defuzzification of the system by assigning each pixel to a single class, namely the one to which it has the highest degree of membership. The concept of this operation is as follows: Each of the N pixels belongs to each of the C classes with a membership coefficient U ; the set of membership degrees is stored in the FCM matrix U . This algorithm is often used in fuzzy image segmentation.

- Principle of the FCM algorithm

The FCM algorithm [32] is a fuzzy segmentation technique applicable to different types of images. To partition the image, we need to minimize the criterion of the sum of intra-class distances generalized to the fuzzy case and given by the following formula:

$$J_{FCM}(V, U, X) = \sum_{k=1}^K \sum_{i=1}^N U_{ki}^m d^2(x_i, v_k) \quad (1)$$

Under the following constraints:

$$0 < \sum_{i=1}^N U_{ki} < N \quad (2)$$

$$\sum_{i=1}^K U_{ki} = 1 \quad (3)$$

Where:

- $m \in]1, +\infty[$ is a parameter that characterizes the degree of fuzziness,
- K : represents the number of classes,
- N : represents the number of pixels to be classified,
- V is the feature vector of the center of gravity of the index class K .
- $d(x_i, v_k)$ characterizes the distance between a pixel x_i and the center of gravity of the class v_k

$d(x_i, v_k)$ is the Euclidean distance given by the following formula:

$$d(x_i, v_k) = \sqrt{\sum_{j=1}^D (x_{ij} - v_{kj})^2} \quad (4)$$

The basic idea of FCM classification is to assign to each vector x_i a degree of membership u_{ki} , to each class centered in v_k . The algorithm minimizes a certain error between classes by iteratively computing the degree of membership and the class centers using previously denoted relations. The update v_k and u_{ki} presented by the following expressions:

$$u_{ki} = \sum_{l=1}^K \left(\frac{\|x_i - v_k\|}{\|x_i - v_l\|} \right)^{\frac{-2}{m-1}} \quad (5)$$

The function to update the centers is:

$$v_k = \frac{\sum_{i=1}^N u_{ki}^m x_i}{\sum_{i=1}^N u_{ki}^m} \quad (6)$$

The FCM is based on the update of the membership function during the iteration of the algorithm. The FCM thus makes the partition examine by minimizing the fitness function J_{FCM}

The FCM algorithm is as follows:

Algorithm 1: FCM

Fix: $c, m, \text{iterMax}$ and stop criterion ε

Initialize randomly the cluster centers $v_k, k=1, \dots, c$.

For $t \leftarrow 1$ to iterMax **do**

 Update the membership function u_{ki} according Eq. (5)

 Compute the cluster centers v_k with the formula Eq. (4)

 Compute the objective function according Eq. (1)

if $|J^t - J^{t+1}| < \varepsilon$ **then**

break

end if

end for

- Cluster validity Indices

Verification of the results obtained by the clustering algorithm is an essential part of the clustering process. The most important method of cluster validation is based on internal cluster validity indicators. Clustering will be good if the clusters are maximally separated from each other and if the objects within the clusters are increasingly close (compact) to the center of gravity. Thus, this operation separates data objects into different clusters with the goal of maximizing intra-cluster similarity and minimizing inter-cluster similarity. To evaluate the quality of the partitions of clustering algorithms, we will use the validity indices which are numerous and very well known in the literature. In our paper, we will implement two indices [33] to examine the new objective function used in this paper, which are:

- The Subarea Coefficient (SC) measures the ratio of the sum of cluster compactness and cluster separation:

$$SC = \sum_{i=1}^c \frac{\sum_{k=1}^n (u_{ik})^m \|x_k - v_i\|^2}{n_i \sum_{j=1}^c \|v_j - v_i\|^2} \quad (7)$$

- The Partition Coefficient (PC) determines the amount of overlap between clusters:

$$PC = \frac{1}{n} \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^2 \quad (8)$$

A clustering approach is considered better and efficient if the PC values are high while the SC values are low.

B. Cuckoo Search Algorithm

The metaheuristic methods are a new generation of powerful and general approximate methods that consist of a set of fundamental concepts. Among of which we find the Cuckoo Search algorithm which is a very recent meta-heuristic, inspired by the parasitism of cuckoo birds by laying their eggs in the nests of other birds, other species created by the authors Yang and Deb in 2009. This algorithm aims at breeding high quality solutions for optimization problems. In Cuckoo Search, an egg refers to a solution of the optimization problem at hand. A cuckoo egg refers to a solution just generated, and a nest means a set of possible solutions. It is based on the aggressive cuckoo breeding strategy complemented by a behavior called Levy flights [34]. The latter is a class of random walks in which the jumps are determined according to the Levy distribution which is based on a power law with infinite variance and a mean of the type [35].

$$Lévy(\beta) \sim y = X^\beta, (1 < \beta \leq 3) \quad (9)$$

The use of the Levy flight by CSA optimizes the search, this process is carried out as follows: the new solutions are generated by a random walk of Levy around the best solution obtained until now, which accelerates the global search.

The fitness function is a function that gives each solution in the search space a numerical value to show its quality. In our treatment, a better quality nest will give us access to new generations. Thus, the quality of a cuckoo egg is necessarily related to its ability to produce a new cuckoo.

Yang and Deb incorporated the Levy flight present in relation (10) to get a new solution $X(t+1)$ generated at each cuckoo i :

$$X_i(t+1) = X_i(t) + \alpha \oplus Lévy(\lambda) \quad (10)$$

Where $\alpha > 0$ is the displacement step size.

Generally, $\alpha = 1$.

The Levy flight represents a random walk whose random steps are defined from the Levy distribution given in the following equation:

$$Levy \sim u = t^{-\lambda} \quad (11)$$

In general, the CSA steps are summarized in Algorithm 2 presented as follows:

Algorithm 2: CSA

Input: Nest population $x_i=(x_{i1}, \dots, x_{iD})$ T for $i= 1, \dots, Np$

Output: Best Solution ($xbest$)

Begin

Generate initial population of host nests ()

eval=0

While (*stopping condition not met*) **do**

For $i=1: Np$ **do**

if ($f_j < f_i$) **then**

$x_i = x_j$;

$f_i = f_j$;

end

if ($\text{rand}(0,1) < pa$) **then**

 Init_nid ($xWorst$);

end

if ($f_i < f_{min}$) **then**

$xbest = x_i$;

$f_{min} = f_i$;

end

end for

end while

end begin

IV. THE PROPOSED METHOD

This paper proposes a new image segmentation technique using CSA and FCM algorithms. The CSA has strong overall optimization capability and a hybridization of FCM and CSA has improved performance over traditional FCM clustering. In the approach we propose, the CSA is used to find the optimal clustering of data taking into account a new objective function of having initial cluster centers. Then the centers found by CSA are used as input for the FCM algorithm. The process of the proposed technique treats the populations as host nests and then the sum of the population gives a better solution to each generation. In other words, the better cuckoo egg is considered to be the optimal solution and will be passed on to the next generation. In our paper, the image segmentation technique is performed on a clustering method that forcefully depends on the cluster center. However, the processing is initiated by generating the cluster centers randomly and after integrating the CSA to refine the location of the class center. Then, CSA updates the class centers to minimize the fitness function of FCM to find near-optimal centers.

A. Fitness Function

The new objective function proposed in this paper to evaluate the quality of clustering results is presented as follows:

$$fitness = \frac{intra_{cluster} + SC}{PC} \quad (12)$$

The parameters of this function are presented by:

- SC is the subarea coefficient determined by equation (7),
- PC is the partition coefficient presented in equation (8).
- The $intra_cluster$ [36] is computed using the equation given below:

$$intra_{cluster} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^c \|x_i - z_j\|^2 \quad (13)$$

The primary objective of our proposed method is to provide a cooperative technique to globally improve the performance of image segmentation results and overcome the limitations of FCM alone. For this purpose, we use the CSA algorithm to minimize the function shown in equation (12) to get the near-optimal initial cluster centers. Then, these centers are applied as initial inputs to the FCM. However, for the fitness function to be minimized, it must have the term value ($intra_cluster + SC$) low and the parameter value PC high.

The proposed approach can be summarized in the following points:

- The CSA finds the near-optimal centroids, after this stage, FCM algorithm operation starts with these cluster centers generated by CSA.
- The performance of the image segmentation results is evaluated by a fitness function given by equation (12).
- The role of CSA is to search for the optimized centroids
- The use of FCM on the input image by the optimal centers generated by CSA.
- The clustering is done by merging the results and gives the final segmented image.

The main steps of the hybrid algorithm of our method are presented in Fig. 1.

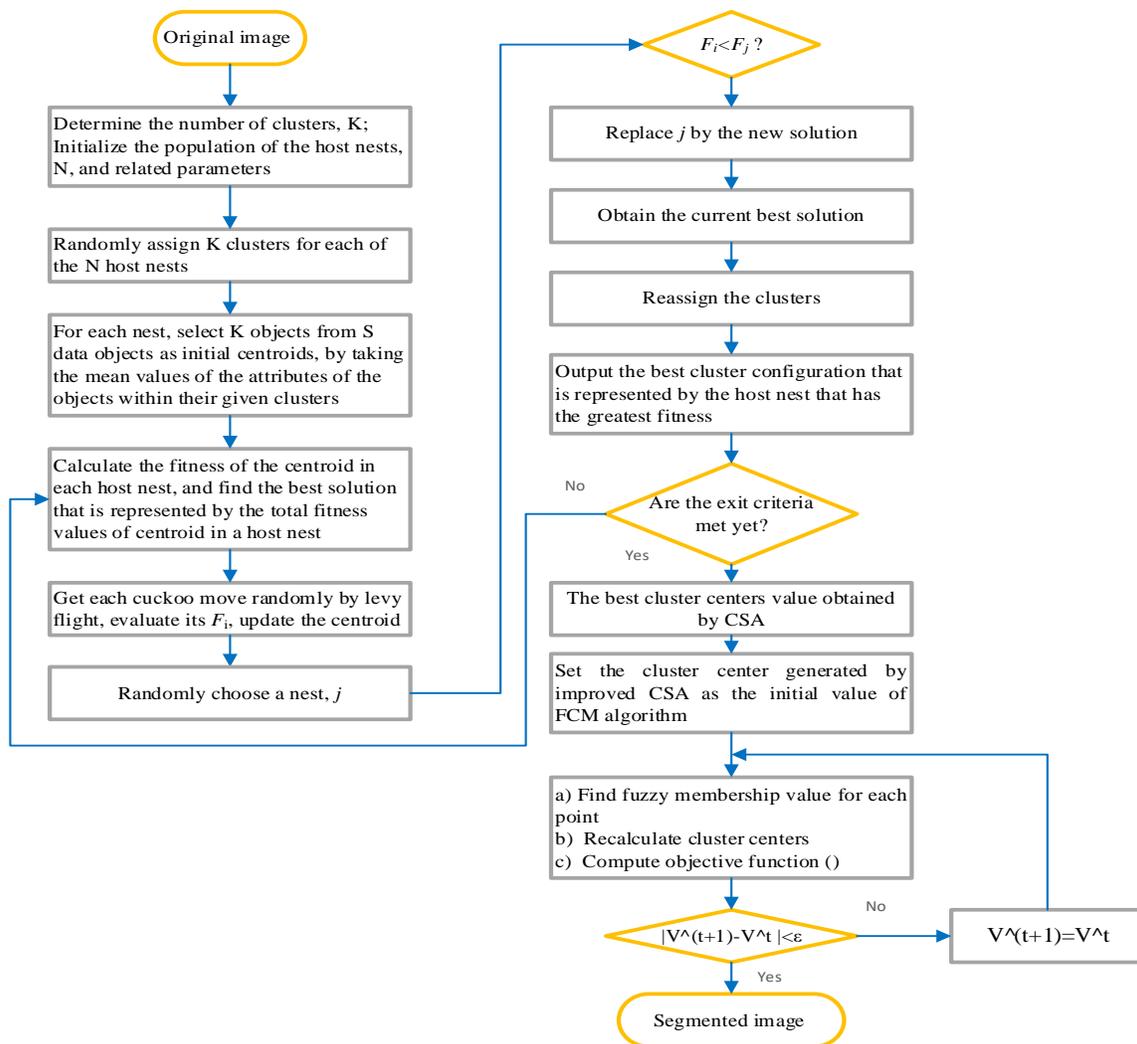


Fig. 1. The Clustering Flowchart of the Proposed Approach.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

For a better segmentation of images, it is necessary to minimize the fitness function. The minimum value of fitness corresponds to segmentation with a minimum distance between pixels belonging to the same region. So to evaluate the proposed approach and better experiment the performance and robustness in image processing, especially image segmentation, we have performed several tests on different reference images. Furthermore, we compared the proposed method with other existing clustering-based segmentation techniques that perform well, such as: FCM based on genetic algorithms [37], FCM based on particle swarm optimization [38] and the standard FCM algorithm. The algorithms used in the experiment are implemented in the MATLAB 2014b platform and run on a computer containing the following configuration: a 4th generation Intel Core (TM) i5 processor at 2.5 GHz, 4G of RAM and running Microsoft Windows 10 64 bits. The effectiveness of various image clustering approaches is analyzed and discussed by different evaluation indices to examine the quality of image segmentation. These are mean square error (MSE) [39], peak signal to noise ratio

(PNSR)[40], Rand Index (RI)[41], global coherence error (GCE)[42], boundary displacement error (BDE)[42], information of variation (VOI)[42], and correlation coefficient (CC)[43].

The FCM, FCM-PSO, FCM-GA, FCM-CSA algorithms are implemented in their original versions. Thus, the parameters have to be adjusted for each algorithm, in order to get the best matching values that can produce good image segmentation results with a short execution time.

First, we perform a series of experiments based on the modification of the number of clusters k , to search for good image segmentation results based on the evaluation parameters mentioned above. Then and in order to optimize the results obtained by CSA, we will apply the fraction $P_a = 0.25$ which allows us to have the optimal solution. In order to approach the best image segmentation, we followed practically to choose the value of each parameter. The experiments show that the choice of cluster number k is influential on the quality of the segmented image i.e. the choice of k is dependent on the image to be segmented; therefore, to present the performance of the proposed technique and the measures of the evaluation criteria

of the algorithms used in this article, we will focus on the choice of cluster number k which is equal to 4 on several images selected from the BSD300 database [44]. Table I shows the best values of the parameters that were optimized for the algorithms used in this paper (npop is the population size and MaxIt is the number of iterations).

TABLE I. THE VALUES OF THE PARAMETERS USED IN THE EXPERIMENT

Algorithm	Parameter	Value
GA	pc, pm	0.7, 0.3
	Mutation rate	0.02
	npop, MaxIt	50, 100
PSO	$C_1, C_2, V_{max}, V_{min}$	1.49, 1.49, 3, -3
	npop, MaxIt	50, 100
	Pa	0.25
CSA	npop, MaxIt	50, 100

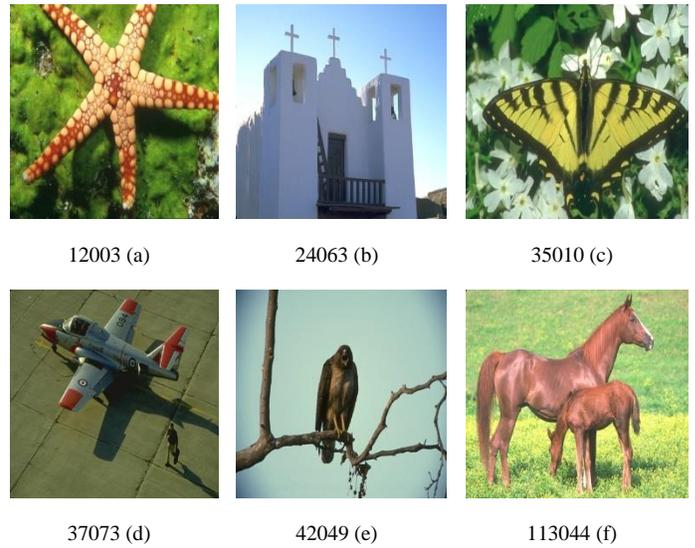


Fig. 2 shows reference images that are selected from the Berkeley 300 database (BSD300).

The image segmentation results of the different used algorithms are shown in Fig. 3.

Fig. 2. Examples of Original Images of the Berkeley 300 (BSD300) Segmentation Base.

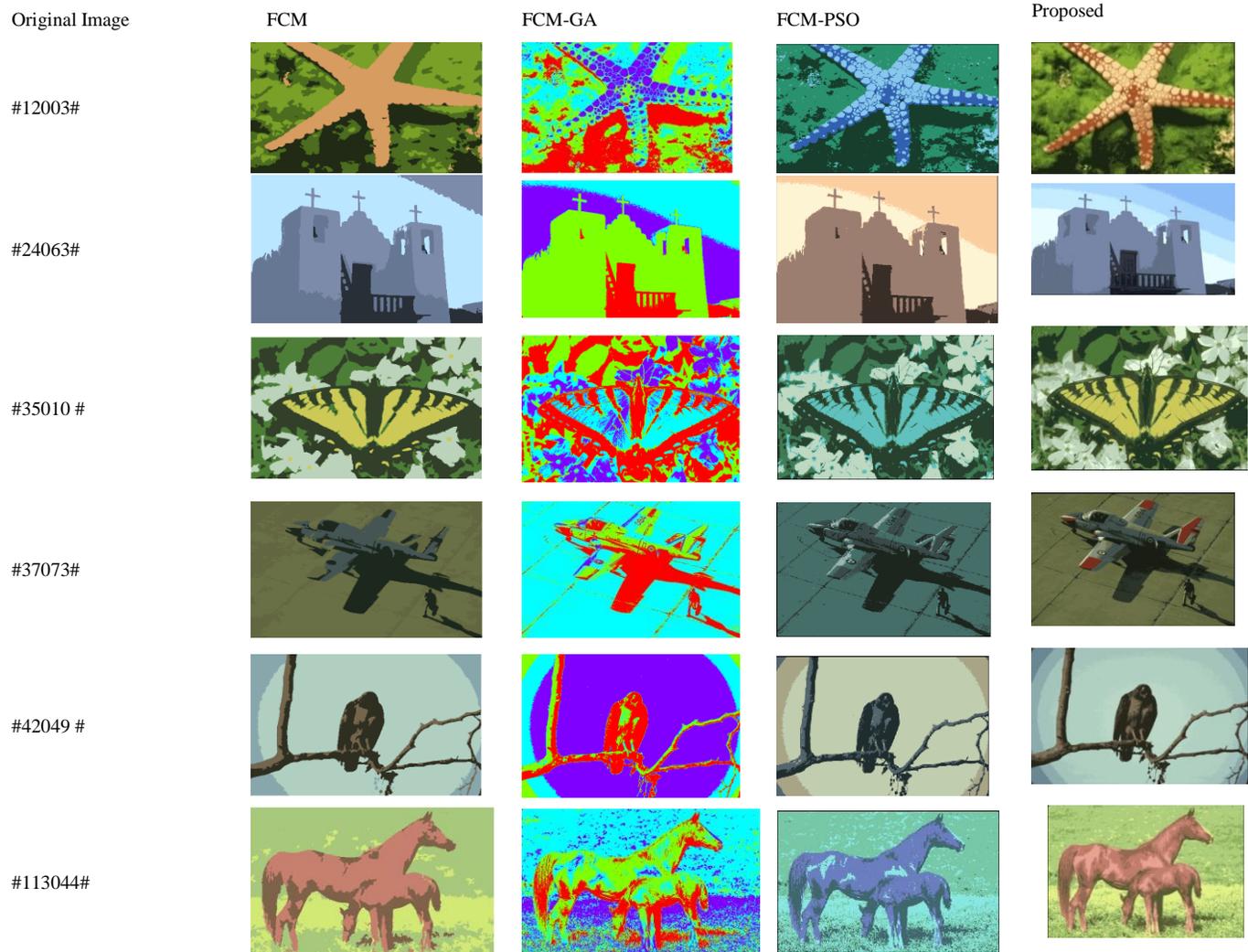


Fig. 3. Experimental Results of Clustering using different Algorithms.

The results obtained by the various techniques are analyzed and discussed on the basis of the excellent fitness values, MSE, PSNR, CC, RI, GCE, BDE and VOI measures.

A clustering technique is considered to be effective and good performance to evaluate the result of the segmented image if the PSNR metric value is large and the MSE value is small as well as the CC parameter value is high. The MSE, PSNR and CC parameter values of the segmented images are measured by the algorithms used in this paper. From the results shown in Fig. 4, 5 and 6, we can see that the MSE values obtained by the proposed approach are very small. On the other hand, the PSNR values obtained by our method are very high, while the correlation coefficient values are high, which clearly show that the proposed approach with the use of the objective function proposed in this paper, can generate correct segmentation results compared to other comparison algorithms.

According to the obtained results, we can conclude that the proposed hybrid algorithm shows good performance and gives better results, because the image segmented by the proposed approach generates well detailed segmentation results, the different regions of the image are visible.

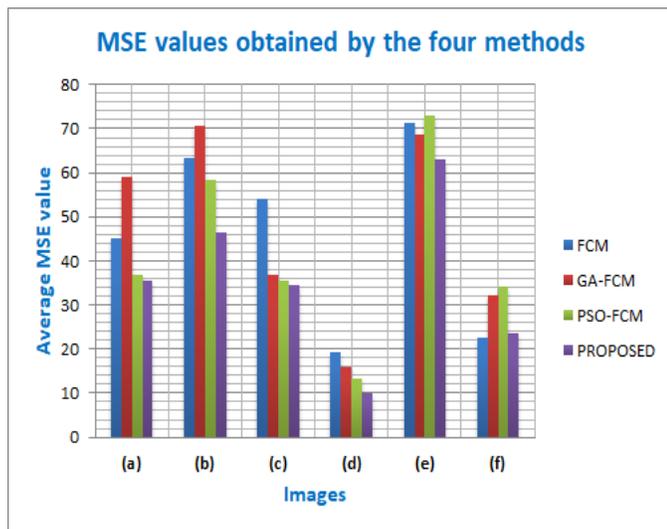


Fig. 4. MSE Values Obtained by the Four Methods.

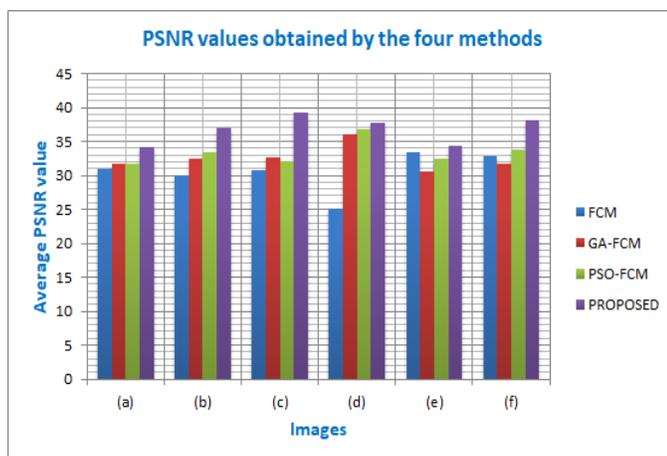


Fig. 5. PSNR Values Obtained by the Four Methods.

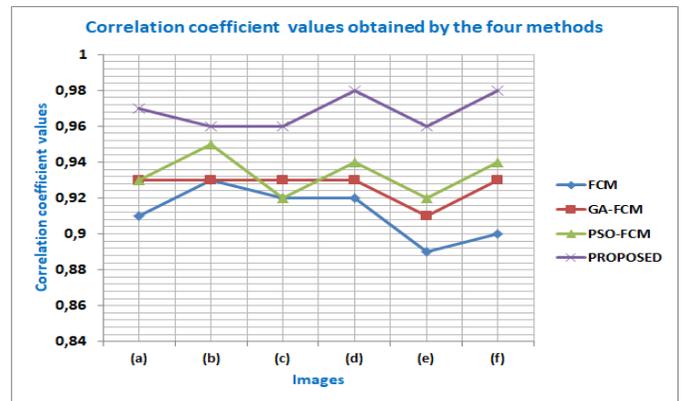


Fig. 6. Correlation Coefficient Values Obtained by the Four Methods.

In order to examine and present the effectiveness of the proposed technique, we compared the segmentation results obtained by different test images with all the algorithms used in this paper. We also evaluated the performance of the segmented image results using four well-recognized image segmentation evaluation indexes in the literature: PRI, VOI, GCE, and BDE which are mentioned earlier. Furthermore, the experiments show that the segmented image result is of good quality and closer to the ground truth, if the value of RI is larger, and the values of VOI, GCE and BDE are smaller.

According to the results displayed in Fig. 2, we can say that our approach gives better results compared to other methods, knowing that each of the segmented image results is related to its content and the number of classes we choose. For the comparison experiments, the value of the parameter K (number of clusters) was changed several times for the segmentation of different images. From the experiments performed on the different algorithms, we can see that the objects in each image can be identified or not depending on the image content and the choice of K. And according to these experiments, we chose the number of clusters equal to 4 for all the selected images in the Berkeley 300 database in order to properly present the performance of our approach and clearly visualize the quality of the segmented image, as well as the measures of the evaluation indices of the methods used.

In the BSD300 database, each image corresponds to several field truth segmentations, which leads to a segmentation result corresponding to several performance index groups. Therefore, the average value of several performance index groups is generally considered as the final performance index of the segmentation result.

Table II shows a comparison of the various evaluation criteria values using the clustering method on the different techniques used in this paper for each test image. From this table, we found that the performance varies depending on the image. But our method of cooperation between FCM and CSA achieves better results than the other algorithms. It is clear that the proposed approach generates well-detailed segmentation results; the different regions of the image are visible and gives satisfactory results as it obtains the best values of RI, VOI, BDE and GCE.

TABLE II. STATISTICAL PARAMETERS OF IMAGE SEGMENTATIONS. THE BEST VALUES ARE IN BOLD

	METRIC	FCM	FCM-GA	FCM-PSO	PROPOSED
(a)	RI	0.6992	0.7531	0.7910	0.8673
	VOI	3.3652	2.4731	2.5537	1.8596
	GCE	0.4320	0.3378	0.3051	0.2434
	BDE	14.4953	14.3853	13.2981	12.2113
(b)	RI	0.7419	0.7531	0.7542	0.7875
	VOI	2.9940	2.5493	2.4433	2.1532
	GCE	0.2315	0.2278	0.2015	0.2064
	BDE	13.7532	13.2016	12.1931	12.0124
(c)	RI	0.7431	0.7415	0.7476	0.7882
	VOI	2.8315	2.8003	2.5821	2.0020
	GCE	0.4153	0.4051	0.3451	0.2603
	BDE	14.3156	14.0251	12.7328	12.0031
(d)	RI	0.6851	0.7631	0.7542	0.7558
	VOI	3.0420	2.8749	2.2074	2.1306
	GCE	0.5003	0.4207	0.4008	0.3120
	BDE	12.7892	12.7112	12.1005	12.1106
(e)	RI	0.7427	0.7462	0.7752	0.7882
	VOI	2.2732	2.2653	2.2124	2.2007
	GCE	0.4167	0.4054	0.3921	0.2521
	BDE	13.9821	12.7629	12.2368	12.0701
(f)	RI	0.7591	0.7698	0.7834	0.7899
	VOI	2.6210	2.3521	2.2750	2.0193
	GCE	0.3281	0.2872	0.2645	2.5698
	BDE	13.9823	12.8902	12.3482	12.331

We also note that the values of RI, VOI, GCE and BDE obtained by our technique are better than those obtained by the other techniques. In detail, we notice that the values of VOI, GCE and BDE of our algorithm are smaller, and the RI value is larger than that obtained by the other methods.

Based on the results of the statistical calculations presented in the previous Fig. 3, 4 and 5 and the values of the parameters of the evaluation indices, indicated in Table II applying the different image segmentation techniques used in this paper, it can be seen that the quality of the segmentation image varies from one method to another depending on the optimization algorithm used to improve the classical FCM method. In summary, the cohesion within clusters is very high by our clustering technique compared to other clustering methods. The clustering technique used in this paper which is based on the FCM optimized by CSA gives good values in terms of cluster quality measures according to the experimental results. Therefore, the detailed analysis of these results on several reference and real images shows the robustness and high efficiency of our method in terms of accuracy and reliability.

VI. CONCLUSION

FCM is the most widely used clustering algorithm in classification problems, especially in image segmentation because it is efficient and simple. However, FCM has the limitation of being sensitive to prior values and often falls into local optima. To overcome this drawback we proposed a new

image segmentation method that relies on the optimization of segmentation by cuckoo search. CSA has a strong global optimization capability and hybridization of FCM with CSA will give an increased performance compared to traditional FCM clustering. Our method has been used on various images, and despite their complexity, the segmentation performed by FCM gives quite good results, and with the help of CSA, it makes a jump and gives us the optimal solution. The performance of the method has been evaluated based on the best values of the cluster evaluation indices and the values of the fitness function used in this paper. We also compared the proposed technique with other existing clustering-based segmentation algorithms such as FCM-GA and PSO-GA. The results indicate that a perfect initialization of the classes gives better results by the proposed algorithm. The experimental results showed the efficiency of our method on the different types of images used in our work and proved its robustness by making a visual analysis of the different results obtained.

Nevertheless, our approach requires the knowledge of the number of classes and it relies on the Euclidean distance to measure the similarity between an observation and the center of a class which makes it usable only to detect spherical classes. To overcome these drawbacks, I propose as a perspective of this work, to apply other hybrid methods based on recent metaheuristics for image segmentation in order to improve the quality of classification and reduce the execution time.

REFERENCES

- [1] L. KHRISSI, N. EL AKKAD, H. SATORI, K. SATORI, "Simple and Efficient Clustering Approach Based on Cuckoo Search Algorithm," 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), 1–6, 2020, doi:10.1109/ICDS50568.2020.9268754.
- [2] L. Khriissi, N. El Akkad, H. Satori, K. Satori, Image Segmentation Based on K-means and Genetic Algorithms, 2020, doi:10.1007/978-981-15-0947-6_46.
- [3] L. Yanling, S. Yi, "KNN-based mean shift algorithm for image segmentation," Journal of Huazhong University of Science and Technology (Nature Science Edition), 10, 2009.
- [4] A. El Idrissi, Y. El Merabet, Y. Ruichek, R. Touahni, A. Sbihi, C. Meurie, A. Moussa, "A Multiple-Objects Recognition Method Based on Region Similarity Measures: Application to Roof Extraction from Orthophotoplans," IJACSA) International Journal of Advanced Computer Science and Applications, 6(11), 2015.
- [5] J. SHEN, H. GUO, H. LI, F. QIU, Z. ZHANG, "A Water Edge Extraction Method from Images Based on Canny Operator and GAC Model [J]," Journal of Geomatics Science and Technology, 30(3), 264–268, 2013.
- [6] L. Khriissi, N.E. Akkad, H. Satori, K. Satori, "Color image segmentation based on hybridization between Canny and k-means," in 7th Mediterranean Congress of Telecommunications 2019, CMT 2019, 2019, doi:10.1109/CMT.2019.8931358.
- [7] M.H. Hesamian, W. Jia, X. He, P. Kennedy, "Deep learning techniques for medical image segmentation: achievements and challenges," Journal of Digital Imaging, 32(4), 582–596, 2019.
- [8] N. El Akkad, M. Merras, A. Baataoui, A. Saaidi, K. Satori, "Camera self-calibration having the varying parameters and based on homography of the plane at infinity," Multimedia Tools and Applications, 77(11), 14055–14075, 2018.
- [9] M.D. Hossain, D. Chen, "Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," ISPRS Journal of Photogrammetry and Remote Sensing, 150, 115–134, 2019.
- [10] X. Yao, Y. Liu, K.-H. Liang, G. Lin, Fast evolutionary algorithms, Springer: 45–94, 2003.

- [11] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, 4(2), 65–85, 1994.
- [12] J. Kennedy, R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks, 1942–1948, 1995*.
- [13] S. Karthikeyan, T. Christopher, "A hybrid clustering approach using artificial bee colony (ABC) and particle swarm optimization," *International Journal of Computer Applications*, 100(15), 2014.
- [14] L. Khrissi, N. El Akkad, H. Satori, K. Satori, "Clustering method and sine cosine algorithm for image segmentation," *Evolutionary Intelligence*, 2021, doi:10.1007/s12065-020-00544-z.
- [15] A. Arjmand, S. Meshgini, R. Afrouzian, A. Farzammia, "Breast Tumor Segmentation Using K-Means Clustering and Cuckoo Search Optimization," in *2019 9th International Conference on Computer and Knowledge Engineering (ICCKE)*, 305–308, 2019.
- [16] X.-S. Yang, S. Deb, "Cuckoo search via Lévy flights," in *2009 World congress on nature & biologically inspired computing (NaBIC)*, 210–214, 2009.
- [17] A.K. Jain, R.C. Dubes, *Algorithms for clustering data*, Prentice-Hall, Inc., 1988.
- [18] J. MacQueen, others, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 281–297, 1967.
- [19] M.A. Sasa, L. Dongqing, X. Jia, F. Xingqiao, "Research on continuous function optimization algorithm based on swarm-intelligence," in *2009 Fifth International Conference on Natural Computation*, 61–65, 2009.
- [20] C. Li, L. Liu, X. Sun, J. Zhao, J. Yin, "Image segmentation based on fuzzy clustering with cellular automata and features weighting," *EURASIP Journal on Image and Video Processing*, 2019(1), 1–11, 2019.
- [21] T.M. Silva Filho, B.A. Pimentel, R.M.C.R. Souza, A.L.I. Oliveira, "Hybrid methods for fuzzy clustering based on fuzzy c-means and improved particle swarm optimization," *Expert Systems with Applications*, 42(17–18), 6315–6328, 2015.
- [22] W.K. Alomoush, S.N.H.S. Abdullah, S. Sahran, R.I. Hussain, "Segmentation of MRI brain images using FCM improved by firefly algorithms," *Journal of Applied Sciences*, 14(1), 66–71, 2014.
- [23] A. Mostafa, A.E. Hassanien, M. Houseni, H. Hefny, "Liver segmentation in MRI images based on whale optimization algorithm," *Multimedia Tools and Applications*, 76(23), 24931–24954, 2017.
- [24] S. Kapoor, I. Zeya, C. Singhal, S.J. Nanda, "A grey wolf optimizer based automatic clustering algorithm for satellite image segmentation," *Procedia Computer Science*, 115, 415–422, 2017.
- [25] A. Ibrahim, A. Ahmed, S. Hussein, A.E. Hassanien, "Fish image segmentation using salp swarm algorithm," in *International Conference on advanced machine learning technologies and applications*, 42–51, 2018.
- [26] V. Tiwari, "Face recognition based on cuckoo search algorithm," *Image*, 7(8), 9, 2012.
- [27] X.-S. Yang, S. Deb, "Cuckoo search: recent advances and applications," *Neural Computing and Applications*, 24(1), 169–174, 2014.
- [28] X.-S. Yang, S. Deb, "Engineering optimisation by cuckoo search," *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(4), 330–343, 2010.
- [29] P. Manikandan, S. Selvarajan, "Data clustering using cuckoo search algorithm (CSA)," in *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012)*, December 28–30, 2012, 1275–1283, 2014.
- [30] M.K. Naik, R. Panda, "A novel adaptive cuckoo search algorithm for intrinsic discriminant analysis based face recognition," *Applied Soft Computing*, 38, 661–675, 2016, doi:https://doi.org/10.1016/j.asoc.2015.10.039.
- [31] S. Suresh, S. Lal, "An efficient cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions," *Expert Systems with Applications*, 58, 184–209, 2016, doi:https://doi.org/10.1016/j.eswa.2016.03.032.
- [32] J.C. Bezdek, C. Coray, R. Gunderson, J. Watson, "Detection and characterization of cluster substructure i. linear structure: Fuzzy c-lines," *SIAM Journal on Applied Mathematics*, 40(2), 339–357, 1981.
- [33] B. Balasko, J. Abonyi, B. Feil, "Fuzzy clustering and data analysis toolbox," Department of Process Engineering, University of Veszprem, Veszprem, 2005.
- [34] C.T. Brown, L.S. Liebovitch, R. Glendon, "Lévy flights in Dobe Ju'hoansi foraging patterns," *Human Ecology*, 35(1), 129–138, 2007.
- [35] P. Feyel, *Optimisation des correcteurs par les métaheuristiques. Application à la stabilisation inertielle de ligne de visée.*, CentraleSupélec, 2015.
- [36] J. Ye, Z. Zhao, H. Liu, "Adaptive Distance Metric Learning for Clustering," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–7, 2007, doi:10.1109/CVPR.2007.383103.
- [37] A. Halder, S. Pramanik, A. Kar, "Dynamic Image Segmentation using Fuzzy CMeans based Genetic Algorithm," *International Journal of Computer Applications*, 28(6), 15–20, 2011, doi:10.5120/3392-4714.
- [38] F. Zhao, Y. Chen, H. Liu, J. Fan, "Alternate PSO-Based Adaptive Interval Type-2 Intuitionistic Fuzzy C-Means Clustering Algorithm for Color Image Segmentation," *IEEE Access*, 7, 64028–64039, 2019, doi:10.1109/ACCESS.2019.2916894.
- [39] T. Chai, R.R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, 7(3), 1247–1250, 2014, doi:10.5194/gmd-7-1247-2014.
- [40] Z. Liu, R. Laganière, "Phase congruence measurement for image similarity assessment," *Pattern Recognition Letters*, 28(1), 166–172, 2007, doi:https://doi.org/10.1016/j.patrec.2006.06.019.
- [41] R. Unnikrishnan, C. Pantofaru, M. Hebert, "Toward Objective Evaluation of Image Segmentation Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6), 929–944, 2007, doi:10.1109/TPAMI.2007.1046.
- [42] B. Peng, L. Zhang, D. Zhang, "A survey of graph theoretical approaches to image segmentation," *Pattern Recognition*, 46(3), 1020–1038, 2013, doi:10.1016/j.patcog.2012.09.015.
- [43] H. Sima, A. Mi, X. Han, S. Du, Z. Wang, J. Wang, "Hyperspectral image classification via joint sparse representation of multi-layer superpixels," *KSII Transactions on Internet and Information Systems*, 12(10), 5015–5038, 2018, doi:10.3837/tiis.2018.10.021.
- [44] D. Martin, C. Fowlkes, D. Tal, J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 416–423 vol.2, 2001, doi:10.1109/ICCV.2001.937655.

Prediction of Cantilever Retaining Wall Stability using Optimal Kernel Function of Support Vector Machine

Rohaya Alias¹

School of Civil Engineering
College of Engineering
Universiti Teknologi MARA Pahang
26400 Bandar Tun Razak Jengka,
Pahang, Malaysia

Siti Jahara Matlan²

Civil Engineering Programme,
Faculty of Engineering
Universiti Malaysia Sabah
88400 Kota Kinabalu
Sabah, Malaysia

Aniza Ibrahim³

Faculty of Civil Engineering
Universiti Pertahanan Nasional
Malaysia
57000 Kuala Lumpur
Malaysia

Abstract—The Support Vector Machine is one of the artificial intelligence techniques that can be applied to forecast the stability of cantilever retaining walls. The selection of the right Kernel function is very important so that the Support Vector Machine model can make good predictions. However, there are no general guidelines that can be used to select Kernel functionality. Therefore, the Kernel function which consists of Linear, Polynomial, Radial Basis Function and Sigmoid has been evaluated to determine the optimal Kernel function by using 10 cross-validation (V-fold cross-validation). The achievement of each function is evaluated based on the mean square error value and the squared correlation coefficient. The mean square error value is closer to zero and the squared correlation coefficient closer to the value of one indicates a more accurate Kernel function. Results show that the Support Vector Machine model with Radial Basis Function Kernel can successfully predict the stability of cantilever retaining walls with good accuracy and reliability in comparison to the various other Kernel functions.

Keywords—Cantilever retaining wall; kernel function; prediction; stability; support vector machine

I. INTRODUCTION

Cantilever retaining walls were introduced by the Burlington and Quincy Railroad in the 1880s. This retaining wall is constructed using reinforced concrete to withstand high tensile strength. It consists of two main parts namely, the walls, and the base made of reinforced concrete. Typically, the height of this wall ranges from 2.5 m to 6.0 m and is usually in the shape of either an inverted L or T. Its size is wider and flatter, and its construction cost is cheaper because the building materials are less compared to gravity walls [2]. The walls can be built on the construction site or pre-cast concrete that has been made in the factory can be used and only installed on the construction site which saves more time and energy. For heights exceeding 6 m, the use of prestressing techniques will be used. The wall part of this cantilever retaining wall is built protruding out of its large and solid site.

The cantilever retaining walls are very widely used in geotechnical engineering practice. Therefore, engineers play an important role in ensuring the construction of cantilever retaining wall is stable and safe. The stability of cantilever

retaining walls involved the checking of stability in terms of overturning, sliding, and bearing capacity. However, trainee engineers may find it difficult to get optimum stability results because of the lack of experience on the behavior of cantilever retaining walls. A prediction method using conventional mathematical models is utilized to estimate the stability of the cantilever retaining walls.

The development of studies on non-linear data analysis is growing with the revolution of the artificial intelligence (AI) methods. AI method is a great and versatile computing tool for solving complex problems, series, and its ability in identifying irregular arrangements and clusters of data. This method is popular for prediction and is able to predict the data that is non-linear, and not uniform. Among the AI methods introduced are artificial neural networks (ANN), adaptive neuro-fuzzy inference systems (ANFIS) and support vector machines (SVM).

The support vector machine (SVM) is the latest machine learning technique after neural network machine. Boser, Guyon and Vapnik introduced the SVM method in 1992 at the Annual Workshop on Computational Learning Theory. SVM is a new generation of statistical learning method which aims to recognize the data structures. It contains algorithmic learning using statistically based theories [4, 22, 15]. Learning is done by using input data and output data as the desired target this is known as supervised learning.

The SVM model has been extensively used in several fields of study for classification and prediction. Initially, SVM was developed to solve the classification problem, after which the use of SVM was extended for regression [27]. SVM regression is considered a nonparametric technique because it relies on kernel functions. According to Li et al. [6], SVM was developed as a regression device and is recognized as support vector regression (SVR). SVM is also worked to reduce overfitting and decreases the expectations of machine learning errors [26]. Smola and Scholkopf [23] stated that SVM is a method that can overcome overfitting and will result in good performance. In summary, it can be agreed that SVM is a technique that can make classifications and predictions with maximum accuracy by using the machine learning theory.

SVM method can solve regression and pattern recognition problems effectively and can also be used to make predictions and stability assessments [28]. According to Lu et al. [7], SVM has many advantages when utilized to solve small samples, nonlinear, and high-dimensional pattern recognition problems when compared to other algorithms.

II. LITERATURE REVIEW

In the last decade, the use of SVM to solve problems that cannot be solved using traditional methods in geotechnical engineering has been deeply investigated. The SVM model can improve the weaknesses found in traditional models such as ease of overfitting, single consideration factor, and inability to make predictions with long periods. Several previous studies have found that SVM is a potential method and has become the most desired in recent studies because of its ability to solve nonlinear and time series regression problems.

In the geotechnical field, reviews show that SVM is used effectively to predict soil shear strength, landslides, slope stability, deformation displacement, etc. Ly and Pham [8] proposed the use of the SVM model to predict soil shear strength using physical properties of soil as input parameters. The results of the study found that the SVM model showed good performance for soil shear strength prediction with an R value of 0.9 to 0.95. It was found that the three parameters that most influenced the prediction of soil shear strength were moisture content, liquid limit, and plastic limit. Shi et al. [21] used the SVM method to predict the deformation of the surrounding rock in the shallow-buried tunnels. The research indicates that the method of SVM produces good accuracy when utilized in making rock deformation predictions around shallow-buried tunnels. Besides, SVM is also an easy method to implement. Ramya and Vinodhkumar [18] used the SVM technique to predict the minimum factor of safety (FOS) based on upper and lower bound theorems for slope stability. The study confirmed that SVM has the capability to calculate the FOS with an adequate degree of exactness and suitable to use for the prediction of the stability of slopes. Rachel and Lakshmi [16] introduced the SVM prediction model to predict landslides by predicting rainfall data sets using big data concepts. The study summarizes that SVM is proven to be an effective technique for predicting landslides by first predicting rainfall. Samuil and Sitharam [20] compared the SVM and ANN on the liquefaction susceptibility of soil under earthquake. The research proves that SVM can produce good performance for the prediction of soil liquefaction susceptibility under earthquakes. Samui [19] investigated the effectiveness of the SVM method when compared with the ANN method to predict the frictional capacity of driven piles in clay. The results showed that SVM gave better performance than the ANN method. However, not many previous studies have been found on the use of the SVM method for predicting the stability of retaining walls. Mohamed et al. [10] used SVM to predict the external stability of segmental retaining walls reinforced with backfill with residual soil and geogrid. The results proved that the SVM based on the Radial Basis Function method and based on the specific data selection can predict the external stability factor of segmental retaining wall with a good accuracy. Cheng and Wu [3] studied the efficacy of the Evolutionary Support Vector Machine Inference Model

(ESIM) to forecast wall deformation in deep excavations. ESIM is a model that uses the SVM method and fast messy genetic algorithm (FMGA). The study found that the ESIM model successfully predicts wall deflection and deformation.

SVM provides two properties that are not found in other learning algorithms, namely the process of maximizing margins and the transformation of non-linear input space into feature space using the Kernel method [4]. Many experts have known that the capabilities of SVM are exactly correlated with Kernel function. The Kernel function transforms a data set into a hyper-plane [13]. Additionally, kernel variables need to be calculated accurately because it determined the structure of high dimensional features when the final solution was performed. Commonly, the types of Kernel functions found in the SVM model are Linear, Polynomial, Gaussian or Radial Basis Function (RBF), Laplace, Sigmoid, Exponential, etc. [9]. The Linear Kernel used when data is linearly separable. The Polynomial Kernel is well suited for problems where all the training data is normalized. The RBF and Laplace Kernel used when there is no prior knowledge about the data. The Sigmoid kernel used as the proxy for neural networks [24].

An accurate prediction model can be obtained by using the optimal Kernel functions. However, there are no general guidelines that can be used to select Kernel functionality. Based on the previous studies, it was found that most of the research conducted until recently has focused on the advantages of the SVM models. Very few studies were conducted to obtain the right Kernel function. Nanda et. al [12] conducted termite detection studies using different Kernel functions in support vector machines and found that Polynomial Kernel has produced the best accuracy. Hong et. al [5] compared the effectiveness of four Kernel functions in a support vector machine in a landslide mapping study. Findings emanated from the research indicated that the SVM-RBF model is the most suitable for landslide susceptibility assessment.

Therefore, this study intends to compare the Kernel function model between Linear, Polynomial, Radial Basis Function (RBF), and Sigmoid in their ability to predict the cantilever retaining walls stability and suggest the optimal Kernel function among them. Only four types of Kernel functions were selected since it was most frequently used compared to other Kernel functions and produced satisfactory results.

III. METHODOLOGY

The Statistica software was used to develop SVM models using data mining methods. According to Thuraisingham [25], Radhakrishnan et al. [17] and Mohammed and Wagner [11], data mining is a process of answering all questions and identifying information, forms and trends found in large quantities of data. Furthermore, in the use of data mining, there are various techniques that can be used to help make decisions. In this study, machine learning techniques have been used for SVM model making predictions. SVM model prediction is performed using input and output data. The data used for this study were 280 different designs of cantilevers retaining wall designs. The input parameters used for prediction contains walls of various heights, slope angles, and surcharges. On the

other hand, the output parameters involve the external stability of the retaining wall i.e., the safety factor (FOS) for sliding, overturning, and bearing capacity. The output parameters are applied as a target for the model prediction.

A. Kernel Function Selection

The selection of the right Kernel function is very essential so that the SVM model can make a good performance. Basically, SVM is employed as a quadratic optimization to solve the case of linear classification. Applying Kernel functions which are presented into the combined type of the optimization model can easily expand linear to nonlinear SVM via of functional mapping [14]. SVM uses a technique called the Kernel trick to provide a bridge from linear to nonlinear. The equation of Kernel trick is shown in equation (1).

$$K(x, y) = \langle f(x), f(y) \rangle \quad (1)$$

Where K is the Kernel function, x and y are n-dimensional inputs, and f is a map from n-dimension to m-dimension space.

In this study, Kernel function consisting of Linear, Polynomial, Radial Basis Function (RBF) and Sigmoid function was evaluated to determine the best Kernel function by using 10 cross-validations (V-fold cross-validation). Based on the mean square error value (MSE) and the squared correlation coefficient (R^2), the achievement of each function is determined. The MSE measures the differences in values between the target data and the predicted data in the existing predicted models. The MSE value is closer to zero and the R^2 value is closer to the value of one indicates a more accurate Kernel function. The equations of MSE and R^2 for the performance evaluation of the prediction model are shown in equation (2) and equation (3) as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2 \quad (2)$$

Where n is the total of data points, y_i is the target data, and \tilde{y}_i is the predicted data.

$$R^2 = \frac{(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

Where n is the total of predicted data, x_i is the target data, y_i is the predicted data, \bar{x} is the average of target data series, and \bar{y} is the average of predicted data series.

B. Optimum Parameters Selection

The Epsilon-RBF model was used to obtain the optimum model parameters such as gamma, capacity, epsilon, and Nu parameters. The optimal use of parameters in the Kernel function will produce an accurate prediction model [1]. The most common method used to obtain optimal model parameters is through the 10-cross-validation method (V-fold cross-validation). The cross-validation method is used in SVM to check the overfitting of data and it gives more correct predicted results.

C. Model prediction

All machine learning algorithms must be tested and validated to select those with the highest performance and prediction accuracy. After obtaining the optimal model parameters, the data set was randomly divided by 70% for the

training and 30% for the testing. To speed up training time, a selection of training data was used. Only training data close to the boundary was used. Data far from the boundary were eliminated to shorten training time.

SVM algorithm attempt to predict a target data or known as a dependent variable using features, which is the dependent variable. Prediction is implemented by mapping an independent variable to a dependent variable known as a process mapping function. The process of a mapping function for an SVM is a boundary that distinguishes two or more classes.

After getting all the prediction output, root mean squared error (RMSE) and regression square was computed for every model and compared. The root mean squared error (RMSE) can be obtained from equation (4).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (4)$$

Where n is the total of data points, y_i is the target data, and \tilde{y}_i is the predicted data.

IV. RESULT AND DISCUSSION

Kernel functions consisting of Linear, Polynomial, Radial Basis Function (RBF) and Sigmoid function were evaluated in the SVM model analysis. The results show that the Kernel function plays an important role in producing a good SVM model. The Kernel functions have the advantage of converting nonlinear input space to linear feature space. Table I and Table II show a comparison of the prediction performance of the SVM model based on MSE and R^2 values for the external stability safety factors using four different Kernel functions. The RBF kernel was found to produce the best prediction accuracy for the external stability factor of cantilever retaining wall, with the MSE value being closest to zero and the R^2 value being closest to one. This finding seems to be agreed with the study conducted by Hong et. al [5], where the RBF Kernel showed better performance than other Kernel functions. Followed next by the Linear, Polynomial, and finally the Sigmoid function which produces the least accurate prediction. Therefore, the Radial Basis Function (RBF) was chosen to be applied to the SVM model for all the external stability safety factors of the cantilever retaining wall.

TABLE I. THE PERFORMANCE OF THE SVM MODEL BASED ON MSE VALUES USING FOUR KERNEL FUNCTIONS

External Stability	Linear	Polinomial	RBF	Sigmoid
FOS for sliding	2.071	7.156	0.396	684.920
FOS for overturning	12.844	37.771	2.187	1930.534
FOS for bearing capacity	0.089	0.153	0.050	6.832

TABLE II. THE PERFORMANCE OF THE SVM MODEL BASED ON R^2 VALUES USING FOUR KERNEL FUNCTIONS

External Stability	Linear	Polinomial	RBF	Sigmoid
FOS for sliding	0.955	0.840	0.992	0.264
FOS for overturning	0.914	0.747	0.987	0.264
FOS for bearing capacity	0.708	0.478	0.716	0.479

A. SVM Model Prediction

Prediction of the cantilever retaining wall stability using the SVM model was performed by implementing the RBF Kernel and the optimal gamma, capacity, epsilon, and Nu parameters. The prediction accuracy of SVM model was compared based on the RMSE and R^2 values. Table III shows a comparison of RMSE and R^2 values for the external stability safety factors. It can be clearly observed that the prediction of the cantilever retaining wall stability using the SVM model with RBF Kernel function for safety factor of overturning is more accurate because it produces an RMSE value that is closer to zero and an R^2 value that is closer to one.

TABLE III. THE PERFORMANCE OF THE SVM MODEL USING RBF KERNEL FUNCTION BASED ON RMSE AND R^2

External Stability	RMSE	R^2
FOS for sliding	0.172	0.9985

FOS for overturning	0.044	0.9992
FOS for bearing capacity	0.236	0.9766

The prediction results in terms of the comparison of target with predicted output for the safety factor of sliding, overturning, and bearing capacity are shown in Fig. 1. It seemed that the target together with predicted output values were mostly overlapping for safety factors of sliding and overturning because of a very good prediction result of R^2 which is 0.99985 and 0.9992, respectively. The most significant difference of target and output values can be seen at the FOS for bearing capacity.

Fig. 2 presented the R^2 plot for a safety factor of sliding, overturning, and bearing capacity. The results proved that the SVM based on the RBF Kernel function can predict the cantilever retaining wall external stability with a good accuracy and reliability because the R^2 value greater than 0.97. The graph shows the dots were scattered close to the 45° line.

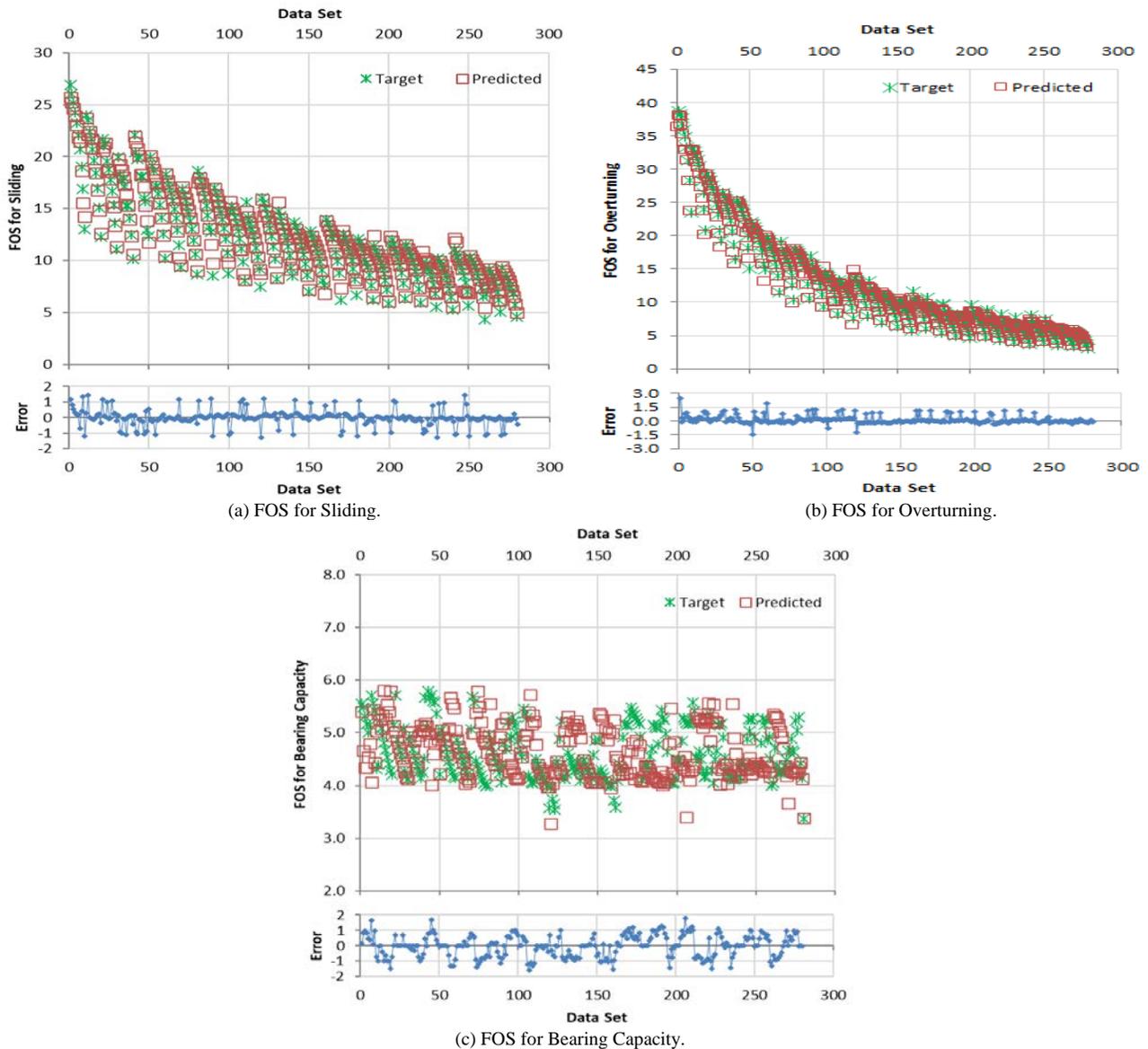


Fig. 1. Comparison of Target with Predicted Output for a Safety Factor of (a) Sliding, (b) Overturning, and (c) Bearing Capacity.

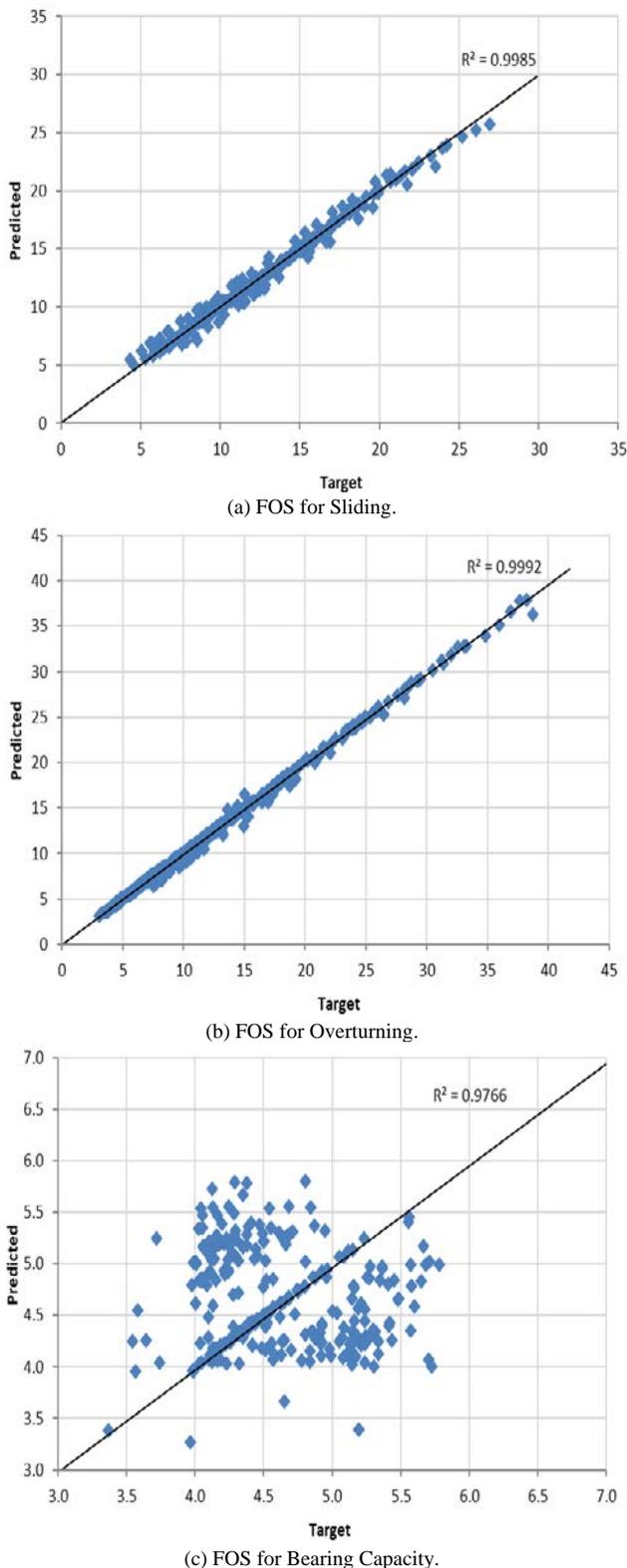


Fig. 2. Regression Square (R^2) Plot for a Safety Factor of (a) Sliding, (b) Overturning, and (c) Bearing Capacity.

V. CONCLUSION

SVM model prediction based on four Kernel functions, namely Linear, Polynomial, Radial Basis Function (RBF) and

Sigmoid were applied successfully to predict 280 data sets of external stability factors for cantilever retaining walls. The perfect prediction result was for RBF Kernel if compared with the other Kernel function because of the good R^2 for the output value. SVM model prediction based on RBF Kernel was able to predict the cantilever retaining wall stability with good accuracy and nearly to the target data. The prediction of the external stability of the cantilever retaining wall using the SVM model has successfully produced an accurate prediction by performing nonlinear regression for high-dimensional data sets.

As a conclusion, the results of the study can contribute to researchers for the current literature, especially in the field of retaining walls stability with the SVM approach. The optimal solution found in this study is that the right Kernel function has been obtained for the SVM model by producing the best prediction accuracy for the external stability of the cantilever retaining wall. The SVM is a technique that can solve complex problems with the application of appropriate Kernel functions. This study can help to forecast the stability of cantilever retaining walls used in civil engineering problems quickly and accurately. Prediction model developed provides advantages to geotechnical engineers in producing a more stable, and safe cantilever retaining wall.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the parties who indirectly involved in this study.

REFERENCES

- [1] F. Budiman, "SVM-RBF parameters testing optimization using cross validation and grid search to improve multiclass classification," *Scientific Visualization*, vol. 11, no. 1, pp. 80-90, 2019.
- [2] J.N. Cernica, *Geotechnical Engineering: Soil Mechanics*, John Wiley & Sons, Inc., 1995.
- [3] M.Y. Cheng and Y.W. Wu, "Prediction of Diaphragm Wall Deflection in Deep Excavations Using Evolutionary Support Vector Machine Inference Model (ESIM)," in *Proceedings of the 26th ISARC, Austin, USA*, pp. 176-182, 2009.
- [4] C. Cortes and V. Vapnik, "Support vector networks," *Machine Learning*, vol. 20, pp. 273 - 297, 1995.
- [5] H. Hong, B. Pradhan, D.T. Bui, C. Xu, A.M. Yousseff, and W. Chen, "Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Suichuan area (China)," *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 544-569, 2017.
- [6] L. Li, B. Wang, and S.O. Meroueh, Support Vector Regression Scoring of Receptor-ligand Complexes for Rank-ordering and Virtual Screening of Chemical Libraries, *J Chem Inf Model*, 51(9), pp. 2132-2138, 2011.
- [7] Y. Lu, N. Zeng, X. Liu, and S. Yi, "A new hybrid algorithm for bankruptcy prediction using switching particle swarm optimization and support vector machines," *Discrete Dynamics in Nature and Society*, pp. 1-7, 2015.
- [8] H.B. Ly and B.T. Pham, "Prediction of shear strength of soil using direct shear test and support vector machine model," *The Open Construction and Building Technology Journal*, vol. 14, pp. 41-50, 2020.
- [9] S.A. Mahmoud and S.O. Olatunji, "Automatic recognition of off-line handwritten Arabic (Indian) numerals using support vector and extreme learning machines," *International Journal of Imaging*, vol. 2, no. 9, pp. 34-53, 2009.
- [10] T. Mohamed, A. Kasa, N. Abdul Rahman, and M. Gaber, "Support vector machine prediction of external stability for segmental retaining wall," *International Journal of Civil Engineering and Technology*, vol. 10, no. 2, pp. 2483-2490, 2019.

- [11] J.Z. Mohammed and M.J. Wagner, *Data Mining and Analysis: Fundamental Concepts and Algorithms*, Cambridge University Press., 2014.
- [12] M.A. Nanda, K.B. Seminar, D. Nandika, and A. Maddu, "A comparison study of kernel functions in the support vector machine and its application for termite detection," *Information*, vol. 9, no. 5, pp. 1-14, 2018.
- [13] S.O. Olatunji, "Comparison of extreme learning machines and support vector machines on premium and regular gasoline classification for arson and oil spill investigation," *Asian Journal of Engineering, Sciences & Technology*, vol. 1, no. 1, pp. 1-7, 2011.
- [14] K.E. Pilario, M. Shafiee, Y. Cao, L. Lao, and S.H. Yang, "A review of kernel methods for feature extraction in nonlinear process monitoring," *Processes*, vol. 8, no. 24, pp. 1-47, 2020.
- [15] J.C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Advances in Large Margin Classifiers*, pp. 61-74, 1999.
- [16] N. Rachel and M. Lakshmi, "Landslide prediction with rainfall analysis using support vector machine," *Indian Journal of Science and Technology*, vol. 9, no. 21, pp. 1-6, 2016.
- [17] B. Radhakrishnan, G. Shineraj, and M.K.M. Anver, "Application of data mining in marketing," *International Journal of Computer Science and Network*, vol. 2, no. 5, pp. 41-46, 2013.
- [18] D. Ramya and S. Vinodhkumar, "Development of support vector machine model to predict stability of slopes based on bound theorems," *International Journal of Engineering and Technology*, vol. 9, no. 2, pp. 1231-1237, 2017.
- [19] P. Samui, "Prediction of friction capacity of driven piles in play using the support vector machine," *Canadian Geotechnical Journal*, vol. 45, no. 2, pp. 288-295, 2008.
- [20] P. Samui and T.G. Sitharam, "Machine learning modelling for predicting soil liquefaction susceptibility," *Nat. Hazards Earth Syst. Sci.*, vol. 11, pp. 1-9, 2011.
- [21] S. Shi, R. Zhao, S. Li, X. Xie, L. Li, Z. Zhou, and H. Liu, "Intelligent prediction of surrounding rock deformation of shallow buried highway tunnel and its engineering application," *Tunnelling and Underground Space Technology*, vol. 90, pp. 1-11, 2019.
- [22] A.J. Smola and B. Scholkopf, "On a kernel-based method for pattern recognition, regression, approximation and operator inversion," *Algorithmica*, vol. 22, pp. 211-231, 1998.
- [23] A.J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Statist. and Comput.*, vol. 14, pp. 199-222, 2004.
- [24] C.R. Souza, *Kernel Functions for Machine Learning Applications*, 2010. <http://crsouza.blogspot.com/2010/03/kernel-functions-for-machine-learning.html>
- [25] B. Thuraisingham, "Data mining, privacy, civil liberties and national Security," *SIGKDD Explorations*, vol. 4, pp. 28-34, 2002.
- [26] V. Vapnik and S. Mukherjee, *Support Vector Machine for Multivariate Density Estimation*, In: *Advances in Neural Information Processing System*, Leen, T., Solla, S. & Muller, K.R. (eds.), Cambridge, MA: MIT Press, pp. 659-665, 2000.
- [27] V. Vapnik, S. Golowich, and A. Smola, *Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing*, In: *Advances in Neural Information Processing System*, Mozer, M., Jordan, M. & Petsche, T. (eds.), Cambridge, MA: MIT Press., 1997.
- [28] I. Yimaz and A.G. Yuksek, "An example of artificial neural network (ANN) application for indirect estimation of rock parameter," *Rock Mechanics and Rock Engineering*, vol. 41, no. 5, pp. 781-795, 2008.

Analyzing the Performance of Anomaly Detection Algorithms

Chiranjit Das¹, Akhtar Rasool², Aditya Dubey³, Nilay Khare⁴

Department of Computer Science Engineering
Maulana Azad National Institute of Technology
Bhopal, India

Abstract—An outlier is a data observation that is considerably irregular from the rest of the dataset. The outlier present in the dataset may cause the integrity of the dataset. Implementing machine learning techniques in various real-world applications and applying those techniques to the healthcare-related dataset will completely change the particular field's present scenario. These applications can highlight the physiological data having anomalous behavior, which can ultimately lead to a fast and necessary response and help to gather more critical knowledge about the particular area. However, a broad amount of study is available about the performance of anomaly detection techniques applied to popular public datasets. But then again, have a minimal amount of analytical work on various supervised and unsupervised methods considering any physiological datasets. The breast cancer dataset is both a universal and numeric dataset. This paper utilized and analyzed four machine learning techniques and their capacity to distinguish anomalies in the breast cancer dataset.

Keywords—Anomaly; machine learning; outlier detection; minimum covariance determinant

I. INTRODUCTION

Anomalies can detect a fault in a system or a network, abnormalities in healthcare data, or fraud [1]. The fast growth in the amount, dimension, and complexity of data in the dataset has made it compulsory to automate the outlier detection in analytical processes. Those analytical results can then be used in decision-making processes by various algorithms to identify the health condition [2]. Description of these anomalous values can provide a better and new understanding of the dataset. Using different examples can visualize the whole phenomenon in a better way for the researchers. Particularly in healthcare, this could be used to develop good knowledge about the patient's health condition and complications, whereas in the case of predicted value, it does not focus only on the data description but also the future assumption about numerous states of a patient's health condition. That could help healthcare employees to treat the patient in an improved way by detecting the disease correctly. These outliers must not be considered an error or noise. Still, these are also important with respect to train the existing detection models and prepare to predict the new data added to the dataset.

Undoubtedly, the anomaly detection procedure depends on an effective definition of data point that must be considered an anomaly [3]. Anomaly detection approaches usually undertake two things, namely, (1) anomalies are rare and minor in

amount, and (2) anomalies are dissimilar in some sense or another from other normal data. The presence of multiple kinds of outliers or anomalies further results in complicating the operational definition. According to the occurrence, data points can be subdivided into anomalous concerning neighboring data points (local anomalies) or the entire dataset (global anomalies). It is challenging to differentiate between classes from which the anomaly belongs. The dissimilarity is still convenient for reminding the data scientists that anomalies can differ from both the data point and each other. A vast variety of technology is present in the healthcare system. However, the application of both machine learning and physiological data-set still needs some attention and in its early stages. Adding to the already problematic and challenging work of health care specialists is the scarcity of employees in the public healthcare field. To enhance the massive burden of healthcare specialists, it requires analyzing the minor units of physiological data collected, and it also helps improve patient care facility. Faulty events in the healthcare sector and minor medical errors may increase the chances of accidental deaths. Anomaly detection is a vital task, especially in the healthcare sector, where errors are rare. This paper has tried to better understand a patient's safety by predicting anomalous behaviors in breast cancer data through different anomaly detection techniques and errors in predicting the disease. Whether anomaly detection techniques are instigated to identify that patient is sick, look after the health conditions through sensors that can be wearable, or support another patient care-related work, these processes are vital components of the automated patient care system.

Nevertheless, what is considered by the system as an anomaly is often challenging to define? In most cases, choosing the proper machine learning techniques to implement depends on both investigation and the variability of the outliers present in the dataset. For example, there exist numerous works where the neural network has been used in healthcare-related problems. At the same time, some researchers have utilized clustering approaches and the multi-layer perceptron [4], [5]. Undoubtedly, choosing the algorithm depends partly on both the type of dataset and the dataset's size. Classifying an ideal machine learning model [6] for a particular problem relies on either or not the ground truth labels for data exist. In labeled data, supervised methods are suitable, while unsupervised methods allow analyzing the unlabelled data.

Nevertheless, it must be momentarily stated that semi-supervised anomaly detection approaches provide yet another vital option comprised of datasets labeled and comparably large amounts of unlabelled datasets. While countless models can be effectively adjusted to process the anomaly detection problems, the present study focuses on the subsequent four approaches: Local Outlier Factor (LOF), Isolation Forests (IF), Minimum Covariance Determinant (MCD), and One-Class SVM. The reason to choose these approaches are (1) their extensive occurrence in the area of machine learning and anomaly detection works fundamentally, and (2) because they comprise both supervised and unsupervised techniques. In the following subdivided sections, the techniques are briefly described.

II. RELATED WORKS

A. Local Outlier Factor (LOF)

The LOF technique was first presented by Breunig et al. as per the name recommends [18]. The technique primarily evaluates the separation between each of the data points concerning its neighborhood. This technique partially depends on the k-nearest neighborhood of the data point. Though other approaches classify outliers in a binary manner, LOF can also generate the degree of outliers of a data point. This local outlier density calculation is summarized with k-nearest neighbors. Data points in comparatively lower density zones are classified as outliers or assign a higher degree of outliers.

B. Isolation Forest (IF)

Isolation forest is an unsupervised type of machine learning technique for outlier detection implemented based on isolating anomalies instead of the most popular techniques of profiling normal points [19]. In this type of detection technique, the outlier is isolated by randomly partitioning them by selecting a random feature at a time. As a result, the outliers are more manageable to isolate than normal data. However, the node with the regular data needs time to isolate and requires many partitions.

C. Minimum Covariance Determinant

The input variables in the Gaussian distribution can deal with normal statistical methods to identify the outliers or anomalies. For example, suppose a dataset has two input variables. Minimum Covariance Determinant (MCD) is a detection technique that deals with those datasets to form the multi-dimensional Gaussian distribution [20]. Information of this distribution can be cast off to detect entries far from the distribution. Thus, the Minimum Covariance Determinant (MCD) technique is a vastly robust detector of multivariate location and scatter.

D. One-Class SVM (OC-SVM)

One-class SVM is a kind of unsupervised learning technique formally based on the classical Support Vector Machine (SVM) [21]. It is an algorithm where the training is only done with regular data or the dataset's negative data. Generally, this algorithm classifies the boundary with these training data subsets and can classify data point that does not belong to that boundary.

Fig. 1 and 2 shows the example features of various anomaly detection techniques on a 2D dataset. For every dataset, 15% of samples are produced as uniform random noise. This ratio is the value assigned to the nu parameter of One-Class SVM and the contamination boundary for other anomaly detection techniques. Considering a manually generated dataset, it tries to visualize the different outlier detection models for better understanding.

Outliers are generally the data points that do not comply with (other data points) general behaviors of the dataset. If the outliers are graphically represented, they fall typically somewhat apart from the regions consisting of normal data points. Outliers typically show variability in the dataset, errors present in measurement, or a novelty.

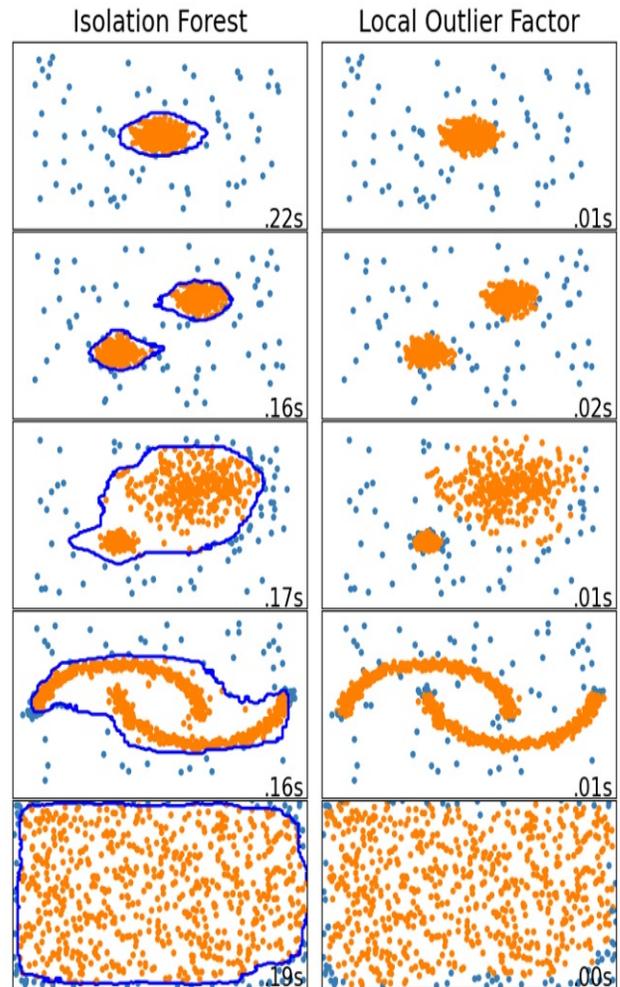


Fig. 1. Scatter Plot Showing the Graphical Representation of Outliers using Isolation Forest and LOF.

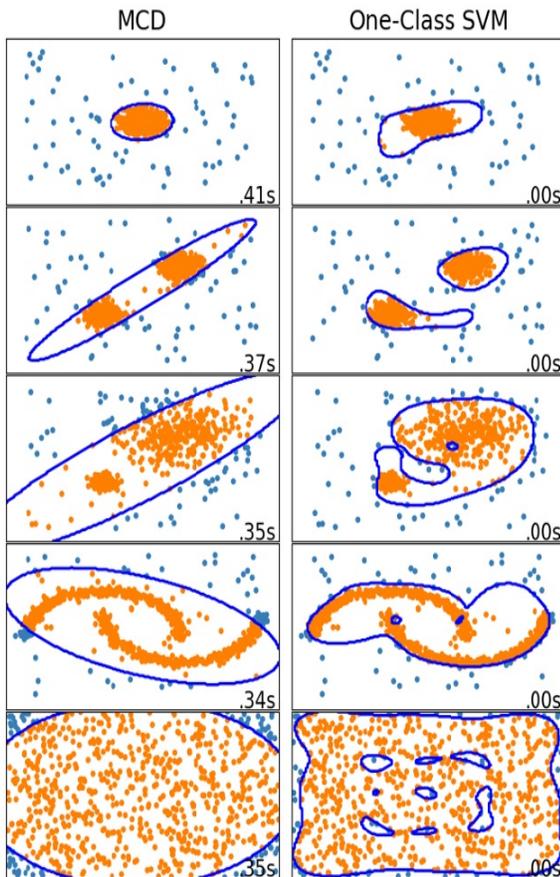


Fig. 2. Scatter Plot Showing the Graphical Representation of Outliers using MCD and One-Class SVM.

Random forest was a detection technique implemented by Leo Breiman [7]. It is an unsupervised learning technique and uses a Gini index classification of the anomalous data point. Many of the trees are generated randomly in this approach, selecting features, and then the most popular one is voted out. Outlier detection was implemented with machine learning techniques by Hadi et al. [8]. It is a supervised technique based on the regression analysis model trained by the training subset and predicts the outliers' test subset. Accordingly, Ramaswamy et al. come with a partition-based supervised technique capable of handling the numeric dataset [9]. It is a proximity-based algorithm. Pawlak et al. use the roughest methodology to derive the outlier from the numeric dataset [10]. It is a supervised detection algorithm based on the model of soft computing. Peterovskiy et al. proposed a fuzzy approach based on kernel function, a supervised learning technique that requires no separate training data [11]. It is capable of handling the numeric type of dataset. The class represents the degree of membership is used to classify the outliers from the normal data point. A trendy outlier detection technique known as Local Outlier Factor (LOF) was introduced by Kriegel et al. [12]. It is a density-based detection technique. It is a supervised learning technique that can process datasets with both numeric and categorical data.

Benjamin Nachman et al. introduced an approach to find outliers using density estimation [13]. A density-based outlier

detection technique uses conditional probability density to classify outliers in the dataset containing categorical and numerical data.

Bo Tang et al. introduced an outlier detection algorithm based on the Local Density of data [14]. It is a density-based detection technique that calculates outlier in a dataset containing both numerical and categorical data. Duan et al. proposed (LDBSCAN) a clustering-based model and used Euclidian distance as a proximity measure [15]. It is capable of the handle only numeric datasets. Kriegel has discussed many subspace-based detection algorithms that consider the subspace of the feature set [16]. All these works are generally based on a subset of the different existed detection techniques. Rank Based Detection Algorithm (RBDA) (2011) was introduced by H. Huang et al. [17]. It is a rank-based model that uses the density of the data points to identify the outliers. Our paper will provide an analysis of four popular detection techniques with different analytical measures.

III. MATERIALS AND MODEL IMPLEMENTED

A. Approaches Involved

Outlier detection is significant and essential in data mining, which can be used to preprocess the dataset, increasing the final result's accuracy. If the database is vast, then it always consists of some abnormal dataset. These data points are known as outliers which are irrelevant from the regular data points. Therefore, these data points must be removed from standard data in data mining. Hence, outlier detection, also known as anomaly detection, is required to classify the outliers to improve the data quality.

1) *Local Outlier Factor (LOF)*: LOF is a type of score that tells either a data point is an outlier or not [18]. Initially, let us start with the basic introduction of a parameter k , determined by calculating the distance between a point, say, p , and its k th nearest neighbors. This detection algorithm works by considering the neighborhood density of a specific data point to its k nearest neighbor. This density of the data points is then considered to determine the relative densities. Now choosing the correct value of k is a bit tricky, for selecting a too-small value of k it will focus more on local data, i.e., concentrated more on neighborhood points giving erroneous result in terms of noisy data, and for a large value of k it will completely miss out the local outliers. Breunig et al. introduced a LOF detection algorithm to understand the algorithm more precisely, which consists of three important terms discussed below:

a) *k-Distance*: The term k -distance can be defined as a distance is between the data point from its k th neighborhood data point. Let us assume k has a value of 4; k -distance will be the distance from a point to its 4th nearest neighbor data point.

Reachability Distance: To determine the reachability distance, the k -distance has been used. The reachability distance is a measure to calculate the maximum distance between two data points and the k -distance value of the second point. Mathematically,

$$reach_dist = \max\{-distance(p, q), distance(p, q)\} \quad (1)$$

Suppose p is a k -nearest neighbor of point q , then $reach_distance(p, q)$ can be defined as k -distance of q or the distance between these two points.

b) Local Reachability Distance (lrd): This value of reachability distance calculates another key concept called local reachability distance (lrd). To determine the value of lrd first step, calculate all the $reach_distance$ of k -nearest neighborhood points and find the mean of these values; use the inverse of this mean will be the lrd. Here, for densities, the higher the distance between the points lesser will be the density. Local reachability distance can be represented as:

$$lrd(p) = 1 / \left(\frac{mean(reach_distance(p,n))}{k} \right) \quad (2)$$

Parameter $lrd(p)$ represents the distance from a point to its nearest neighbors. Each point's lrd and the neighbor's lrd are then related. Through this, the ratio for each point is determined, and calculates the average. The LOF is determined by calculating the average ratio of lrd of a particular point by its neighbor's lrd. Lastly, the density of that data point is correlated with the density of its nearest neighbors. Less density of a data point than its neighbors indicates that point as an outlier.

2) Isolation forest: A feature is chosen randomly at once in an isolation forest from a subset of the dataset. It then isolates each data point until they are entirely isolated from one another [19]. This algorithm is based on the idea that anomalous data points are easier to isolate than normal data points. The detection algorithm enables building all the possible collections of isolation trees (Itree). Those Itrees are made considering all random subsets of data points and assigning each Itree with an outlier score that sums up an outlier score for each dataset's data points. The most important part of this algorithm is building the Itree. The process that comprises choosing a random subset of the whole dataset and then selecting one random feature at a time separates each point until they are entirely isolated. It is required to conduct a binary search operation to predict a new data point's anomalous behavior for that point. That operation must follow top to bottom order. Then assign the anomaly score to that data by calculating the total path length to reach the data point. Then determine the collective outlier score of the data point from the outlier score of individual Itree. The required mathematical equation to compute the outlier score is given by:

$$S(p, m) = 2^{-\frac{E(h(p))}{c(m)}} \quad (3)$$

Where $S(p, m)$ is the outlier score of the data point p and m is the sample size, $h(p)$ gives the average depth to reach the point p from the Itree. $c(m)$ is the average value of $h(p)$. Now, if outlier score, i.e., when $E(h(p)) \ll c(m)$ indicates the point to be an anomalous point and if outlier score, when indicates a regular point. The points with a value close to 1 are outliers, and values close to 0.5 are normal.

The whole detection process can be sub-divided into two stages- the training stage and the evaluation stage. In the training

stage, the building of Itree is done by recursively separating every data point up to a precise height is reached, which is approximately the average depth of Itree. Then, each test data point's outlier score is set from the expected path distance $E(h(p))$ in the evaluating stage. Finally, the value of $E(h(p))$ is determined by passing the test data point through Itree, a part of an isolation forest.

3) Minimum Covariance Determinant (MCD): Robust estimation of the multivariate mean ($\hat{\mu}$) and covariance ($\hat{\Sigma}$) in MCD, searching data subset from the h data points is concluded in MCD [20]. The robust estimate must comply with a minimum determinate value the covariance matrix. where h data point must lie between $\frac{k+p+1}{2}$ to k . The MCD is based on Mahalanobis Distance (MD) [21] and given by:

$$m^2 = (x - \hat{\mu})^T \hat{\Sigma}^{-1} (x - \hat{\mu}) \quad (4)$$

When the determinate covariance is equal to zero, then the inverse of the covariance ($\hat{\Sigma}$) will not exist, and also, the value of m will be undefined in the case of ($h < p$). So, it must be necessary to have more observation than the number of variables to determine the value of m for the dataset. It is complex and costly to implement exact MCD instead; fast-MCD is very popular in practice. Fast-MCD starts with randomly choosing a subset from observation equal to $p+1$ from dataset x [22]. The subset x_n of real dataset x , simultaneously calculating their Mahalanobis Distance m_n for k number of observations with subsequent mean and variance $\hat{\mu}$ and $\hat{\Sigma}$, respectively. Then the number of h observations are separated from a whole new subset of the dataset x , having the lowest value of MD m_0 . The value of h is determined by-

$$h = \left\lfloor 2 \left\lfloor \frac{k+p+1}{2} \right\rfloor - k + 2\alpha \left(n - \left\lfloor \frac{k+p+1}{2} \right\rfloor \right) \right\rfloor \quad (5)$$

The value of the parameter α always lies between 0.5 and 1, representing the desirable robustness. Equating a lower value of α may cause in increasing the robustness but also cost inefficiency and resulting in a potentially large outlier set. Then all x_i and m_i is computed for all k number of observations each time choosing a new subset of h , unless the subset at the present iteration is the same as the previous iteration, the process will repeat. At that point, the local minimum value of determinate of covariance is attained. This algorithm repeats itself up to a maximum number of the subset of x . x_{Mcd} is then defined by a subset of h observation with the lowest value of covariance determinate. The robust estimation is calculated by:

$$\hat{\mu}_{Mcd} = \frac{1}{h} \sum_{i=1}^h x_{Mcd_i} \quad (6)$$

$$\hat{\Sigma}_{Mcd} = c_0 \frac{1}{h} \sum_{i=1}^h (x_{Mcd_i} - \hat{\mu}_{Mcd})(x_{Mcd_i} - \hat{\mu}_{Mcd})^T \quad (7)$$

To handle correctly comparatively small sample of a scalar consistency factor (c_0) is introduced in the equation of estimation, following the estimate of m_{Mcd} . Then the threshold value of m_{Mcd} is calculated, which indicates the data point as an outlier when it falls beyond that threshold. A popular approach is followed to do so introduced by Dovoedo and Chakraborty [23]. In this approach, the first step is to transform the value of each x_{Mcd} to Robust Mahalanobis

Distance (r_{Mcd}) [20] outliers, to limit the distance distribution in the range of zero to one.

$$r_{Mcd} = 1 - \frac{1}{1+m_{Mcd}} \quad (8)$$

Then on the second step, some regular multivariate sample of k observation is simulated with $\hat{\mu}_{Mcd}$ and $\hat{\Sigma}_{Mcd}$. Finally, the outliers of those simulated k observations are determined and $\hat{\epsilon}_{Mcd}$ percentile is being used from these simulations to limit the outliers. The user-defined value $\hat{\epsilon}_{Mcd}$ lies between zero to one. More reliably, the detect outlier is needed to implement a robust estimation to resist a possible outlier set. The robust MCD estimator is more helpful to deal with the multivariate dataset. The distance called robust distance [20] is used to flag anomalies. The robust distance is given by-

$$RD^2 = (x - \hat{\mu}_{Mcd})^T \hat{\Sigma}_{Mcd}^{-1} (x - \hat{\mu}_{Mcd}) \quad (9)$$

It is more reliable to define a precise hypersphere capable of covering all regular data points, and the points out of this distance from the origin are termed outliers.

4) *One-Class Support Vector Machine (OCS)*: A One-Class SVM is a type of unsupervised machine learning technique [24]. The One-Class SVM is only trained with regular points or negative examples. It automatically develops learning boundaries of those negative points and can successfully estimate the classification of any data that belongs outside that defined boundaries and identify them as outliers. The training stage of any unsupervised machine learning technique is challenging, and so with the One-Class SVM. Crucial parameter ν in this detection technique controls the amount of outlier or contamination one user expects to identifies as outliers. The gamma is another parameter used to calculate the smoothing of boundary lines.

In One-Class, SVM defines an optimized boundary used to differentiate regular data from the anomalous data points in higher dimensionality by maximizing the difference between regular and anomalous points. For example, in One-Class SVM, a hyperplane [25] is defined as a separate anomalous and regular data point from the origin.

The objective function to generate a hyperplane for One-Class SVM is given by-

$$\min_{w, \xi, \rho} \frac{1}{2} \|w\|^2 + \frac{1}{\nu n} \sum_{i=1}^n (\xi_i - \rho) \quad (10)$$

$$\text{Subject to: } (w \cdot \phi(x_i)) \geq \rho - \xi_i, \xi \geq 0 \quad (11)$$

Where w is weight vector, ξ is represented slack variables, ρ is the distance between origin and hyperplane, $\phi(x_i) \rightarrow F$ is feature space mapping for input dataset x . The regularization parameter is represented by ρ , which controls the boundary around regular data and leaves a fraction of data classified as anomalous. The main objective is to attain a hyperplane with less distance from the origin, which calculates the optimized value of w & r so that some miss-classification allows for few data points. Moreover, the hyperplane-based One-Class SVM has a low-performance ability. So, to overcome this, another approach is developed where a hypersphere is constructed aiming to separate regular and anomalous point by a sphere of

radius R , center C , and feature space F [26]. the objective function is given by-

$$\min_{R, \xi, C} R^2 + \frac{1}{\nu n} \sum_{i=1}^n \xi_i \quad (12)$$

$$\text{Subject to: } \|\phi(x_i) - C\|^2 \leq R^2 + \xi_i \quad (13)$$

Here the main issue is to reduce the value of R of hypersphere and to compute this. It has to determine the value R & C so that most data points have a comparative Euclidean distance to radius R so that some miss-classifications are allowed.

B. Model Implemented

To analyze the different outlier detection techniques, make a baseline algorithm performance and compare it with other algorithm's performance. We developed the baseline performance by fitting a linear regression for the given dataset. The evaluation of method performance is then carried out by training a particular algorithm on the training data subset. The Mean Absolute Error, Root Mean Square Error, and model score [26] are determined based on the algorithm's prediction on the test dataset. Then we fit the dataset with each of the detection algorithms after defining them. The previously fitted algorithm predicts the anomalous and non-anomalous data points. The anomalous data points are dropped from the training part of the dataset. Then the remaining data points are fitted on the detection algorithm, and finally, the prediction on the test part of the dataset is analyzed.

This is the analytical model use to analyze breast cancer data. Four different outlier detection techniques are used to detect outliers. The various analytical measures are used to analyze the predicted outcomes. Then a classifier is called a decision tree classifier. The proposed algorithm is given below:

Input: Breast cancer dataset.

Output: Analysis and classification of data.

Step 1: Prepare the breast cancer dataset.

Step 2: Prepare a Base performance using the linear regression technique.

Step 3: Apply different detection methods to identify outliers.

Step 4: Analyze the results with different analytical measures.

Step 5: ROC value and the precision rank of different detection are measured.

Step 6: Out of the above outlier detection, the best one is used to detect the outlier from the breast cancer dataset, and using the confusion matrix as a classifier, the prediction of breast cancer has been made.

IV. RESULT AND DISCUSSION

A. Data Description

The breast cancer dataset is considered is downloaded from the data world repositories. The breast cancer dataset contains information about the women who are likely to have cancer. The dataset consists of a total of 11 attributes, namely sample id, clump thickness (1-10), uniformity of cells (1-10),

marginal adhesion (1-10), single epithelial cell size (1-10), bare nuclei (1-10), bland chromatin (1-10), normal nucleoli (1-10), mitoses (1-10) and finally the class attribute. The dataset consists of 698 instances. The dataset also comprises 10 feature attributes and one class attribute. The original dataset is split into training and test dataset. The training dataset is 70% of the total dataset, i.e., 488 instances, and the test data is 30% of the entire dataset, i.e., 210 cases.

B. Analytical Measures

The different detection algorithms' performance is evaluated by some analytical measures like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Model Score of each detection algorithm [27],[28],[29]. The measurements are described in the following Table I.

TABLE I. ANALYTICAL MEASURES

Evaluating Measures	Description	Formula
Mean Absolute Error (MAE)	It helps to evaluate the mean error in the prediction set, not considering its direction. It indicates the error between prediction and actual observations.	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $
Root Mean Square Error (RMSE)	It is a type of quadratic grading method that uses to determine the average magnitude of the error.	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$
Model score	It takes input as a feature matrix and probable target values. Then both predictions made on the feature matrix are compared with the target to attain a score.	mdl.fit (X_train, y_train) mdl.score(X_test, y_test)
Precision	It is the ratio of the true anomalies present with the total number of anomaly candidates.	$Tp / (Tp + Fp)$ where Tp is the number of true positives and Fp is the number of false-positive cases.

Both MAE and RMSE are the average error of the prediction made by the detection algorithms. This error measure works on negatively oriented scoring, which implies the lowest values are better values. For considering larger errors, RMSE is much useful. The more excellent value of RMSE does not mean the larger value of variance in the errors. Using the MAE value, we can always bound the values of RMSE.

- The RMSE error values are either greater than or equal to MAE, i.e., $RMSE \geq MAE$.
- When the prediction errors are calculated from the same test dataset, then the difference in MAE and RMSE will be the highest. In that respect, the total squared error will always equal $(MAE^2 \times n)$, where n is the data in the test dataset. Thus, RMSE is less than or equal to the absolute squared error's square root, i.e., $RMSE \leq [MAE^2 \times n]$.

Finally, RMSE is more beneficial for penalizing a large number of errors, but both are equally important.

TABLE II. EVALUATING RESULTS OF OUTLIER DETECTION TECHNIQUES

Analytical Measures	Anomaly Detection Approaches			
	Isolation Forest	MCD	OC-SVM	LOF
ROC (Training dataset)	0.9833	0.9864	0.9838	0.8863
ROC (Test dataset)	0.9804	0.9832	0.88	0.8123
Precision	0.8888	0.9028	0.8880	0.9164
Model Score	0.802	0.835	0.814	
Computation Time	0.3	0.3	0.3	<0.1

C. Result Analysis

The evaluating results are tabulated in Table II, representing the ROC and the precision of the different algorithms and the algorithm's model score.

Fig. 3 and 4 represent the MAE and RMSE of the different outlier detection techniques. The figure shows the performance analysis of different detection algorithms; we can derive that the best performing algorithm is Minimum Covariance Determinant (MCD). It has lower MAE and RMSE values of 0.241 and 0.371 than other detection algorithms. It also has a higher prediction accuracy rate of 84%.

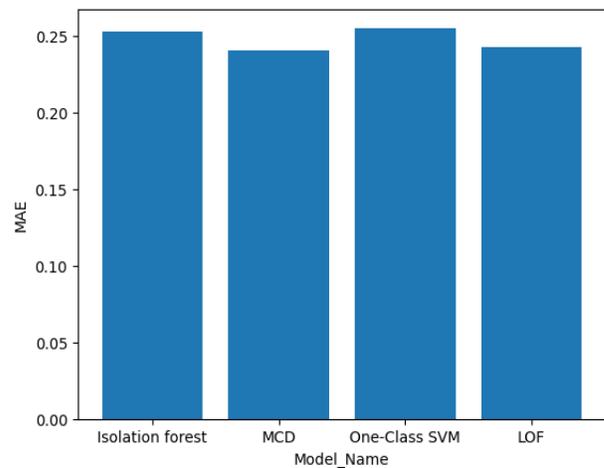


Fig. 3. Mean Absolute Error of Different Detection Algorithm.

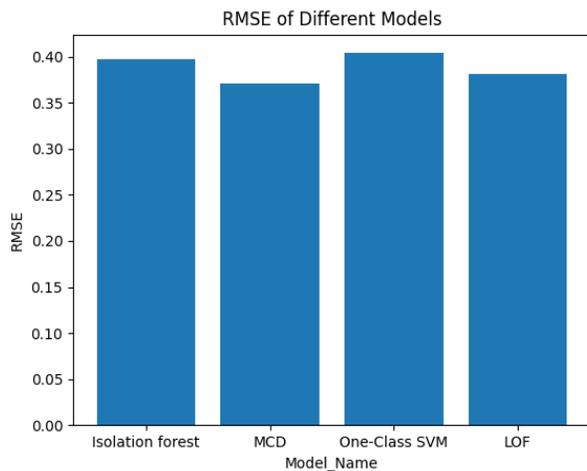


Fig. 4. Root Mean Square Error of the Different Detection Algorithm.

V. CONCLUSION

This paper discussed the details of the four most popular and versatile outlier detection algorithms. It is a performance analysis-based study where various analytical measures are considered to derive a final result. The breast cancer dataset was considered in our research from multiple datasets available in the real world. The detection algorithms considered are also very diverse. The MAE and RMSE are used for analyzing the performance of different algorithms. Also, the model score of all the algorithms is calculated. The returned value of the model score is measured as these algorithms' accuracy to predict the test dataset after training the model with the training dataset. The slight variance in both the error values can make a lot of difference in analyzing algorithms' performance with the particular dataset. The inclusive conclusion observed that the algorithm that produced a minor error in predicting the test dataset is Minimum Covariance Determinant (MCD) with MAE and RMSE of 0.241 and 0.371. MCD also has the highest accuracy among the other algorithm.

REFERENCES

- [1] C. Romero, and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 3, e. 1355, Jan. 2020.
- [2] A. Dubey, and A. Rasool, "Clustering-Based Hybrid Approach for Multivariate Missing Data Imputation," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 11, no. 11, pp. 710-714, 2020.
- [3] A. Boukerche, L. Zheng, and O. Alfandi, "Outlier Detection: Methods, Models, and Classification," *ACM Computing Surveys (CSUR)*, vol. 53, no. 3, pp. 1-37, Jun. 2020.
- [4] D. Xu, and Y. Tian, "A comprehensive survey of clustering algorithms," *Annals of Data Science*, vol. 2, no 2, pp. 165-193, Aug. 2015.
- [5] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S.Y. Philip, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4-24, Jan. 2021.
- [6] M. Fatima, and M. Pasha, "Survey of machine learning algorithms for disease diagnostic," *Journal of Intelligent Learning Systems and Applications*, vol. 9 no. 1, p.1-16, 2017.
- [7] L. Breiman, Random forests. *Machine learning*, vol. 45, no. 1, pp. 5-32, Oct. 2001.
- [8] A. S. Hadi, "A new measure of overall potential influence in linear regression," *Computational Statistics & Data Analysis*, vol. 14, pp. 1-27, 1992.
- [9] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," *ACM SIGMOD Record*, vol. 29, pp. 427-438, 2000.
- [10] Z. Pawlak, J. Grzymala-Busse, and W. Ziarko, "Rough sets", *Communications of the ACM*, Vol. 38, pp. 88-95, 1995.
- [11] M. I. Petrovskiy, "Outlier detection algorithms in data mining systems," *Programming and Computer Software*, Vol. 29, pp. 228-237, 2003.
- [12] M. M. Breunig, H. P. Kriegel, R. T. Ng, and J. Sander, "Lof: Identifying density-based local outliers," *ACM SIGMOD*, Vol. 29, pp. 93-104, 2000.
- [13] B. Nachman, and D. Shih, "Anomaly detection with density estimation," *Physical Review D*, vol. 101, no. 7, pp.075042-1-16, Apr. 2020.
- [14] B. Tang, and H. He, 2017, "A local density-based approach for outlier detection," *Neurocomputing*, vol. 241, pp.171-180, 2017.
- [15] L. Duan, L. Xu, Y. Liu, and J. Lee, "Cluster-based outlier detection," *Annals of Operations Research*, Vol. 168, pp. 151-168, 2009.
- [16] H. P. Kriegel, P. Kröger, E. Schubert, A. Zimek, (2, "Outlier detection in axis-parallel subspaces of high dimensional data," In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, pp. 831-838, 2009.
- [17] H. Huang, K. Mehrotra, C. K. Mohan, "Rank-based outlier detection," *Journal of Statistical Computation and Simulation*, Vol. 83, No. 3, pp. 518-531, 2011.
- [18] M. M. Breunig, H. P. Kriegel, R. T. Ng, and J. Sander, "Lof: Identifying density-based local outliers," *ACM SIGMOD*, Vol. 29, pp. 93-104, 2000.
- [19] A. Mensi, and M. Bicego, "A novel anomaly score for isolation forests," in *proc. International Conference on Image Analysis and Processing*, Springer, University of Verona, Verona, Italy, Sep. 2019, pp. 152-163.
- [20] M. Hubert, and M. Debruyne, "Minimum covariance determinant," *Wiley interdisciplinary reviews: Computational Statistics*, Vol. 2, No. 1, pp. 36-43, 2010.
- [21] C. Leys, O. Klein, Y. Dominicy, and C. Ley, "Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance," *Journal of Experimental Social Psychology*, Vol. 74, pp.150-156, Jan. 2018.
- [22] M. Hubert, M. Debruyne, and P. J. Rousseeuw, "Minimum covariance determinant and extensions," *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 10, No. 3, e. 1421, 2017.
- [23] Y. H. Dovoedo, and S. Chakraborti, "Outlier detection for multivariate skew-normal data: a comparative study," *J Stat Comput Simul*, Vol. 83, pp. 773-83, 2011.
- [24] S. Amraee, A. Vafaei, K. Jamshidi, and P. Adibi, "Abnormal event detection in crowded scenes using one-class SVM," *Signal, Image, and Video Processing*, vol. 12, no. 6, pp. 1115-1123, Mar. 2018.
- [25] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognition*, Vol. 58, pp. 121-134, 2016.
- [26] Q. Li, "Covariance modelling with hypersphere decomposition method and modified hypersphere decomposition method," *The University of Manchester, Manchester, United Kingdom*, 2018.
- [27] A. Dubey, and A. Rasool, "Time-Series Missing Value Prediction: Algorithms and Applications," *International Conference on Information, Communication and Computing Technology ICICCT*, pp. 21-36, 2020.
- [28] A. Dubey, and A. Rasool, "Local Similarity-Based Approach for Multivariate Missing Data Imputation." *International Journal of Advanced Science and Technology*, Vol. 29, No. 06, pp. 9208 - 9215, 2020.
- [29] A. Dubey and A. Rasool, "Data Mining based Handling Missing Data," *International Conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud)*, pp. 483-489, 2019.

Current Perspective of Symbiotic Organisms Search Technique in Cloud Computing Environment: A Review

Ajoze Abdulraheem Zubair¹
Shukor Bin Abd Razak², Md. Asri Bin Ngadi³
School of Computing, Faculty of Engineering
Universiti Teknologi Malaysia, 81310 Skudai
Johor Bahru, Johor Malaysia

Aliyu Ahmed⁴
Department of Mathematics
Bauchi State University Gadau
751105 Itas-Gadau Bauchi, Nigeria

Abstract—Nature-inspired algorithms in computer science and engineering are algorithms that take their inspiration from living things and imitate their actions in order to construct functional models. The SOS algorithm (symbiotic organisms search) is a new promising metaheuristic algorithm. It is based on the symbiotic relationship that exists between different species in an ecosystem. Organisms develop symbiotic bonds like mutualism, commensalism, and parasitism to survive in their environment. Standard SOS has since been modified several times, either by hybridization or as better versions of the original algorithm. Most of these modifications came from engineering construction works and other discipline like medicine and finance. However, little improvement on the standard SOS has been noticed on its application in cloud computing environment, especially cloud task scheduling. As a result, this paper provides an overview of SOS applications in task scheduling problem and suggest a new enhanced method for better performance of the technique in terms of fast convergence speed.

Keywords—Cloud computing; cloud resource management; cloud task scheduling; symbiotic organisms search; entrapment; convergence speed

I. INTRODUCTION

The term "cloud computing" refers to a fully centralized, scalable, on-demand computing network with distributed virtual infrastructure, storage, and pay-per-use services[1][2][3]. Resource management in infrastructure as a service (IaaS) is one of the most important challenges of cloud computing especially cloud task scheduling. Task Scheduling is an optimization problem that falls into the NP-hard class[4][5][6]. Therefore, using conventional search algorithms to find an optimal mapping between tasks and resources in a short period of time becomes almost impossible.

Many metaheuristic algorithms have been proposed to solve optimization problems in recent decades. Metaheuristic research around the world has resulted in optimization techniques that have proved to be superior to conventional gradient-based approaches[7][8][9][10]. Almost all metaheuristic algorithms have the following features in common: nature-inspired, probability distributions, do not need significant gradient information, and many parameters that must be tailored to the problem at hand. Several metaheuristic algorithms that mimic the actions of insect or animal classes in

nature have been developed using a combination of deterministic rules and randomness[9][11]. Many of the optimization problems that can be formulated are extremely nonlinear, with multimodal objective landscapes and a series of dynamic, nonlinear constraints[12][13][14].

Optimization algorithms are intelligent self-learning algorithms derived from the observation and mimicking of intelligent processes and actions observed in nature, usually based on swarm intelligence[9][15]. The optimization problems that have piqued interest in metaheuristic approaches range widely in complexity, from single to multi-objective, continuous to discrete, constrained to unconstrained and large-scale global optimization problems. Therefore, solving these problems is not always easy due to their complicated behaviour[1][9][16][17][18]. Studies of swarm behavior can perhaps be regarded as a field of artificial intelligence (AI), is saddled with the responsibilities of collecting information about the behaviour of organisms in self organized environment. Symbiotic Organisms Search algorithm (SOS) is an SI(Swarm Intelligence)-based recently developed metaheuristic algorithm that was inspired by nature[19]. The SOS algorithm mimics the collaborative behavior observed in nature among individuals. Symbiotic relationships between paired organisms in an ecosystem define cohabitation behavior of different species which include mutualism, commensalism, and parasitism. Symbiosis is a close relationship between organisms that last over a period[20][21][22][23]. Overview of the relationship among distinct species is depicted in Fig. 1.

Mutualism is the interaction between two species with mutual benefit, which means that both benefit from the relationship. Commensalism happens when one species forms a bond with a different species, with one species gaining while the other is unaffected. When two species form a relationship in which one benefits and the other hurts, it is referred to as parasitism. Both mutualism and commensalism operations focus on creating new species for the next generation by allowing the search procedure to find solutions to the problems within the solution search space, thereby improving the algorithm's exploratory capability. The parasitism phase, on the other hand, focuses on increasing exploitative capability by allowing the search process to escape entrapment at local optimality.

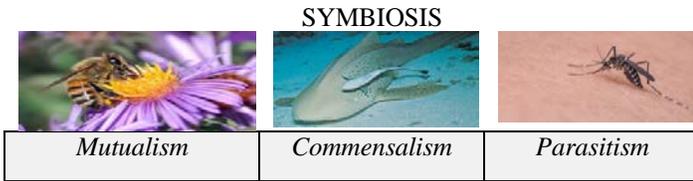


Fig. 1. Vector Illustration of Symbiosis.

The optimization analysis group and other similar domains have been paying considerable interest to the algorithm because of its deployment simplicity and consistency.

Furthermore, in comparison to other alternative algorithms or competitors such as GA, ACO, PSO, and DE, the SOS algorithm has drastically risen in its area of application to several optimization realms but little in cloud computing environments. Therefore, this paper provides a comprehensive analysis of the standard SOS algorithm, its basic concepts and structures, variants, and hybrid implementations for addressing constrained, unconstrained, single objective, multiple objectives, large scale global optimization problems, and practical oriented real world optimization problems in a cloud computing setting. This paper's major contributions are as follows:

- A comprehensive review of the traditional SOS algorithm.
- A look at SOS varieties and hybridization strategies.
- Identification and presentation of the different applications of the SOS algorithm in task scheduling problems within cloud computing context.

The rest of the paper is structured as follows: Section 2 introduces the fundamental concept and summarized work on standard SOS algorithm; Section 3 presents a review of research works on SOS algorithms, including variants and hybrid versions of the regular algorithm, while Section 4 concludes the study and provides viable future directions.

II. SYMBIOTIC ORGANISMS SEARCH ALGORITHM (SOS)

Cheng and Prayogo [19] developed a framework for simulating the symbiotic relationship between species in an environment, using it to solve mathematical and engineering design problems. Three distinct search processes govern the next possible solution, which are modeled after three basic symbiotic interactions: mutualism, commensalism, and parasitism.

A. Mutualism Phase

This is the first stage which entails selecting organisms based on their fitted values. In this phase, two organisms Z^i and Z^j (where $i \neq j$) are selected to interact with each other. The two distinct organisms benefit from one another. The mathematical formulations of these procedures are presented as in (1) and (2).

$$Z^{i_new} = Z^i + \mathfrak{r}(0,1) \times [Z^{best} - M_v \times B^{f-1}] \quad (1)$$

$$Z^{j_new} = Z^j + \mathfrak{r}(0,1) \times [Z^{best} - M_v \times B^{f-2}] \quad (2)$$

The joint relation vector between Z^i and Z^j denoted by M_v and benefit factor (B^f) are evaluated using the mathematical formulars in equations (3), (4), and (5) respectively.

$$M_v = \frac{z^i + z^j}{2} \quad (3)$$

$$B^{f-1} = 1 + \text{round}(\mathfrak{r}(0,1)) \quad (4)$$

$$B^{f-2} = 1 + \text{round}(\mathfrak{r}(0,1)) \quad (5)$$

The new species Z^{i_new} and Z^{j_new} are created by sculpting their structure from M_v and B^f matching to the existing population's fittest species (Z^{best}). However, while M_v specifies the mutual bond between different species, B^f denotes the degree of value gained by each species as a result of their interaction. $\mathfrak{r}(0,1)$ denotes a vector of uniformly distributed random numbers ranging from 0 to 1. The fitness values of these new species Z^{i_new} and Z^{j_new} are each evaluated and compared with their respective predecessor to select the next generation of the ecosystem. If the fitness values of the new species are better than their predecessor, then they automatically move to the next generation otherwise they are rejected, and the old species remain. Note that worst fitness values are rejected and replaced.

B. Commensalism Phase

In this phase, only one organism Z^i can increase his or her fitness value. This is achieved by randomly select organism Z^j which is neutrally affected in their relationship. The commensalism relationship between these two organisms is established using the mathematical equation as in (6).

$$Z^{i_new} = Z^i + \mathfrak{r}(-1,1) \times (Z^{best} - Z^j) \quad (6)$$

So, the beneficial contribution of organism Z^j to organism Z^i is computed using $(Z^{best} - Z^j)$ as in (6). Similar to the mutualism process, the newly created organism (Z^{i_new}) can only substitute its predecessor (Z^i) if and only if its fitness value is higher; otherwise, (Z^i) is retained.

Algorithm 1 below outlines the processes of the standard symbiotic organism search (SOS) algorithm.

Algorithm 1: Pseudo-code of Symbiotic Organism Search algorithm

```

Define input variables, objective function, and problem dimension.
//Population size (ecosize), maximum iterations (Maxitern)
Initialize population with respect to population size(ecosize)
Identify the best organism in the ecosize ( $Z^{best}$ )

Start SOS Looping:
  While itern<maxitern
    For i = 1: ecosize
      Mutualism Phase
      Select species  $Z^i$  and  $Z^j$  ( $i \neq j$ )
      compute the common vector ( $M_v$ ) as in (3) and the Benefit
      factor ( $B^{f-1}$  and  $B^{f-2}$ ) as in (4) and (5)

      Using (1) and (2) to create the new species
       $Z^{i\_new}$  &  $Z^{j\_new}$  and assess their strengths.
      If the strengths of the new species are better, then
        replace the predecessors

      Commensalism phase
      Select species  $Z^j$  randomly ( $i \neq j$ )
      Using (6) to create new species  $Z^{i\_new}$  and assess
      its strength
      If the strength of the new species is better, then
        replace the predecessor.

      Parasitism Phase
      Select species  $Z^j$  randomly ( $i \neq j$ )
      Produce parasite vector  $Z^p$  by amending  $Z^i$ , assess the strength
      If the strength of the parasite vector ( $Z^p$ ) is better,
      then replace  $Z^j$  with  $Z^p$ 

    End for
    Update best species  $Z^{best}$  of the existing population
  End while

```

C. Parasitism phase

The phenomenon of parasitism can be best described by the interaction between a mosquito and human. Once the human is

bitten by mosquito, the process create parasite in the human body. The parasite present in the human body reproduces itself and overwhelms the human host through disease infection which may lead to death. On the contrary, if the human body has better immunity, then it protects itself and the parasite gets eliminated from the body.

The human host is harmed in this type of interaction, while the anopheles mosquito, which is the parasite carrier, remains unharmed, the plasmodium parasite thrives and replicates within the body system. The parasitism interaction between two organisms Z^i and Z^j with ($i \neq j$) is introduced in the SOS model as follows:

If Z^i is assigned to the anopheles' mosquito which is selected randomly from the search space, fine-tuned its dimension to create an artificial vector or (parasite vector) denoted by Z^p . Similarly, the organism Z^j is equally selected randomly from the search space and serves as host to Z^p . The parasite vector, Z^p , tries to replace the host Z^j in the ecosystem. If Z^p has a higher fitness value than Z^j , Z^p replaces Z^j otherwise, Z^j gains immunity, fights back, and eliminate Z^p from the ecosystem.

The mathematical formulation of selection of species guided by their strength is as in (7) and (8): -

Randomly select Z^j with $i \neq j$

Create a parasite vector Z^p from Z^i using a random number.

If $f(Z^p) > f(Z^j)$ then

$$Z^j \leftarrow Z^p \tag{7}$$

else

$$Z^{j_new} = Z^j \tag{8}$$

III. THE MOST RECENT SOS VERSIONS FOR TASK SCHEDULING PROBLEMS

So far, different implementations of SOS have been introduced due to various type of task scheduling optimization problems. Therefore, it is critically important to examine the enhanced and improved variants of the SOS algorithm as they contribute to task scheduling issues in the cloud. Fig. 2 shows the classification of variant symbiotic organisms search algorithm. Some of the modified SOS algorithms considered in this review are highlighted in Table 1.

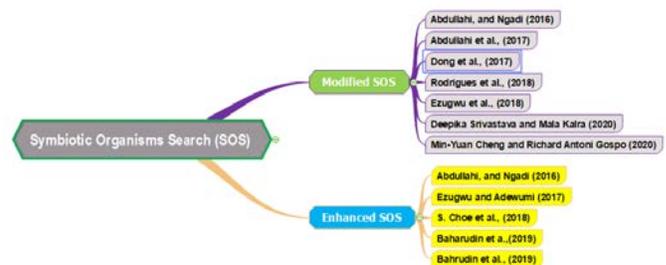


Fig. 2. Taxonomy of Variant Symbiotic Organisms Search Algorithm.

TABLE I. THE SUMMARY OF LITERATURES ON MODIFIED SYMBIOTIC ORGANISM SEARCH (SOS) ALGORITHMS

S/N	Authors	Title	Objectives	Modified technique	Benchmarked Technique	Journal/Conference	Publishers	Remarks
1	Abdullahi, and Ngadi (2016)	Symbiotic Organism Search optimization-based task scheduling in cloud computing environment	Minimizing makespan	DSOS	PSO	Future Generation Computer Systems	Elsevier	T-test analysis of the proposed method revealed that DSOS performs significantly better than PSO, especially for large search spaces.
2	Abdullahi et al., (2017)	“Chaotic symbiotic organisms search for task scheduling optimization on cloud computing environment”	Minimizes cost and makespan	CSOS	SOS, PSO	6th ICT International Student Project Conference	UTM/RUG/15H99 RMC Universiti Teknologi Malaysia.	The key idea is to incorporate a chaotic map, which enlarges the search space and provides diversity, to avoid premature convergence of SOS at early stages of the optimization process.
3	Dong et al., (2017)	CTS-SOS: Cloud task scheduling based on the symbiotic organisms search	Minimizing makespan	CTS-SOS	RR and ACO	Communications in Computer and Information Science	Springer	The results of the experiments demonstrate that CTS-SOS can have improved resource optimization and scheduling, efficiently minimize the makespan, and increase the performance of processing tasks and user satisfaction.
4	Rodrigues et al., (2018)	A Search Algorithm for Flow Shop Scheduling Issues based on Modified Symbiotic Organisms Search Technique	Minimizing makespan	ISOS	SOS	IEEE Congress on Evolutionary Computation (CEC)	IEEE	The proposed modification improved the performance of the SOS algorithm in the search for the global optimum value in most of the instances.
5	Ezugwu et al., (2018)	Symbiotic organisms search algorithm for the unrelated parallel machines scheduling with sequence-dependent setup times	Makespan minimization	SOS-LPT	ACO, ACOII, GADP, SADP, SOS	PLoS ONE	PLOS	The incorporation of the local search improvement mechanism into the SOS scheme has been shown to introduce diversity in the search process and avoid premature convergence.
6	Deepika Srivastava and Mala Kalra (2020)	Improved symbiotic organism search based approach for scheduling jobs in cloud	Makespan minimization	ISOS	PSO	(ICIIL 2019) International conference on IoT Inclusive Life	Springer	The results of the experiments show a better performance of the improved technique over the benchmarked PSO algorithm
7	Min-Yuan Cheng and Richard Antoni Gosno (2020)	SOS 2.0 : an evolutionary approach for SOS algorithm	Minimization of task execution time	SOS 2.0	SOS	Evolutionary Intelligence	Springer	The integration of SPU and SOS 2.0's chaotic sequence provide a significantly improved balance within the algorithm's searching pattern exploration and exploitation.

A. Modified SOS

Modified symbiotic organism search technique is one in which either one or two of the operators are improved to enhance the efficiency of the algorithm in order to optimize fast convergence time while avoiding entrapment at the local optimal region[15][21]. In most cases, either the mutualism or commensalism phases, or both, are strengthened by adding a random weighted reflective parameter or converting continuous outcomes to discrete results among approaches.

A symbiotic organism search optimization-based task scheduling technique has been proposed in the work of [24] to schedule a collection of independent tasks in a cloud computing environment. The proposed technique used various types of distributions to gain insight into the proposed method's performance trend. In addition, to account for the finite nature of cloud resources, a discrete version of SOS was introduced. To assess the suitability of the proposed approach, four types of data set were used in the experimental setup: regular, left-skewed, right-skewed, and uniform distributions. Furthermore, the performance of DSOS over SAPSO improves with increasing search space. According to [25], an investigation to a scheduling issue using a unique chaotic SOS algorithm to save time and expenses has been presented. The key idea is to perform a chaotic map, which expands the search area and generates diversity, to prevent the SOS algorithm from prematurely converging in the early stages of the optimization process.

In the work of [26], an optimized SOS algorithm for use in workflow-scheduling problems has been suggested. The proposed algorithm selects three steps of the SOS algorithm without regard to the population's pre-defined symbiotic relationship. Following the selection of solutions, a special method is selected to assign each solution to a symbiotic relationship. The proposed algorithm's performance was tested using twenty standard samples of the workflow-scheduling problem, and it was found to be more effective than the standard SOS. In another study, Improved Symbiotic Organism Search-Based Approach for Cloud Job Scheduling has been presented by [27]. The technique presents a discrete algorithm based on Improved Symbiotic Organism Search (ISOS) for scheduling independent tasks with the sole aim of mapping of various tasks to the best available cloud resources so as to minimize makespan and computational cost. Experimental results show a better performance of the improved technique over the benchmarked PSO algorithm.

According to [28] a reconstructed mathematical model for cloud task scheduling with additional approximate end time to the proposed design has been presented. By implementing a cloud task scheduling protocol based on the SOS (CTS-SOS) process, a continuous search space is obtained. The CTS-SOS algorithm more accurately optimizes and schedules time, and it effectively reduced the makespan, improved the efficiency of processing operations, and increased customer loyalty. Furthermore, the results of the simulation using CloudSim toolkit shows that the CTS-SOS algorithm outperform Round

Robin and ACO algorithms. Another similar work proposed an SOS algorithm for scheduling unrelated parallel machines with sequence-dependent startup times [29]. The technique developed a new solution of representative and decoding procedure to enhance the capacity of SOS algorithm for heterogeneous parallel machine scheduling problems. The problem of minimizing makespan in the scheduling of unrelated parallel machines with sequence-dependent initialization times has been solved using the proposed technique.

SOS 2.0: An Evolving Solution to the SOS Algorithm has been presented by [30]. The proposed technique integrates self-parameter updating (SPU) method into the search process to enhance its exploration capability. Furthermore, a chaotic map for generating chaotic behaviour was adopted to enhance its exploitation process thereby providing faster convergence. The combination of SPU and chaotic sequence provide adequate balance between local and global search process as well as avoiding premature convergence. The experimental result shows that SOS 2.0 performed better than the basic SOS and other several improved techniques.

B. Hybrid Symbiotic Organisms Search Algorithm

The modified/improved versions of SOS are faced with some challenges related to solving some complex optimization problems that needs extra synergy to resolve. Since there is no single technique that can address all optimization problems, the hybridized algorithms are more likely to be better in dealing with complex optimization issues and as such can produce a better result. The main aim of hybridizing two or more optimization techniques is to harness the complementary benefits of the techniques involved to produce a better solution[29][31][32]. Current literatures on the usage of SOS to address a variety of complex optimization and real-life problems have shown that the implementation of the SOS algorithm needs additional effort to produce high-quality performance. Table 2 highlight the summary of the literatures on hybridized SOS algorithm, while Table 3 presents analysis of some of the performance metrics used to measure the algorithms performance.

An optimization algorithm for task scheduling in cloud computing environment using hybrid symbiotic organisms search technique has been presented in [31]. The proposed technique integrated the Simulated Annealing (SA) method with the basic SOS. The SOS employs fewer control parameters and is highly flexible, has excellent exploration and convergence capabilities. The proposed technique employs SA's systematic reasoning ability to find better solutions on the local region identified by SOS, thereby enhancing SASOS's exploitative ability. Furthermore, a fitness function was introduced to increase the degree of resource utilization thereby minimizing makespan while optimizing the degree of imbalance among the virtual resources. The result of the experiment indicates a better performance of the hybridized technique (SASOS) over the basic SOS algorithm.

TABLE II. THE FOLLOWING IS A SUMMARY OF THE LITERATURES ON HYBRID SYMBIOTIC ORGANISM SEARCH (SOS) ALGORITHMS

S/N	Authors	Title	Objectives	Hybridized Technique	Benchmarked Technique	Journal/Conference	Publishers	Remarks
1	Abdullahi, and Ngadi (2016)	Hybrid Symbiotic Organisms Search Optimization Algorithm for Scheduling of Tasks on Cloud Computing Environment	Minimizing makespan, optimizing job scheduling, and reducing the degree of imbalance between virtual machines	SOSSA	SOS	PloS one	PLOS	In a cloud computing world, the proposed hybrid algorithm (SOSSA) has been shown to outperform the current standard SOS algorithm in terms of task scheduling.
2	Esugwu and Adewumi. (2017)	Soft sets based symbiotic organisms search algorithm for resource discovery in cloud computing environment	Maximizes resource utilization	SSSOS	PSO, SOS, PSOSA,	Future generation computer systems	Elsevier	The hybrid SSSOS algorithm for optimizing VMs resource selection in cloud in minimal time, have displayed a fared well in finding high number of global solutions.
3	Choe et al., (2018)	Improved hybrid symbiotic organism search task-scheduling algorithm for cloud computing	Minimizing makespan	SA-CLS-SOS	SOS, SA-SOS, CLS-SOS	KSII Transactions on Internet and Information Systems	KSII	An acute makespan reduction was achieved.
4	Baharudin et al., (2019)	A Quasi- Oppositional- Chaotic Symbiotic Organisms Search algorithm for global optimization problems	Makespan minimization	QOCSOS	SOS	Applied Soft Computing Journal	Elsevier	The technique QOCSOS, is more viable in dealing with global optimization problem
5	Baharudin et al. (2019)	Multi-Objective Task Scheduling Problems in Cloud Computing Environment based on Symbiotic Organisms Search Algorithm with Chaotic Optimization Strategy.	Minimizing makespan and reduction of task processing cost	CMSOS	MSOS, EMS-C, ECMSMOO and BOGA	Journal of Network and Computer Applications	Elsevier	In terms of Pareto fronts for makespan and cost, the CMSOS algorithm outperformed the compared algorithms.

TABLE III. COMPARISON OF THE EXISTING MODIFIED AND HYBRID SOS TECHNIQUES IN CLOUD TASK SCHEDULING PROBLEM

S/N	Modified /Hybrid Techniques	Settings	Makespan	Cost	VMs Utilization	Response Time	Convergence speed	Performance
1	[24]	Dynamic	Minimizes	Less	High	More	Low	Less
2	[25]	Dynamic	Minimizes	Less	More	More	Low	Less
3	[26]	Dynamic	Minimizes	Less	more	High	Low	Less
4	[27]	Dynamic	Minimizes	Less	Less	More	Low	Less
5	[28]	Dynamic	Minimizes	Less	More	High	Low	Less
6	[29]	Dynamic	Minimizes	Less	Less	Less	Low	Less
7	[30]	Dynamic	Minimizes	Less	High	more	Low	Less
8	[31]	Dynamic	Minimizes	Less	High	Less	Low	More
9	[32]	Dynamic	Minimizes	More	Less	Less	Low	Less
10	[33]	Dynamic	Minimizes	More	Less	more	Average	More
11	[34]	Dynamic	Minimizes	Less	Less	Less	Average	Less
12	[35]	Dynamic	Minimizes	Less	Less	Less	Low	More

An improved task scheduling algorithm for cloud computing using hybrid symbiotic organisms search technique has been presented by [32]. The proposed technique combines the Simulated Annealing (SA) and Chaotic local search (CLS) with SOS optimization technique to improve the convergence rate as well as to avoid being trapped at the local optimal region. Simulation results show that the hybrid SOS outperforms SOS, SA-SOS, and CLS-SOS in terms of convergence speed and makespan.

A soft sets-based symbiotic organisms search algorithm for resource detection has been built in a cloud computing environment [33]. The proposed technique tackles the uncertainty problems of resource selections in cloud computing environment. Furthermore, the hybridized approach incorporates the capabilities of the underlying approaches to efficiently handle tasks that must be performed during cloud resource discovery. As compared to the PSO, SOS, and PSOSA algorithms, the algorithm performed well in terms of finding a large number of good global solutions while minimizing resource collection time, search accuracy, search depth distribution, and the ability to escape premature convergence.

For global optimization problems, a Quasi-Oppositional-Chaotic Symbiotic Organisms Search algorithm has been developed by [34]. For a higher standard of performance and faster convergence, the technique combines Quasi-Opposition-Based Learning (QOBL) and Chaotic Local search methods with traditional SOS. QOBL is used for population initialization and generation hopping, while CLS is used to target the new best population. The results of the experiments show that QOCSOS achieve a better global optimum solution with remarkable convergence speed as compared to the original SOS.

In another study, [35] presents an efficient search algorithm for a multi-objective task scheduling problems based on symbiotic organisms search with a chaotic optimization strategy in a cloud computing environment. The proposed technique employed chaotic strategy to generate initial ecosize (population) as well as using its chaotic sequence to replace the random sequence based component of the SOS phases. This process ensures diversity among species for global convergence. Furthermore, the approach guarantees optimum cost-makespan trade-offs. Therefore, the CMSOS algorithm outperformed the other algorithms in terms of Pareto fronts for task execution time and cost.

The analysis of Table 3 shows that more work needs to be done on Symbiotic Organism Search (SOS) algorithms to improve its convergence speed so as to enhance its performance in cloud task scheduling issues. This can be achieved by modifying the mutual features parameter that exist between two distinct species and by taking into cognizance the contribution of each species in the relation measured by equity rather than equality in the ecosystem. By this improvement, it will enhance equitable allocation/mapping of task to heterogenous VMs as well as balance the search process between local search and global search. This modified technique named G_SOS is in its incubation stage awaiting further improvement and implementation.

IV. CONCLUSION AND FUTURE WORK

Many metaheuristic algorithms have been proposed to solve optimization problems in recent decades. Thus, metaheuristic research around the world has resulted in optimization techniques that have proved to be superior to conventional gradient-based approaches. Symbiotic Organisms Search algorithm (SOS) is a recent SI-based metaheuristic algorithm that was inspired by nature[19]. The SOS algorithm mimics the collaborative behavior observed in nature among distinct species in ecosystem. Though, there have not been any review work on symbiotic organism search algorithm based on task scheduling problems. The modified and enhanced versions of SOS are faced with some challenges when solving some complex optimization problems that needs extra synergy to resolve. Therefore, these algorithms, modified or enhanced, can still be improved by looking inward to the features of the basic (SOS) algorithm.

ACKNOWLEDGMENT

I would like to appreciate the Nigerian Tertiary Education Trust Fund (TETFund) in partnership with Kogi State Polytechnic Lokoja for their support of this work.

REFERENCES

- [1] S. H. H. Madni, M. S. A. Latiff, Y. Coulibaly, and S. M. Abdulhamid, "Recent advancements in resource allocation techniques for cloud computing environment: a systematic review," *Cluster Comput.*, vol. 20, no. 3, pp. 2489–2533, 2017, doi: 10.1007/s10586-016-0684-4.
- [2] S. H. H. Madni, M. S. Abd Latiff, S. M. Abdulhamid, and J. Ali, "Hybrid gradient descent cuckoo search (HGDCS) algorithm for resource scheduling in IaaS cloud computing environment," *Cluster Comput.*, vol. 22, no. 1, pp. 301–334, 2019, doi: 10.1007/s10586-018-2856-x.
- [3] L. Liu, T. Luo, and Y. Du, "A new task scheduling strategy based on improved ant colony algorithm in IaaS layer," *CITS 2019 - Proceeding 2019 Int. Conf. Comput. Inf. Telecommun. Syst.*, no. 61370132, pp. 6–10, 2019, doi: 10.1109/CITS.2019.8862055.
- [4] K. Dubey, M. Kumar, and S. C. Sharma, "Modified HEFT Algorithm for Task Scheduling in Cloud Environment," *Procedia Comput. Sci.*, vol. 125, pp. 725–732, 2018, doi: 10.1016/j.procs.2017.12.093.
- [5] G. Natesan and A. Chokkalingam, "Multi-Objective Task Scheduling Using Hybrid Whale Genetic Optimization Algorithm in Heterogeneous Computing Environment," *Wirel. Pers. Commun.*, vol. 110, no. 4, pp. 1887–1913, 2020, doi: 10.1007/s11277-019-06817-w.
- [6] Z. Tong, H. Chen, X. Deng, K. Li, and K. Li, "A novel task scheduling scheme in a cloud computing environment using hybrid biogeography-based optimization," *Soft Comput.*, vol. 23, no. 21, pp. 11035–11054, 2019, doi: 10.1007/s00500-018-3657-0.
- [7] A. Shabani, B. Asgarian, S. A. Gharebaghi, M. A. Salido, and A. Giret, "A New Optimization Algorithm Based on Search and Rescue Operations," *Math. Probl. Eng.*, vol. 2019, 2019, doi: 10.1155/2019/2482543.
- [8] A. E. Ezugwu, "Enhanced symbiotic organisms search algorithm for unrelated parallel machines manufacturing scheduling with setup times," *Knowledge-Based Syst.*, vol. 172, pp. 15–32, 2019, doi: 10.1016/j.knosys.2019.02.005.
- [9] X. S. Yang, "Nature-inspired optimization algorithms: Challenges and open problems," *J. Comput. Sci.*, vol. 46, p. 101104, 2020, doi: 10.1016/j.jocs.2020.101104.
- [10] D. Gabi, A. S. Ismail, A. Zainal, and Z. Zakaria, "Quality of service task scheduling algorithm for time-cost trade off scheduling problem in cloud computing environment," *Int. J. Intell. Syst. Technol. Appl.*, vol. 18, no. 5, pp. 448–469, 2019, doi: 10.1504/IJISTA.2019.101952.
- [11] X.-S. Yang, *Nature-Inspired Algorithms and Applied Optimization*, vol. 744, no. January. 2018.

- [12] E. Cuervas, D. Zaldivar, and M. Perez, "Advances in Metaheuristics Algorithms_ Methods and Applications," vol. 775, no. April, p. 229, 2018, doi: 10.1007/978-3-319-89309-9.
- [13] T. Dokeroglu, E. Sevinc, T. Kucukyilmaz, and A. Cosar, "A survey on new generation metaheuristic algorithms," *Comput. Ind. Eng.*, vol. 137, no. May, p. 106040, 2019, doi: 10.1016/j.cie.2019.106040.
- [14] N. M. A. Samee, S. S. Ahmed, and R. A. A. A. A. Scoud, "Metaheuristic algorithms for independent task scheduling in symmetric and asymmetric cloud computing environment," *J. Comput. Sci.*, vol. 15, no. 4, pp. 594–611, 2019, doi: 10.3844/jcssp.2019.594.611.
- [15] M. Abdullahi, M. A. Ngadi, S. I. Dishing, S. M. Abdulhamid, and M. J. Usman, "A survey of symbiotic organisms search algorithms and applications," *Neural Comput. Appl.*, vol. 32, no. 2, pp. 547–566, 2020, doi: 10.1007/s00521-019-04170-4.
- [16] H. Shayanfar and F. S. Gharehchopogh, "Farmland fertility: A new metaheuristic algorithm for solving continuous optimization problems," *Appl. Soft Comput. J.*, vol. 71, pp. 728–746, 2018, doi: 10.1016/j.asoc.2018.07.033.
- [17] D. Gabi, A. S. Ismail, A. Zainal, Z. Zakaria, and A. Abraham, "Orthogonal Taguchi-based cat algorithm for solving task scheduling problem in cloud computing," *Neural Comput. Appl.*, vol. 30, no. 6, pp. 1845–1863, 2018, doi: 10.1007/s00521-016-2816-4.
- [18] A. A. Zubair, S. B. A. Razak, M. A. B. Ngadi, A. Ahmed, and S. H. H. Madni, "Convergence-based task scheduling techniques in cloud computing: A review," vol. 1073, 2020.
- [19] M. Y. Cheng and D. Prayogo, "Symbiotic Organisms Search: A new metaheuristic optimization algorithm," *Comput. Struct.*, vol. 139, pp. 98–112, 2014, doi: 10.1016/j.compstruc.2014.03.007.
- [20] G. Xiong, J. Zhang, X. Yuan, D. Shi, and Y. He, "Application of symbiotic organisms search algorithm for parameter extraction of solar cell models," *Appl. Sci.*, vol. 8, no. 11, 2018, doi: 10.3390/app8112155.
- [21] A. E. Ezugwu and D. Prayogo, "Symbiotic Organisms Search Algorithm: theory, recent advances and applications," *Expert Syst. Appl.*, vol. 119, pp. 184–209, 2019, doi: 10.1016/j.eswa.2018.10.045.
- [22] Y. Zhou, H. Wu, Q. Luo, and M. Abdel-Baset, "Automatic data clustering using nature-inspired symbiotic organism search algorithm," *Knowledge-Based Syst.*, vol. 163, pp. 546–557, 2019, doi: 10.1016/j.knosys.2018.09.013.
- [23] G. G. Tejani, N. Pholdee, S. Bureerat, and D. Prayogo, "Multiobjective adaptive symbiotic organisms search for truss optimization problems," *Knowledge-Based Syst.*, vol. 161, pp. 398–414, 2018, doi: 10.1016/j.knosys.2018.08.005.
- [24] M. Abdullahi, M. A. Ngadi, and S. M. Abdulhamid, "Symbiotic Organism Search optimization based task scheduling in cloud computing environment," *Futur. Gener. Comput. Syst.*, vol. 56, pp. 640–650, 2016, doi: 10.1016/j.future.2015.08.006.
- [25] M. Abdullahi, M. Asri Ngadi, and S. I. Dishing, "Chaotic symbiotic organisms search for task scheduling optimization on cloud computing environment," 6th ICT Int. Student Proj. Conf. Elev. Community Through ICT, ICT-ISPC 2017, vol. 2017-Janua, pp. 1–4, 2017, doi: 10.1109/ICT-ISPC.2017.8075340.
- [26] L. R. Rodrigues, J. P. P. Gomes, A. R. R. Neto, and A. H. Souza, "A Modified Symbiotic Organisms Search Algorithm Applied to Flow Shop Scheduling Problems," 2018 IEEE Congr. Evol. Comput. CEC 2018 - Proc., p. 8477846, 2018, doi: 10.1109/CEC.2018.8477846.
- [27] D. Srivastava and M. Kalra, "Improved symbiotic organism search based approach for scheduling jobs in cloud," *Lect. Notes Networks Syst.*, vol. 116, no. Iciiil, pp. 453–461, 2020, doi: 10.1007/978-981-15-3020-3_39.
- [28] Z. Liu, X. Liu, Y. Dong, X. Zhao, and B. Zhang, "CTS-SOS: Cloud task scheduling based on the symbiotic organisms search," *Commun. Comput. Inf. Sci.*, vol. 729, pp. 82–94, 2017, doi: 10.1007/978-981-10-6442-5_8.
- [29] A. E. Ezugwu, O. J. Adeleke, and S. Viriri, "Symbiotic organisms search algorithm for the unrelated parallel machines scheduling with sequence-dependent setup times," *PLoS One*, vol. 13, no. 7, pp. 1–23, 2018, doi: 10.1371/journal.pone.0200030.
- [30] M. Yuan, C. Richard, and A. Gosno, "SOS 2.0: an evolutionary approach for SOS algorithm," *Evol. Intell.*, no. 0123456789, 2020, doi: 10.1007/s12065-020-00476-8.
- [31] M. Abdullahi and M. A. Ngadi, "Hybrid symbiotic organisms search optimization algorithm for scheduling of tasks on cloud computing environment," *PLoS One*, vol. 11, no. 6, pp. 1–29, 2016, doi: 10.1371/journal.pone.0158229.
- [32] S. Choe, B. Li, I. Ri, C. Paek, J. Rim, and S. Yun, "Improved hybrid symbiotic organism search task-scheduling algorithm for cloud computing," *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 8, pp. 3516–3541, 2018, doi: 10.3837/tiis.2018.08.001.
- [33] A. E. Ezugwu and A. O. Adewumi, "Soft sets based symbiotic organisms search algorithm for resource discovery in cloud computing environment," *Futur. Gener. Comput. Syst.*, vol. 76, pp. 33–50, 2017, doi: 10.1016/j.future.2017.05.024.
- [34] K. H. Truong, P. Nallagownden, Z. Baharudin, and D. N. Vo, "A Quasi- Oppositional-Chaotic Symbiotic Organisms Search algorithm for global optimization problems," *Appl. Soft Comput. J.*, vol. 77, pp. 567–583, 2019, doi: 10.1016/j.asoc.2019.01.043.
- [35] M. Abdullahi, M. A. Ngadi, S. I. Dishing, S. M. Abdulhamid, and B. I. eel Ahmad, "An efficient symbiotic organisms search algorithm with chaotic optimization strategy for multi-objective task scheduling problems in cloud computing environment," *J. Netw. Comput. Appl.*, vol. 133, no. July 2018, pp. 60–74, 2019, doi: 10.1016/j.jnca.2019.02.005.

Sarcasm Detection in Tweets: A Feature-based Approach using Supervised Machine Learning Models

Arifur Rahaman¹, Ratnadip Kuri², Syful Islam^{3*}, Md. Javed Hossain⁴, Mohammed Humayun Kabir⁵

Dept. of Computer Science and Telecommunication Engineering
Noakhali Science and Technology University, Bangladesh^{1,2,3,4,5}

Dept. of Computer Science and Engineering, Sonargaon University, Bangladesh¹

Abstract—Sarcasm (i.e., the use of irony to mock or convey contempt) detection in tweets and other social media platforms is one of the problems facing the regulation and moderation of social media content. Sarcasm is difficult to detect, even for humans, due to the deliberate ambiguity in using words. Existing approaches to automatic sarcasm detection primarily rely on lexical and linguistic cues. However, these approaches have produced little or no significant improvement in terms of the accuracy of sentiment. We propose implementing a robust and efficient system to detect sarcasm to improve accuracy for sentiment analysis. In this study, four sets of features include various types of sarcasm commonly used in social media. These feature sets are used to classify tweets into sarcastic and non-sarcastic. This study reveals a sarcastic feature set with an effective supervised machine learning model, leading to better accuracy. Results show that Decision Tree (91.84%) and Random Forest (91.90%) outperform in terms of accuracy compared to other supervised machine learning algorithms for the right features selection. The paper has highlighted the suitable supervised machine learning models along with its appropriate feature set for detecting sarcasm in tweets.

Keywords—Machine learning; detection; sarcasm; sentiment; tweets

I. INTRODUCTION

Sarcasm detection in opinion mining is an essential tool with various applications, including health, security, and sales [1, 2]. Several organizations and companies have shown their interest in studying tweets data to know people's opinion towards popular products, political events or movies. Millions of tweets are posted every day, which increase the content of twitter tremendously.

However, microblogging social media (i.e., maximum 140 characters in every tweet) and containing informal language essentially makes it quite tricky to understand users' sentiment and perform sentiment analysis. Additionally, sarcasm poses a challenge in sentiment analysis and causes the misclassification of people's opinions. Hence, it leads to reduced accuracy of sentiment analysis. People use sarcasm to mock or convey contempt through a sentence or while speaking. People apply positive words to reveal gloomy feelings. In recent days, sarcasm or irony is very common in social media, although it is challenging to detect.

The cutting-edge approaches of opinion mining and sentiment analysis are prone to unsatisfactory performances while analyzing social media data. Maynard and others [3] proposed that detecting sarcasm during sentiment analysis might significantly improve performance. Consequently, the necessity for an effective method to identify sarcasm arises.

In this paper, we propose an effective method to identify a sarcastic tweet. Our strategy considers the various types of sarcasm indicating features such as Lexical-based Features, Sarcastic-based Features, Contrast-based Features, Context-based Features, and detects the sarcastic tweets using multiple supervised machine learning models based on extracted features. We suggest an effective machine learning model and feature set to better perform sarcasm detection in sentiment analysis and get better accuracy, which is explained at the end of the evaluation and the result analysis parts of this paper.

The main contributions are as follows:

- 1) To detect sarcasm, we use a set of machine learning classification algorithms along with a variety of features to identify the best classifier model with significant features, which leads to recognize of the sarcasm in tweets to get better performance of sentiment analysis.
- 2) We propose the right set of features that lead to better accuracy, which is presented in the result analysis part of this paper.
- 3) Analysis results present that Decision Tree (91.84%) and Random Forest (91.90%) outperform the accuracy compared to Logistic Regression, Gaussian Naive Bayes, and Support Vector Machine for the different features set up.

The remainder of this paper is arranged as follows: Section 2 explains the literature review, and Section 3 demonstrates the sarcasm recognition process. Section 4 illustrates our findings, and Section 5 concludes this work.

II. LITERATURE REVIEW

Many research articles and publications motivated us to work with this topic; a few of them are discussed here in detail. The authors, Sana Parveen, Sachin N. Deshmukh [4], suggested a methodology to recognize the sarcasm on Twitter using Maximum Entropy and Support Vector Machine (SVM). Firstly, they created two datasets from collected

Twitter data. One dataset has sarcastic tweets and another dataset without sarcastic tweets in training data. Penn treebank is used to tag POS with each word. Maximum Entropy and SVM are used to classify tweets after features extraction related to sentiment, syntactic, punctuation, and pattern. Finally, they got more accuracy for Maximum Entropy than SVM. The authors, Anukarsh G Prasad, Sanjana S, Skanda M Bhat, BS Harish [5], suggested a technique based on the slang and emojis used in tweets to identify sarcastic and non-sarcastic tweets. They took into consideration slang and emoji values according to the slang and the emoji dictionary. Afterward, these values are used to classify sarcasm, applying Random Forest, Gaussian Naive Bayes, Gradient Boosting, Logistic Regression, Adaptive Boost, Logistic Regression, and Decision Tree. They suggested an effective classifier using slang and emoji dictionary mapping to produce the most satisfactory performance. A new technique was suggested by Sreelakshmi K, Rafeeqe P C [6] using context incongruity to define sarcasm on Twitter. To recognize the irony, they considered various features such as linguistic, sentiment, and context features. Both the sarcastic and non-sarcastic tweets are collected and pre-processed. They used Simple Vector Machine (SVM) and Decision Tree as a sarcasm classifier and produced a satisfactory level. A significant study from the linguistic sector tells us that lexical factors such as adjectives and adverbs, interjections, and punctuations play a considerable role in sarcasm detection [7]. In sentiment analysis, many researchers used machine learning methods such as Maximum Entropy, Naïve Bayes, and Support Vector machine because these algorithms tend to outperform the other algorithms in text classification [8,9]. Buscaldi and others (2012) [10] addressed the features which lead to the sarcasm classification. It also provides an in-depth description of how the different features contribute to the classification. Barbieri F. And Saggion H.'s work (2014) [11] dealt with the automatic detection of sarcasm on Twitter data. They divided the results of the classification of tweets into sarcastic or non-sarcastic classes. They do it depending on frequency (the gap between rare and common words), written-spoken (written-spoken style uses), intensity (intensity of adverbs and adjectives). And also, analyze structure (length, punctuation, emoticons, links), sentiments (the gap between positive and negative terms), synonyms (common vs. rare synonyms use), and Ambiguity (a measure of possible ambiguities). Based on the features mentioned above, they proposed a classification algorithm and claimed 71% accuracy on irony detection. The authors, David B. and Noah A. S. (2016) [12], improved the classification method by adding the history of tweets and author profiles, which helps in the classification process. The article presents accuracy for different contexts ranging from 70% and

upwards. Parmar and others (2018) [13] pointed out a few of the present challenges to classify sarcastic tweets, such as the nature of the collected tweet, the presence of uncommon words, abbreviations, and slang, that are more informal and no predefined structure for sarcasm identification in Twitter. Ren, Y., Ji, D., and Ren, H. suggested two distinct context-augmented neural models for detecting sarcasm in the text [14]. Prasad and others [15] examined numerous classification approaches in which they noticed that Gradient Boost (GB) showed the best performance. Another study stated a novel method for identifying sarcasm in tweets by combining two classification techniques, Support Vector Machine with N-gram features [16]. Karthik Sundararajan and others proposed a sarcasm detecting approach and irony type using Multi-Rule Based Ensemble Feature Selection Model [17]. Siti Khotija, Khadijah Tirtawangsa, Arie A Suryani suggested a context-based method to detect sarcasm in tweets based on Long Short-Term Memory (LSTM) [18]. According to another paper, classifier's performance is vital in sarcasm predictions in opinion mining [19].

Moreover, the nature of classifiers can also play an essential role in sarcasm detection. However, in tweets, limited studies have approached the efficiency of sarcasm detection methods with valuable features. Hence, this study investigates the principal sarcasm classifiers, classification performance, and the selection of features that dominate such performance.

Besides, the paper drives motivation from numerous works, as mentioned earlier, and intends to produce better performance in sarcasm detection.

III. PROPOSED METHODOLOGY

The structure of suggested methodology is depicted in Fig. 1 and mainly consists of three modules such as tweet pre-processing, feature engineering, and sarcasm recognition modules.

A. Tweets Dataset

Twitter's streaming API was used to collect tweets. To obtain sarcastic tweets, we request the API for tweets having the hashtag "#sarcasm". Similarly, for no sarcastic tweets, we collected tweets regarding different topics and eliminated ones that include any hashtag indicating sarcasm. We collect a total of 76799 tweets having two categories, sarcastic (37583) and non-sarcastic (39216) tweets. 0(zero) is used to indicate non-sarcastic tweets, while 1(one) for sarcastic tweets. The dataset contains two columns as Label and Tweet. The Label column has 1 or 0 to present, whether it is sarcastic or not while the Tweet column has tweets.

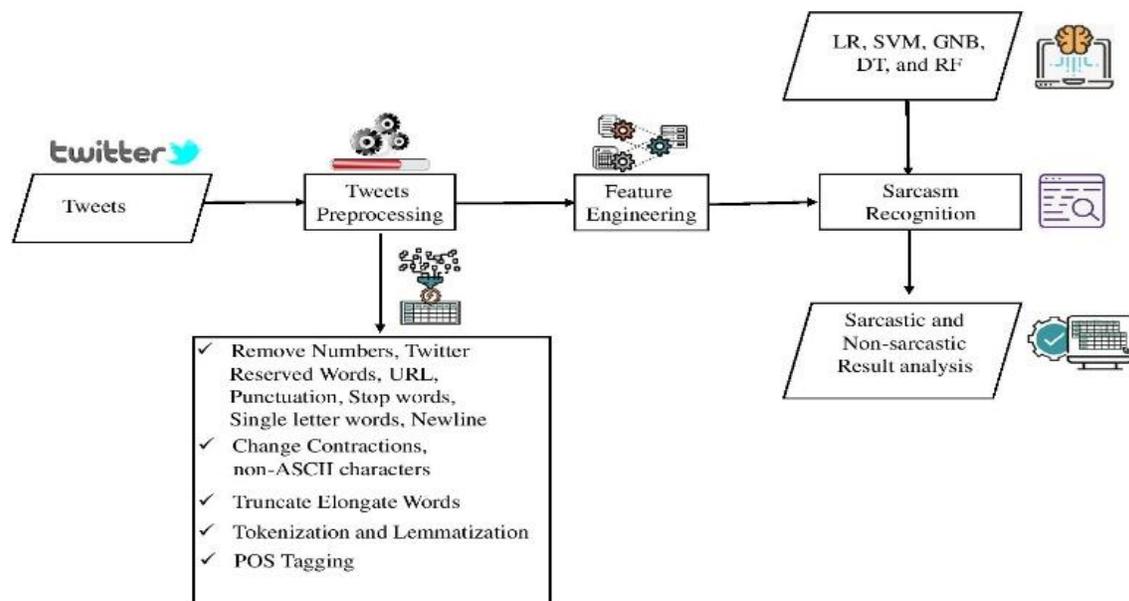


Fig. 1. Proposed Methodology.

B. Tweets Preprocessing

1) *Noise removal*: We have removed numbers, newlines, non-ASCII characters, twitter preserved words, a single word tweet and some common string literals to speed up the feature mining process.

2) *URLs removal*: URLs (Uniform Resource Locators) in the tweets are references to a location on the web but do not provide any additional information. They are removed in the pre-processing phase of the sarcasm detection process.

3) *Stop words removal*: One of the significant forms of pre-processing is to filter out useless data referred to as stop words. Stop words include mainly articles (a, an, the), prepositions (in, of, to, and so on), along with other very commonly used words. They, indeed, don't have any contribution to detect sarcasm in the sentence. Therefore, they are removed before further processing of data. For example, after removing common words from "I don't know how to swim", we have "know and swim" words remaining.

4) *Truncated elongated word*: Tweets contain repeated characters in a word such as goooooood, loooove, moooove and many more styles. These words usually indicate sarcasm in tweets. We did not apply this preprocess step while counting the number of repeated letter segments and vowel repetitions in the tweet. However, we get the base word for these words, for instance, good, love, move, respectively, to extract other features.

5) *Punctuation removal*: Punctuations such as "#\$%&'()*+,-./:;<=>@[\\^_`{|}~" are discarded before extracting features from tweets, though some special punctuations, for example, ?, !, ... are preserved which sometimes indicate sign of sarcasm.

6) *Replace contraction and acronym*: Contractions are n't, aren't, wasn't, can't, couldn't, haven't, won't, shan't, shouldn't and many more. We replaced all of them with their

full form so that we can analyze a single word. Moreover, we use forms of common acronyms used in tweets such as hella, lhh, lmao, jk and so on.

7) *Normalization*: To normalize tweets, we apply tokenization and lemmatization steps. Tokenization helps to create words list from tweets while lemmatization finds the base form of any provided word, for instance, made to make and loved to love. The base form of a word, in most cases, supports identifying tweet sentiment.

8) *POS Tagging*: It is the method of matching a word to its morphological class, which supports learning its role inside the sentence. Necessary parts of speech counted in POS Tagging are Noun, Verb, Adverbs, and Adjectives. Part-of-speech taggers essentially take a series of words as input and produce a list of tuples as output, where every word is connected with the relevant tag.

C. Feature Engineering

Being a modern form of communication, sarcasm is used for various purposes that fall in prevailing in three categories:

1) *Irony as wit*: when used as a wit, sarcasm is applied to be funny; the person applies some particular sorts of speeches, favors to exaggerate, or adopts a distinct tone from that when he speaks usually. In social networks, voice tones are transformed into particular kinds of writing: use of uppercase letter words, ellipsis, letter repetition, quote repetition, question marks, interjections, and exclamation, as well as some sarcasm-related emoticons.

2) *Irony as whimper*: when used as a whimper, sarcasm is used to reveal how annoyed or irritated the person is. Consequently, it stimulates to explain how wicked the circumstance is, using exaggeration and highly positive feelings to express a negative state.

3) *Irony as an escape*: it attributes to the circumstance when the person wishes to pretend to give a precise answer. Hence, it makes the presence of sarcasm. In this case, the person applies perplexing sentences, unusual words, some common expressions, and slangs.

We extracted four sets of features based on the assumption mentioned above: Lexical, Sarcastic, Contrast, and Context-based.

4) *Lexical-based features*: We extract seven lexical or textual-related features. They are Noun, verb, adverb, adjective having a higher impact in a tweet than any other parts of speech. We count them individually according to individual tweets. Moreover, intensifiers such as absolutely, amazingly, awfully, ridiculously, bloody, excessively, outrageously, strikingly, tremendously and so on, are sometimes used to show exaggeration in the tweet to express negative feelings through positive intensifiers and vice versa. Therefore, we also count the number of positive and negative intensifiers present in every tweet. Lastly, the whole tweet's sentiment is calculated, which reveals the overall polarity of the tweet.

5) *Sarcastic-based features*: Generally, people tend to make complicated sentences or use rare words to make them vague to the listener or reader to get a definite answer in other events. Indeed, when people use sarcasm as avoidance, they intend to hide their actual feeling or sentiment by hiding them in fun. Therefore, we derive the following features. People sometimes try to convey the irony message through punctuations, such as exclamation, question marks, and repeated ellipsis. Hence, we count the number of exclamations, question marks and repeated ellipsis. Additionally, interjections are 'ha-ha', 'ho-ho', 'ho-ho-ho', 'oh', 'ouch', 'ow', 'shh', 'super' 'kidding', 'ah', 'aha', 'aww', 'nah', 'yay', 'uh', 'bah', 'bingo', 'boo', 'bravo', 'brilliant' and many more. People use them to express their feeling in different ways, which help to identify sarcastic tweet. Words like loooove, goooooood, mooooove having repeated letters more than three or equals to three also probable indications of mockery tweets and repetition of vowels also denote the same thing. So, we count the number of repeated letter segments and vowel repetition in the tweet. It is worthy to note here, we extract this feature before removing repeated letters. On the other hand, people apply the uppercase word, for instance, GOOD, AWESOME and part of the word as uppercase, for example, gOOD, aWESome, to reveal their irony feeling. We find their number as well. We compute the number of laughter namely 'lol', 'lhh', 'jk', 'wow', 'kidding', 'ha ha', 'ha-ha', 'haha', 'rofl', 'roflmao', 'lmao', 'wtf'. Emoji are the facial emotions such as laughter created by typing a series of keyboard symbols, which are normally used to convey the author's attitude, feeling, or intended tone. In particular, sarcastic Emoji are ones sometimes used with sarcastic or ironical statements (e.g., ":P"). We consider not only rare sarcastic words but also very common words used in sarcastic tweets.

Sarcastic slangs as examples of 'ayfkmwts', 'fubar', 'kyso', 'lhh', 'lmao', 'stfw', 'stf', 'roflmao', 'rofl' are used in the tweet, and we take into consideration them too. We calculate the number of repeated quotes in the tweet as letter repetition. Besides, hashtags also contain emotional content.

In many cases, they are employed to ambiguate the tweet's real purpose carried in the message. Accordingly, we also calculate the sentiment score of hashtags. Finally, we calculate polarity summation (+1 for positive whereas -1 for negative polarity) for n-grams such as bigrams and trigrams.

6) *Contrast-based features*: We then outline four features that interpret whether there is a contrast between the various elements. By comparison, we indicate a negative element's coexistence and a positive one within the identical tweet. We calculate emoji polarity flip, polarity flip of sentiment, the number of positive and negative words. Before counting positive and negative words, we convert the following word of a negation word (e.g., "not", "never", etc.) into antonym, then find the number of positive words and that of negative words.

7) *Context-based features*: For context-related features, we find the number of users mentions and hashtags included in the tweet. People tend to express their feelings by applying various types of hashtags and sometimes use user mention as well.

D. Sarcasm Recognition

Usually, supervised machine learning models and lexicon-based approaches are used in opinion mining and text classification. The former includes various techniques such as K-Nearest Neighbor, Support Vector Machine (SVM), Linear Regression, Logistic Regression (LR) [20], Gaussian Naive Bayes (GNB)¹, Decision Tree (DT), Random Forest (RF), Neural Networks, Linear Discriminant Analysis (LDA) and so on. The latter has two strategies: a dictionary-based approach and a corpus-based approach.

However, for sarcasm detection in our proposed approach, we use five machine learning classifiers, such as Support Vector Machine, Logistic Regression, Gaussian Naive Bayes, Decision Tree and Random Forest, to examine which one performs better with our extracted features.

We commonly need two sets of data in machine learning-based classification, such as training data and test data sets. The classifier uses training data set to learn from the data and build the model, while the test data set is applied to validate the classifier's performance. In our case, we use the 10-fold cross-validation technique to split the dataset into training and test sets. Cross-validation is a method to evaluate predictive models by splitting the dataset into a training data set for training the model and a test data set to evaluate it.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Evaluation

We have evaluated the effective classifier and feature set depending on three metrics: precision, recall, and F1-score.

¹ <https://iq.opengenus.org/gaussian-naive-bayes/>

Precision: Precision presents the ratio between the True Positives and all actual Positives in the dataset. Additionally, precision indicates the number of Positive class predictions which belong to the Positive class. It denotes TP (True Positive) divided by the sum of TP (True Positive) and (FP) False Positive.

Recall: Recall presents the percentage of correctly identifying the True Positive class. It denotes TP (True Positive) divided by the sum of TP (True Positive) and (FN) False Negative.

F1-score: F1-score presents the weighted or Harmonic average of Precision and Recall. Hence, it takes into consideration both False Positive and False Negatives. The formula denotes it: $(2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$.

1) *Evaluating effective sarcasm classifier:* We have evaluated five machine learning classifiers depending on precision, recall and F1-score to investigate which outperforms. According to Table I, Random Forest shows the highest precision, and Naïve Bayes presents the lowest value for precision. Regarding Recall and F1 metrics, the Decision Tree classifier performs better than other classifiers, 91.72% and 91.67%, respectively. Overall, the Decision tree outperforms different classifiers.

2) *Evaluating of effective feature set:* We have evaluated four feature sets with three popular classifiers, such as Support Vector Machine, Decision Tree and Random Forest. Here, also we have investigated three metrics: precision, recall and F1-score. According to our evaluation presented in Table II, the Sarcasmic feature set produces consistently better values for precision, recall and F1 than other feature sets. In contrast, the context-based feature set shows the lowest values in all cases.

B. Results

This section is explained in three separate subsections as follows:

- Investigation of the accuracy metric depending on various features.
- Explanation of our observation for various feature combinations.
- Analysis of the variation in classification accuracy for adding categories incrementally.

We have used accuracy metric to perform results analysis.

Accuracy: Accuracy indicates the ratio between the total number of accurate predictions and the total number of possible predictions in test data. It is calculated as the sum of TP (True Positive) and TN (True Negative) divided by the sum of total Positives and Negative classes.

1) *Performances of each set of features:* According to the given Table III and Fig. 2 above, it is clear that sarcasmic features in our study have a much better contribution in sarcasm detection in tweets than any other features such as lexical, contrast, and context features. For the rest of the

feature sets, lexical features have better accuracy than contrast-based features, while context-based features have far less impact on sarcasm detection. Decision Tree performs consistently better in terms of classification technique, and it reaches around 80 percent accuracy for sarcastic features merely. Logistic Regression shows the lowest accuracy for lexical and sarcastic features, while contrast and context features are less useful for the Random Forest technique. As sarcastic features produce far more accuracy, it is worthy to note that the proper selection of sarcastic feature set can increase the accuracy instead of selecting any other features.

TABLE I. EVALUATION OF EFFECTIVE SARCASM DETECTION CLASSIFIER. RESULT SHOWS THAT THE DECISION TREE OUTPERFORMS OTHER CLASSIFIERS

Classifiers	Precision	Recall	F1-score
LR	72.12	42.43	53.43
SVM	90.06	85.93	85.36
GNB	63.43	84.18	72.35
DT	91.62	91.72	91.67
RF	91.9	91.5	90.33

TABLE II. EVALUATION OF EFFECTIVE FEATURE SET ALONG WITH VARIOUS CLASSIFIERS. RESULT SHOWS THAT THE SARCASMIC FEATURE SET PRODUCES CONSISTENTLY BETTER VALUES FOR PRECISION, RECALL AND F1 THAN OTHER FEATURE SETS

Classifiers	Features	Precision	Recall	F1-score
SVM	Lexical	58.85	62.79	60.75
	Sarcasmic	77.83	68.05	72.61
	Contrast	60.99	35.68	45.02
	Context	0	0	0
DT	Lexical	64.68	60.77	62.67
	Sarcasmic	82.57	77.72	80.07
	Contrast	57.71	74.76	65.14
	Context	55.59	37.5	44.79
RF	Lexical	56.69	53.31	54.95
	Sarcasmic	80.19	74.98	77.5
	Contrast	39.04	41.73	40.34
	Context	21.93	20.36	21.11

TABLE III. ACCURACY (%) ON EVERY FEATURE SET. OUR ANALYSIS SHOWS THAT SARCASMIC FEATURES ARE USEFUL TO OBTAIN HIGHER CLASSIFICATION ACCURACY THAN OTHERS

Classifier	Lexical Features	Sarcasmic Features	Contrast Features	Context Features
LR	51.10	63.83	51.10	51.10
SVM	60.33	74.90	57.39	51.10
GNB	59.83	62.19	59.56	49.96
DT	64.59	81.08	60.87	54.79
RF	57.25	78.71	39.65	25.61

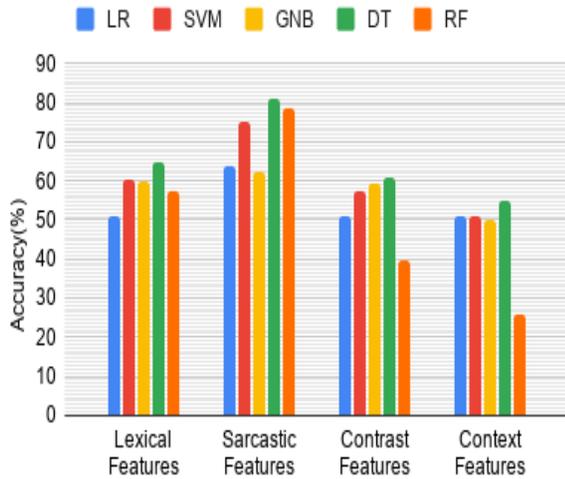


Fig. 2. Graphical view of Accuracy (%) on Every Feature Category. We Observe that Sarcastic Features are Dominating to get Higher Sarcasm Detection Accuracy.

2) *Accuracy for various feature combinations:* As of Table IV and Fig. 3 mentioned below, sarcastic (S) and contrast (C) based features together show the highest accuracy for DT and RF than the other feature combinations. The analysis result shows that lexical (L) and sarcastic feature combinations achieve more accuracy than sarcastic and context features. Mainly, GNB produces the highest accuracy for lexical and sarcastic features set. For all classification techniques, contrast and context (Cx) features together represent the lowest accuracy. As explained, except for lexical and sarcastic feature sets, it is found that the sarcastic and contrast features combinedly are the minimal feature set which leads to getting higher accuracy than the remaining combination sets. Although the Decision Tree classification technique has consistently higher accuracy than random forest and others, it exhibits lower accuracy for lexical and sarcastic feature combinations. Naïve Bayes and Logistic Regression are stabilized with almost 63% accuracy, while accuracy for SVM fluctuates between around 58% and 87%. Overall, the Decision Tree produces maximum accuracy, nearly 91% for the minimal sarcastic and context-based feature set.

TABLE IV. ACCURACY (%) FOR COMBINED FEATURE SETS. SARCASTIC (S) AND CONTRAST (C) BASED FEATURES TOGETHER SHOW THE HIGHEST ACCURACY FOR DT AND RF THAN THE OTHER FEATURE COMBINATIONS

Classifier	L + S	S + C	S + Cx	C + Cx
LR	63.83	63.83	63.83	51.10
SVM	79.34	86.59	74.90	57.39
GNB	67.95	63.44	62.63	60.32
DT	82.79	90.99	81.86	64.54
RF	84.66	89.81	79.38	57.21

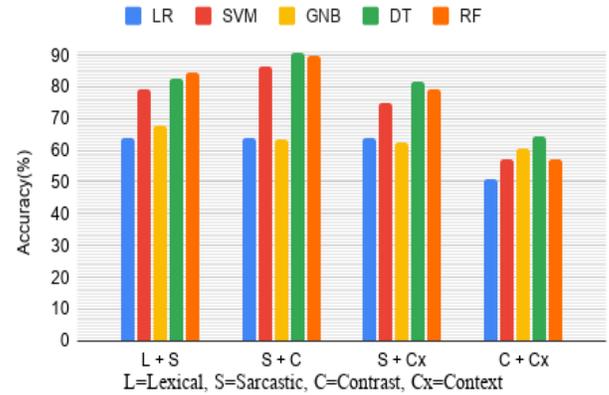


Fig. 3. Accuracy for Various Category Combinations. Sarcastic (S) and Contrast (C) based Features Produce the Highest Accuracy for DT and RF than the other Feature Combinations.

3) *Accuracy for incrementally added category:* Depending on our findings as of Table V and Fig. 4, we observe that the sarcastic-based features dominated the accuracy metric. We fixed them as a base feature set and added feature sets incrementally to understand how accuracy changed with more feature sets. The addition of two (sarcastic, context) and three (sarcastic, contrast and lexical) feature sets show almost similar accuracy. Therefore, it is observable that lexical features with sarcastic and contrast features have far less impact on accuracy. It is not necessary to add more features to get better accuracy at all. Analysis results reveal that Decision Tree (91.84%) and Random Forest (91.90%) exceed the accuracy compared to Logistic Regression, Gaussian Naive Bayes, and Support Vector Machine for the different features selection.

C. Discussion

Looking at the values obtained in our sarcasm recognition experiments, it seems that the sarcastic feature set has more contribution to sarcasm detection in tweets among the four feature sets. In fact, this feature set's accuracy is constantly more excellent than other feature sets with all classifiers, and the Decision Tree shows maximum accuracy (approximately 81%) for the sarcastic feature set.

TABLE V. ACCURACY (%) FOR THE INCREMENTALLY ADDED FEATURE SETS. WE FOUND THAT THE SARCASTIC BASED FEATURES DOMINATED THE ACCURACY METRIC, AND ANALYSIS RESULTS REVEAL THAT DECISION TREE (91.84%) AND RANDOM FOREST (91.90%) PRESENT HIGHER PERFORMANCE

Classifier	S	S + C	S + C + L	S + C + L + Cx
LR	63.83	63.83	63.83	63.83
SVM	74.90	86.59	86.59	88.48
GNB	62.19	63.44	64.16	68.53
DT	81.07	91.04	91.51	91.84
RF	78.71	89.81	90.60	91.90

ACKNOWLEDGMENT

This research work has been supported by Research Cell, Noakhali Science and Technology University, Noakhali-3814.

REFERENCES

- [1] S. Jain, V. Hsu, "The lowest form of wit: Identifying sarcasm in social media," 2015.
- [2] S. Saha, J. Yadav, P. Ranjan, "Proposed approach for sarcasm detection in twitter," *Indian Journal of Science and Technology*, 10(25), 1–8, 2017.
- [3] D. G. Maynard, M. A. Greenwood, "Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis," in *Lrec 2014 proceedings*, ELRA, 2014.
- [4] S. Parveen, S. N. Deshmukh, "Opinion Mining in Twitter–Sarcasm Detection," *Politics*, 1200, 125, 2017.
- [5] A. G. Prasad, S. Sanjana, S. M. Bhat, B. Harish, "Sentiment analysis for sarcasm detection on streaming short text data," in *2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA)*, 1–5, IEEE, 2017.
- [6] K. Sreelakshmi, P. Rafeeqe, "An effective approach for detection of sarcasm in tweets," in *2018 International CET Conference on Control, Communication, and Computing (IC4)*, 377–382, IEEE, 2018.
- [7] R. Kreuz, G. Caucci, "Lexical influences on the perception of sarcasm," in *Proceedings of the Workshop on computational approaches to Figurative Language*, 1–4, 2007.
- [8] A. R. Naradhipa, A. Purwarianti, "Sentiment classification for Indonesian message in social media," in *2012 International Conference on Cloud Computing and Social Networking (ICCCSN)*, 1–5, IEEE, 2012.
- [9] B. Pang, L. Lee, S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," *arXiv preprint cs/0205070*, 2002.
- [10] A. Reyes, P. Rosso, D. Buscaldi, "From humor recognition to irony detection: The figurative language of social media," *Data & Knowledge Engineering*, 74, 1–12, 2012.
- [11] F. Barbieri, H. Saggion, "Automatic Detection of Irony and Humour in Twitter," in *ICCC*, 155–162, 2014.
- [12] D. Bamman, N. Smith, "Contextualized sarcasm detection on twitter," in *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9, 2015.
- [13] K. Parmar, N. Limbasiya, M. Dhamecha, "Feature based composite approach for sarcasm detection using MapReduce," in *2018 second international conference on computing methodologies and communication (ICCMC)*, 587–591, IEEE, 2018.
- [14] Y. Ren, D. Ji, H. Ren, "Context-augmented convolutional neural networks for twitter sarcasm detection," *Neurocomputing*, 308, 1–7, 2018.
- [15] A. G. Prasad, S. Sanjana, S. M. Bhat, B. Harish, "Sentiment analysis for sarcasm detection on streaming short text data," in *2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA)*, 1–5, IEEE, 2017.
- [16] P. Tungthamthiti, K. Shirai, M. Mohd, "Recognition of sarcasm in microblogging based on sentiment analysis and coherence identification," *Information and Media Technologies*, 12, 80–102, 2017.
- [17] K. Sundararajan, A. Palanisamy, "Multi-rule based ensemble feature selection model for sarcasm type detection in twitter," *Computational intelligence and neuroscience*, 2020.
- [18] S. Khotijah, J. Tirtawangsa, A. A. Suryani, "Using lstm for context based approach of sarcasm detection in twitter," in *Proceedings of the 11th International Conference on Advances in Information Technology*, 1–7, 2020.
- [19] S. M. Sarsam, "Reinforcing the decision-making process in chemometrics: Feature selection and algorithm optimization," in *Proceedings of the 2019 8th international conference on software and computer applications*, 11–16, 2019.
- [20] Hilbe, Joseph M., *Logistic Regression Models*. CRC Press, 2009.

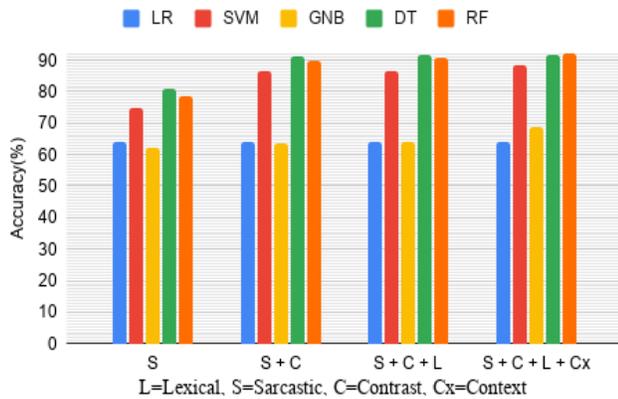


Fig. 4. Accuracy for the Incrementally Added Category. Decision Tree and Random Forest Classifiers show Better Accuracy along with the Sarcastic Feature Set.

Now turning to the experiment on different feature set combinations. We analyzed four distinct combinations to see which combination outperforms others. We combined the sarcastic feature set with lexical, contrast and context-based feature-set separately and contrast and context together for this experiment. We found higher accuracy (around 90%) for the sarcastic and contrast-based feature set. It appears that adding more features does not increase the model performance and feature selection is the central part of the efficient classification.

Lastly, we performed tests to observe which classifier outperforms if we increase the features, considering the sarcastic feature set as the base feature set. According to our observation, all feature sets' accuracy is not far more significant than other feature combinations. Hence, it is not always practical to add more feature sets to get better accuracy. However, the Decision Tree presents overall higher accuracy in all cases.

V. CONCLUSION AND FUTURE WORK

In this paper, we have suggested an improved model for detecting sarcasm in sentiment analysis. According to results, sarcastic features are more dominating than other features in sarcasm detection in tweets. Results show higher accuracy with sarcastic-based features for all classifiers we have used in our study. Moreover, the Decision Tree presents the highest accuracy (around 81%) with sarcastic-based features. Combining contrast-based features with sarcastic features increases the accuracy at approximately 90% for the Decision Tree. Therefore, it seems that it is not enough to add more features to obtain high accuracy. Though, the selection of the suitable feature set is the central part of the effective classification. Finally, we have evaluated all classifiers with incrementally added features, and findings reveal overall higher accuracy for the Decision Tree. We will study how to apply our proposed approach to improve sentiment analysis and opinion mining performances in future work. Additionally, we are also interested in context and pattern-based sarcasm detection in sentiment analysis.

Evaluation of Agent-Network Environment Mapping on Open-AI Gym for Q-Routing Algorithm

Varshini Vidyadhar¹

Research Scholar

Department of Computer Science & Engineering
Bangalore Institute of Technology, Bengaluru, India

Dr. R. Nagaraja

Professor

Department of Information Science and Engineering
Bangalore Institute of Technology, Bengaluru, India

Abstract—The changes in network dynamics demands a routing algorithm that adapts intelligently with the changing requirements and parameters. In this regard, an efficient routing mechanism plays an essential role in supporting such requirements of dynamic and QoS-aware network services. This paper has introduced a self-learning intelligent approach to route selection in the network. A Q-Routing approach is designed based on a reinforcement learning algorithm to provide reliable and stable packet transmission for different network services with minimal delay and low routing overhead. The novelty of the proposed work is that a new customized environment for the network, namely Net-AI-Gym, has been integrated into Open-AI Gym. Besides, the proposed Q-routing with Net-AI-Gym offers optimization in exploring the path to support multi-QoS aware services in the different networking applications. The performance assessment of the NET-AI Gym is carried out with less, medium, and a high number of nodes. Also, the results of the proposed system are compared with the existing rule-based method. The study outcome shows the Net-AI-Gym's potential that effectively supports the varied scale of nodes in the network. Apart from this, the proposed Q-routing approach outperforms the rule-based routing technique regarding episodes vs. Rewards and path length.

Keywords—Reinforcement learning; environment; agent; network; Net-AI-Gym; Q-routing; rule-based routing

I. INTRODUCTION

The collaboration of entities either in the form of computing devices or the people or be it any things through some specific form of connectivity and set of communication protocols forms a network [1]. Examples of networks may include computer networks [2], social networks [3], the network of things as the Internet of Things (IoT) [4]. The adoption of machine learning is a requirement to bring automation in the process of routing.

A. Machine Learning Models

The machine learning models learn to perform a specified task(T). The machine learning models (MLM) are broadly classified into three categories as i) Supervised learning Model (SLM), ii) Unsupervised Learning Model (USLM), and iii) Reinforcement Learning (RL) model, as shown in the Fig. 1. The selection of the MLM depends upon the type of task(T) to be performed by the machine. Whereas the learning experiences (LE) in different MLM comes from the different sources of the data. In the SML, the 'LE' comes from the input

and output mapping of the data. The USML gains the 'LE' from the pattern of the data.

B. Reinforcement Learning

The RL Model is a goal-oriented ML approach, where the 'LE' for performing a 'T' comes by interacting with the uncertain and dynamic environment. The RL enables the computer to make a sequence of such decisions that ensure to maximize their cumulative rewards (CR) automatically even if the computer is not explicitly programmed to complete the 'T.' Fig. 2 illustrates the architectural diagram of the typical RL context.

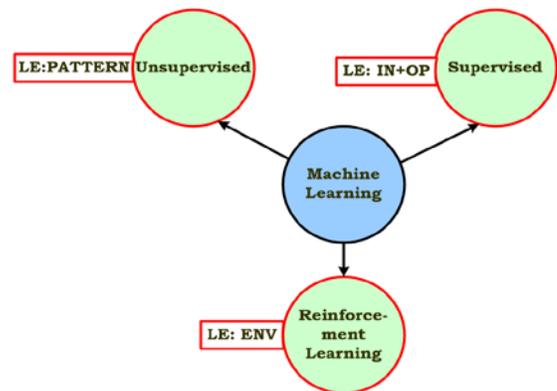


Fig. 1. Classification of Machine Learning Models

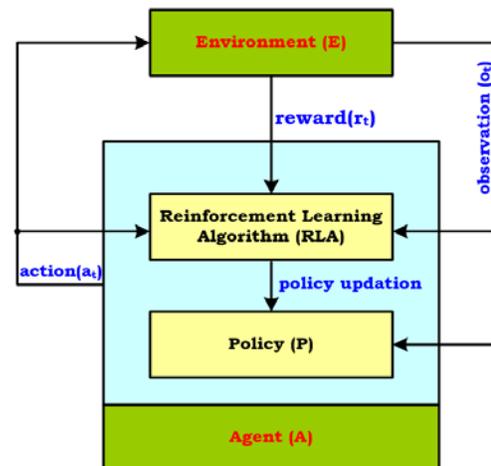


Fig. 2. Architectural Diagram of Typical RL Context.

The 'E' in RL is partially observable Markov decision process (POMDP). The mathematical modeling of any problem as an RL model takes a set of attributes (Sa) as a core building component for building the intuition of $\forall E \in S$. The attribute set $S_a = \{AE-R, MP, MP-MC, MRP, MDP, BE\}$, where, AE-R= agent and environment relationship, MP=Markov process, MP-MC= Markov process and Markov chain, MRP= Markov rewards process, MDP = 'Markov decision process, and finally, BE=Bellman equation.

C. Agent-Environment Relation (AE-R)

Basically, in RL Model, the 'A' is a software component that takes intelligent decisions based on the learning by an iterative cycle of gaining reward and penalty. The 'E' mimics the representation of the problem through a simulated environment to which the 'A' interacts. Another construct, namely 'state' (S), is the 'A' at a definite time step in the 'E.' The process that cannot be changed arbitrarily or randomly is a part of the 'E.' In simplification, action is any decision which the model wants it to learn, and the State is useful in choosing action. The 'A' cannot change the rewards arbitrarily. However, while designing the 'A,' it is assumed that as a part of its functions of taking action under the particular State, the agent knows how the computation of reward takes place in the environment. In some context, though the 'A' may have complete exposure or awareness of the 'E' yet, the 'A' finds it hard to maximize the 'R.' Therefore, the AE-R is the representation of the boundary or the limit of the 'A' control, not the knowledge of the 'A.'

D. Markov Property (MP)

The two intrinsic properties that define the Markov property are i) Transition (T): the process of changing one State to another state, ii) Transition probability (TP): The probability by which the agent can move from one State to another. Therefore, the Markov property states that "Future is independent of the past if present is given." The MP is expressed as in (1).

$$P[S_{t+1} | S_t] = P[S_{t+1} | S_1, \dots, S_t] \quad (1)$$

Where, $S[t]$ =current state of 'A' and $S[t+1]$ = next state. The intuitively meaning is that the current state includes information of the past states. Whereas, the state transition probability (STP) is expressed as in (2).

$$STP = P_{s \rightarrow s'} = P[S_{t+1} = s' | S_t = s] \quad (2)$$

The STP formulates a STP matrix (STPM) where, \forall Row \in STPM, represents the probability of moving next state.

$$\begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}$$

The sum of each row \in STPM, $\sum R = 1$.

E. Markov Process and Markov Chain (MP-MC)

MP is a memoryless random process such that the sequence of the random states $\{S[1], S[2], \dots, S[n]\}$ with a Markov property, so the environment, $E: \{State(S), STPM\}$.

F. Markov Reward Process (MRP)

To understand the MRP, it is essential to understand the rewards concept and the different nature of the task. The agents receive a +ve or a -ve numerical value on acting as some state(S) in the 'E,' whereas the sum of such rewards is called return(G) which is expressed in the equation below as in (3):

$$G_t = \sum_{i=t+1}^T r_i \quad (3)$$

Another essential aspect is the type of task. It is either episodic or continuous. In the episodic task, every time the process initiates from the start-state (S_s) and ends at the terminating-state (S_t), and once it researches the S_t , it is said to be completing one episode (E_t). Again, the process restarts from the S_s to S_t , where every E_t is independent of others. Once the agent research the destination node from the source node in the routing case, it completes one E_t . In contrast to the episodic task, the continuous task (C_t) does not have any terminating state (S_t). Therefore, the return (G) can be easily calculated, whereas the value of 'G' in C_t yields infinity. Thus, to resolve the problem of infinity return (G), the concept of discount

factor (γ) comes into the picture.

The factor ' γ ' basically regulates the immediate reward (R_i) and the future rewards (R_f) and the range of ' γ ': $\{0,1\}$ to avoid R^∞ in the C_t . The value '0' of ' γ ' indicates higher importance to the R_i , whereas the value '1' indicates higher importance. Therefore $\gamma = 0$, in practice no learning condition and $\gamma = 1$, will keep hunting infinitely the R_f , so for all practical purpose ' γ ' is taken as: $0.2 \leq \gamma \leq 0.8$. Therefore, the equation (3) with discount factor(γ) is normalized as in equation (4).

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (4)$$

The Agent (A) operates in the environment and gets rewards for each operation, and the prime goal of the 'A' is to maximize the cumulative reward by taking the most suitable action for the current observation. Therefore, the success rate of the 'A' is decided based on its reward, which shows how well the action being taken. In the existing literature, various research works have been conducted based on RL to solve network-related problems. However, none of the existing research works have introduced a suitable networking environment to evaluate the RL agent. In this paper, the evaluation of a customized environment, namely Net-AI Gym, is carried out with an RL agent algorithm designed based on the Q-Learning approach for network routing. The proposed study also considers a rule-based algorithm to be evaluated in the Net-AI Gym environment to carry out performance assessment in terms of reward VS episode and path length. The remaining sections of this paper are organized as follows: Section II provides a brief discussion on the types of networks and their characteristics. Section III presents the formulation of routing problems in the network. Section IV presents related work for analyzing the existing literature regarding the application of RL to network routing. Section V discusses the Environment Setup and processes involved. Section VI discusses the performance analysis and model validation, and finally, the overall contribution of the proposed work is concluded in Section VII.

II. TYPES OF NETWORK AND CHARACTERISTICS

Networks have become progressively ubiquitous. The network is a system of interconnected devices intended to share digital information. According to the characteristics, communication networks are usually divided into different categories: wired or wireless, energy limitation, network topology, and node mobility. These characteristics have a significant influence on RL-based routing optimization. Fig. 3 highlights the different types of networks based on wired and wireless network types.

In general, the typical characteristics of the network that usually affect the protocol design are based on several factors such as network is infrastructure-based or no infrastructure-based, centralized or distributed, node mobility, variation in network topology, node's energy consumption, link quality, bandwidth, the accuracy of data transmission, and application-specific requirements such as, time sensitivity and reliability-based requirements. The RL-based routing protocol has been gradually extended and improved as the network develops. Thus, 1 RL-based routing protocols have the capability of addressing various network issues.

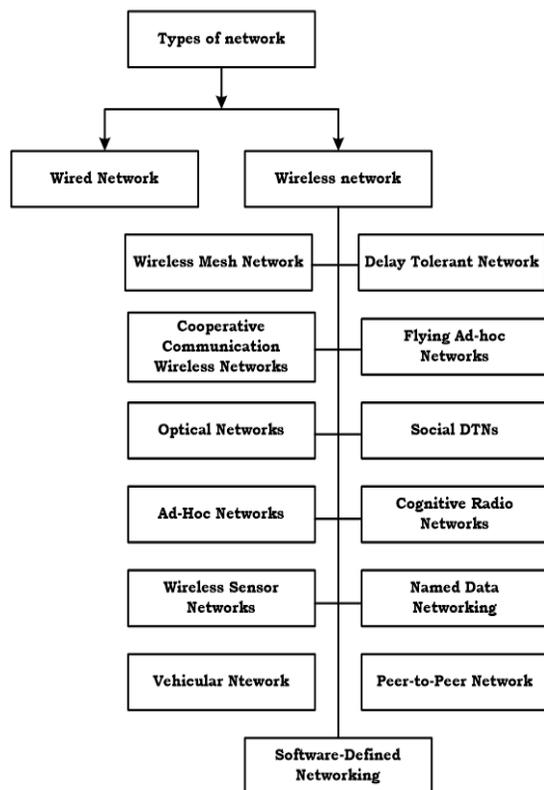


Fig. 3. Types of Networks.

III. FORMULATION OF ROUTING PROBLEM

The typical network 'NW' is a set of nodes(N), whereas a typical graph(G) is a set of vertices(V). In the case of NW, the 'N' is connected by links(L), whereas in the case of G, 'V' is connected by edges(E). Therefore, the mapping of a set: Network (NW) = {Node(N), Link(L)} → Graph(G) = {Vertex(V), Edge(E)}, so NW (N, L) is mapped to G (V, E). Whereas the Route(R) in an NW and a Path (P) in G is defined

as a way to move from the source node (Ns) to the destination node (Nd) in the NW and origin vertex (Vi) to another vertex (Vj), so {R (NW): Ns→Nd} → {P(G): Vi→Vj}. In the Network (NW), the link(L) is either one way or two ways, i.e., either Vi→Vj or Vj→Vi. Fig. 4 illustrates a real-world network as a mathematical model of an NW.

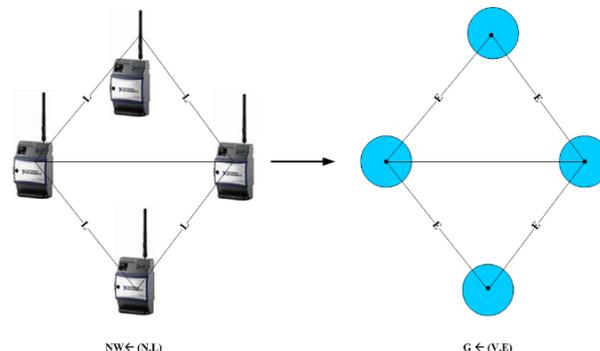


Fig. 4. Mapping Real-world Network (NW) to a Mathematical Model as Graph(G) of Machine Learning Models.

Therefore, in a nutshell, the network routing problem (NRP) is to obtain an optimal path (Po) between Vs. and Vd through intermediate nodes (Vk) under the constraints of weight (Wij) assigned to the link between Vi and Vj. The numerical value of Wij depends upon a set of parameters (Ps) with explicit and implicit relation between various properties of the chosen network such that Ps={Time(T), Cost (C), distance(D), Energy(E), Bandwidth (BW), Signal Strength (SS)}. Generally, in any network delay is induced or experienced due to the traffic condition of congestion, along with it if the nodes and links fail, the network topology dynamically changes; therefore, finding Po in such changing dynamics is a computationally expensive task because the complexity O(N²) of computation is exponential with the number of nodes(N/V).

IV. REVIEW OF LITERATURE

This section discusses the existing literature regarding the application of RL to network routing. A recent work carried out by Guo et al. [5] developed a quality-of-service (QoS) aware secure routing technique in software-defined Internet of Things (SD-IOT). This scheme exploits SDN's exclusive characteristics and considers the routing path planning under malevolent attacks to achieve QoS. A few research literature on underwater sensor network (UWSN) specific RL-based routing congestion control techniques [6]. In this approach, the next forwarder nodes are selected based on the current buffer state, location, and remaining energy of the one-hop neighbor nodes. The adoption of RL based routing scheme is considered for magnetic induction communication in UWSN [7]. In this study, the authors have derived the iterative formula of the Q-table by considering distance and energy path metrics. The RL-based routing scheme to the Mobile Ad-hoc Network (MANET) is considered by [8]. A routing algorithm is designed to select the next hop for packet forwarding based on the joint mechanism of stochastic approximation, function approximation, and RL. The authors in the study of [9] used RL's Q-learning approach to design an efficient and enhanced gradient-oriented routing strategy for balancing energy

consumption and QoS in Wireless Sensor Network (WSN). RL-based opportunistic routing scheme is suggested by [10] to support high-dimension data streaming in the application of multi-hop wireless networks. The application of Q-learning for Unstructured Peer-to-Peer Networks is considered by [11]. A load-balancing intelligent routing is devised to improvise query search efficiency for both intragroup and intergroup peers and reduce loads of queries under higher churns and heavy network workloads. A concept of a multi-agent RL scheme is introduced by [12] to devise an optimal routing scheme for the underwater optical WSN. In this approach, nodes' link quality and remnant energy are considered to develop an efficient routing algorithm suitable to dynamic communication environment and prolong network service duration. Also, Q-value initialization and the variant learning rate are devised to boost the routing algorithm's convergence. Some existing studies have implemented an RL-based routing scheme in Cognitive Radio Networks (CRNs). The authors in [13] examined the routing issue in energy harvesting in a multi-hop CRN communication scenario. RL-based route selection mechanism is developed considering the various factors affecting routings, such as the number of hops, the node's distance, energy consumption during the communication process, and remnant energy. Various routing schemes have been introduced in the context of unmanned aerial vehicle (UAV) communications in complex network environments [14-16]. However, such schemes are associated with bottleneck issues. In this regard, the use of RL for routing algorithms in UAV applications has gained wide attention. The researchers in the study of [17] suggested a Q-learning-based load balancing routing technique to handle relay traffic in UAV communication. The presented technique estimates network load through the queue status obtained from the ground-vehicular nodes. A reward function control is also implemented for quick learning feedback of the reward values under a dynamic communication environment. In [18], a global routing scheme using where each mobile node participates in the route discovery process. The presented global routing protocol is compared with the local routing protocol, where each intermediate node performs routing then based on its energy profile. Both these approaches are designed based on the Q-learning approach of RL. The concept of RL is adopted in routing optimization at the network level for minimizing interferences and delay in channel switching [19]. The routing decision is carried out based on past events and predicted routing decisions of primary users. An RL agent is devised into the cross-layer approach towards assisting the transfer of channel information to the network layer. In [20], adopted deep RL mechanism to address sampling problems and optimal route selection in the highly complex and dynamic network communication scenario. A hierarchical routing scheme based on Q-learning is presented by [21] to enhance message delivery rate performance with less delay and hops in Vehicular Ad-hoc networks (VANET). Here, a network region is divided into different grids, and the presented routing scheme discovers the next optimal grid towards the end-point or target point. It also finds a vehicle moving towards the next optimal grid for communicating data transmission. The authors in [22] presented a collaborative RL model to design a tree-based routing scheme that captures the network's dynamic

characteristics, such as several nodes and uncertain traffic, to provide QoS-aware services in Cloud Content Delivery Networks. In [23], the authors have used RL for network traffic engineering. The presented techniques learn to select critical traffic flows in the matrix and reconstruct optimal routes based on flow information in the matrix to balance link utilization in the network. In [24], RL-based intelligent routing is developed to provide a complete view of the network and fast data forwarding process in SDN-enabled networks. In [25], Cluster-oriented cooperative Scheduling scheme using RL to improve vehicular networks' communication efficiency and reliability.

V. OPEN-AI: GYM AND NETWORK ENVIRONMENT SETUP

With the appropriate function approximation, the methods like Q-Learning and policy gradients may provide better performance even under challenging environments. The popular benchmark for RL includes 1) Arcade Learning Environment (ALE), and 2) RL Lab (RLL).

Recently, Open-AI-Gym is the most popular benchmark with the following essentials into it:

- Combines best elements of ALE and RLL and having a diverse collection of tasks (Environment) with a standard interface.
- OpenAI-Gym provides the episodic setting of RL, where the agent experiences are broken down into a series of episodes.
- The agents' initial State is randomly sampled from distribution in each episode. The interaction proceeds until the environment reaches a terminal state.
- The goal in episodic RL is to maximize the expectation of total reward per episode and achieve a high level of performance in a few episodes as possible.
- OpenAI does not include an agent class.

Process: Function of -Open-AI Gym

1. Sample environment state [return first observation]:

Ob0 = env.reset()

2. Agent chooses first action:

A0=agent.act(Ob0)

3. Environment returns observations:

Info =env.step(A0)

Ob1, Rew0, Done0

4. Reward and a Boolean flag indicates

- if the episode is complete

A1=agent.act(Ob1)

Env.step(A1)

Ob2, Rew1, Done1

[A99 =agent.act(O99)]

Info= env.step(A99)

Ob100, rew 99, Done99

5. Done 99 = = True →Terminal

Design assumptions in OpenAI-Gym include i) the Only environment but no agent, ii) Emphasis on the sample, not just final performance, iii) Encourage peer review, not competition, iv) Strict versioning of environment, v) Monitoring by default.

OpenAI-Gym is a collection of partially observable Markov decision processes (POMDP), and the current environment consists of i) Algorithms, ii) Atari, iii) Box2D, iv) Classic Control, iv) Mujoco, v) Robotics, and vi) Toy text. It does not consider any environment explicitly for a network; therefore, a custom environment for the network is created as SimpleNetwork with a different number of nodes.

VI. Q-ROUTING MODEL VALIDATION WITH RULE-BASED ROUTING

This section presents outcome analysis and performance assessment of the proposed Q-routing on the customized OpenAI Gym environment, namely, Net-AI Gym [26]. The development and design of the proposed system are carried out on the numerical computing platform, and scripting of the proposed technique is done in Python. Different case studies have been considered to evaluate the stability and consistency of both Net-AI gym and Q-routing scheme in the performance analysis.

A. Case:1 Network with 6-Nodes

In this scenario, a network with six nodes is being considered.

Table I presents test case one with a network scenario with six nodes.

TABLE I. NETWORK SCENARIO-1: 6-NODES

Observation space	Discrete (6)
Action space	Discrete (6)

In this scenario, a network with six nodes is being considered. The agent will be examined with various networks containing various nodes. Both rule-based algorithms, as well as Q learning algorithms will be tested. The networks are designed in such a way that they simulate the actual internet.

Fig. 5 shows the network deployment scenario. The network contains six nodes. Randomly two nodes are selected and made into source and destination. The Net-AI gym platform has a unique reward system where if the agent tries to transfer the packet from one node to another, the packet will be dropped if the nodes are not connected. Every transfer agent will get a negative reward. However, the agent will get a positive reward when the packet reaches its destination. So in a way, the reward shows the throughput of the system.

Fig. 6 shows how the throughput is increasing over some time. After many episodes, the throughput starts increasing. The graph clearly shows the increase in the throughput and hence the reward. The agent initially takes some random moves to explore the network.

In Fig. 7 shows an analysis of the proposed Q-routing scheme concerning Epsilon Vs; the episode with six nodes network. The epsilon represents the probability of the agent taking up random moves. As can see, as the episodes progress,

epsilon decays down. This means, in the initial episodes, the agent may take more random moves; the probability of it reduces as the episodes progress.

Fig. 8 shows an analysis of the proposed Q-routing scheme concerning Episodes Vs. path length. The path length of 0 represents that the packet is lost. The path length of 2 is optimal. If we can observe the epsilon plot and the path length plot together, path length reduces to optimal length when epsilon decay below 1% which means the agent stops exploring.

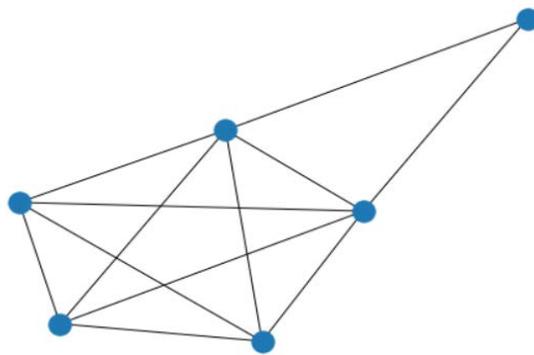


Fig. 5. An Environment with 6- Nodes Network.

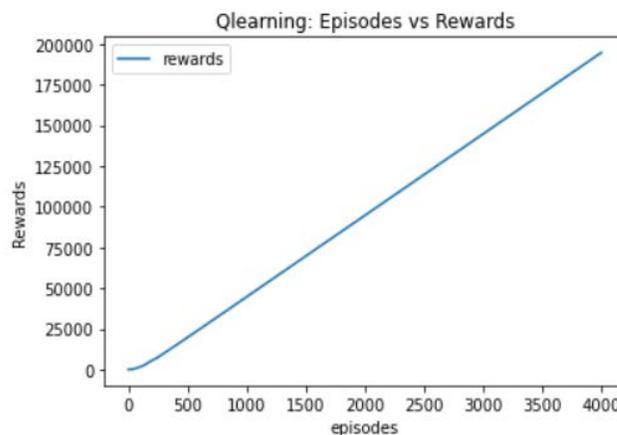


Fig. 6. Network with Node=6, Episodes vs. Rewards.

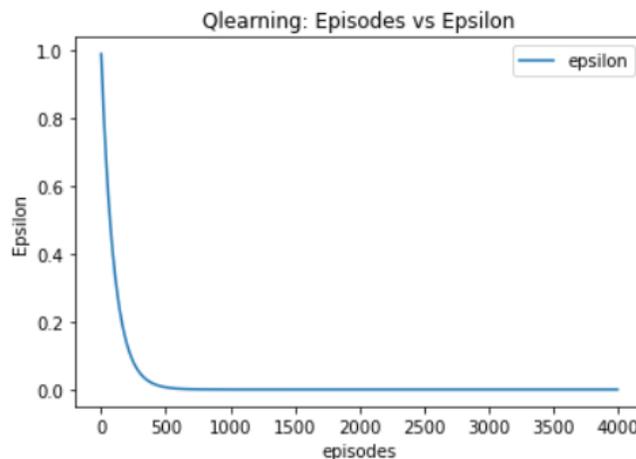


Fig. 7. Network with node=6, Epsilon Vs. Episode Define Episode.

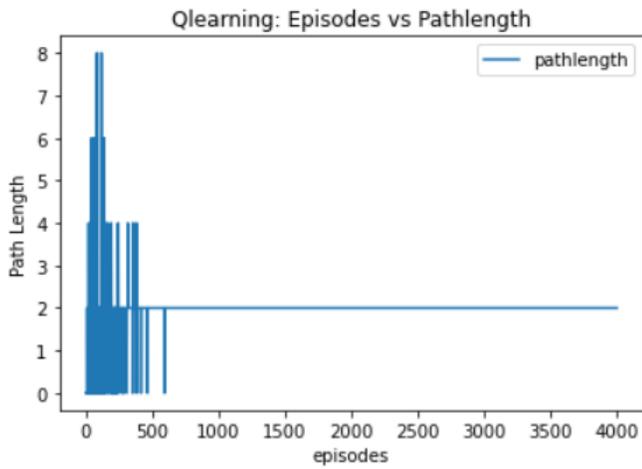


Fig. 8. Network with Node=5, Episode vs. Pathlength.

B. Case:2 Network with 50-Nodes

In this scenario, a network with fifty nodes is being considered. Table II presents test case one with a network scenario with six nodes.

TABLE II. NETWORK SCENARIO-2: 50-NODES

Observation space	Discrete (50)
Action space	Discrete (50)

However, when the network is with higher nodes, as shown in Fig. 9, the study comes across a different type of result. The probability of dropping is much more compared to a smaller network. As it can be observed, the agent will start delivering packets only after 1000 episodes.

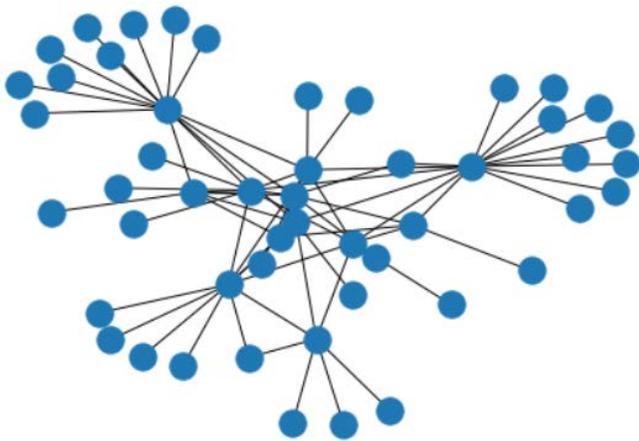


Fig. 9. An environment with 50- Nodes Network

This indicates that the more significant nodes, the more time the agent will take to learn the correct path. This is as expected.

As shown in Fig. 10, 11, and 12, even though the epsilon decay is much similar to that of the lower node network, if we observe the path length graph, the agent takes a very long time to find any let alone the longer path. This is because the network contains many gateway nodes. Due to which there will

be many nodes connected to a common node. This exactly happens in a real-world network scenario. The gateways are usually connected to a high number of other nodes and national links. Very rarely will any node be connected directly to a significant network. As it can be observed, the optimal path length here is 4. Even with 50 nodes, the optimal path length is only 4. This is true even in a real-world scenario.



Fig. 10. Network with node=50, Episode vs. Rewards.

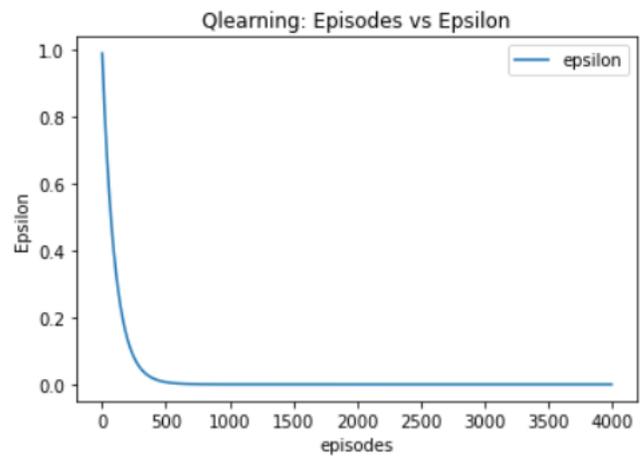


Fig. 11. Network with node=50, Episode vs. Epsilon.

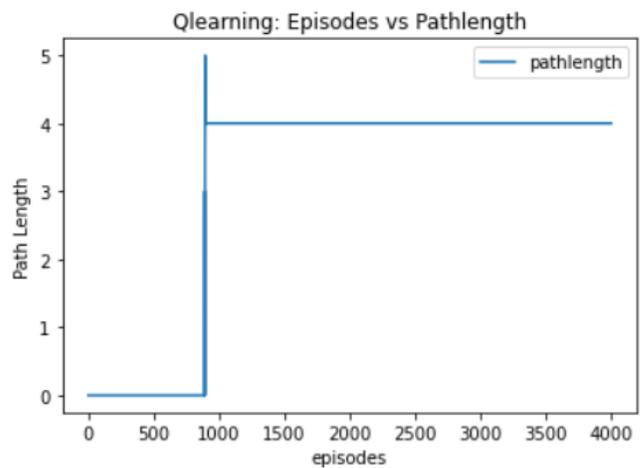


Fig. 12. Network with node=50, Episode vs. Pathlength.

C. Case:3 Network with 100-Nodes

In this scenario, a network with fifty nodes is being considered. Table III presents test case one with a network scenario with six nodes.

TABLE III. NETWORK SCENARIO-3: 100-NODES

Observation space	Discrete (100)
Action space	Discrete (100)

Fig. 13 shows a network environment with 100 nodes. The study has adopted the monte Carlo method here to evaluate the algorithm. A different case study has been considered to see how the algorithm performs with an increasing number of nodes. The scalability of the network routing algorithm is an essential aspect.

In Fig. 14, the 100-node network takes around 2000 episodes to learn the optimal path. As shown from graph trend, till episode 2000, there is no reward at all. Only after the algorithm learns the route, it starts to perform.

Fig. 15 shows an analysis of the proposed Q-routing scheme concerning Epsilon Vs; an episode with a 100 nodes network. The epsilon represents the probability of the agent taking up random moves. As can see as the episodes progress, Epsilon decay is the same as usual. The same rule works here as well.

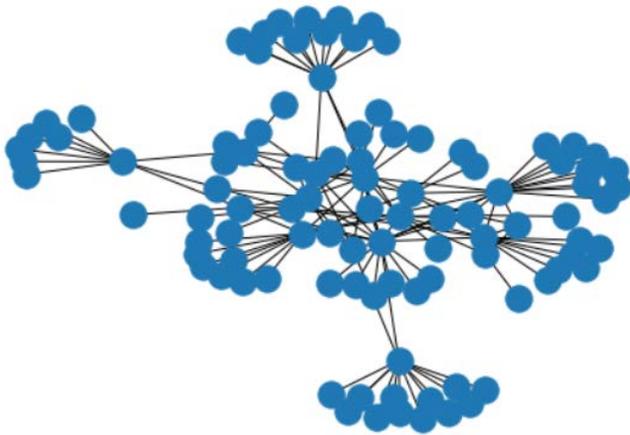


Fig. 13. An Environment with 100- Nodes Network.



Fig. 14. Network with node=100, Episode vs. Rewards.

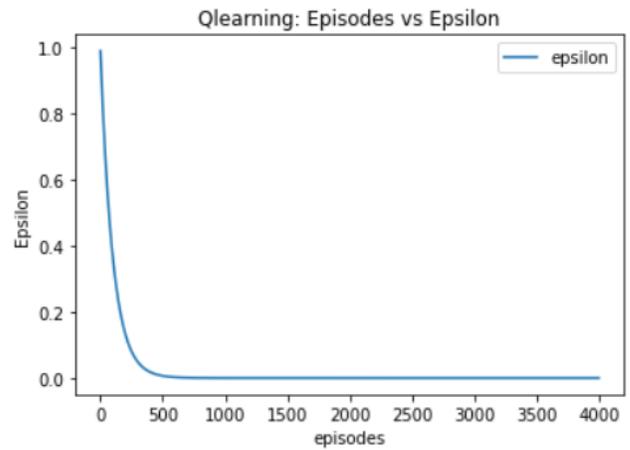


Fig. 15. Network with node=100, Episode vs. Epsilon.

Fig. 16 shows an analysis of the proposed Q-routing scheme concerning Episodes Vs. path length. As shown from the path length plot, the algorithm finds a suitable path after 2000 episodes. Even after epsilon decay fall below 1%, the algorithm has not found a path yet. Even if the algorithm does not explore, the algorithm finds a path in the network. Also, it can be analyzed based on the closer analysis that the optimal path length here is 3, and it has been earlier; there is no relationship between optimal path length and the number of nodes.

Fig. 17 demonstrates a comparative analysis for all three case studies of different network sizes. As it can be observed, once the algorithm learns, then the performance is the same on all the networks. However same is not true with rule-based methods.

As it can be seen in Fig. 18, the rule-based method never settles for a single path. This is due to the dynamic nature of the network. Connections and the weights keep changing. Hence, the path length also keeps varying.

Fig. 19 shows an analysis of Episodes vs. Pathlength for the proposed Q-learning scheme. In contrast to the path length plot of the rule-based method, the Q learning method always follows the optimal path after some time since it can predict the changes in the network before they can occur.

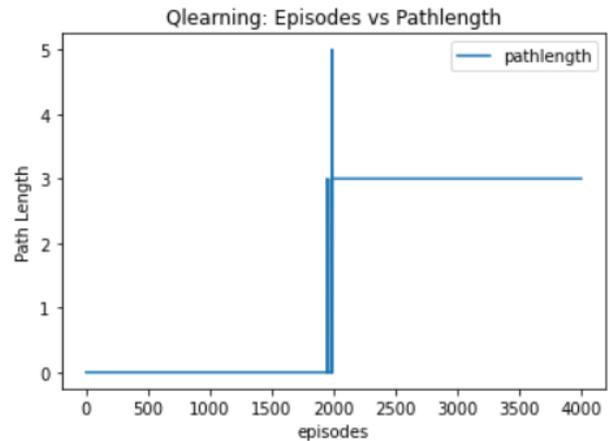


Fig. 16. Network with node=100, Episode vs. Pathlength.

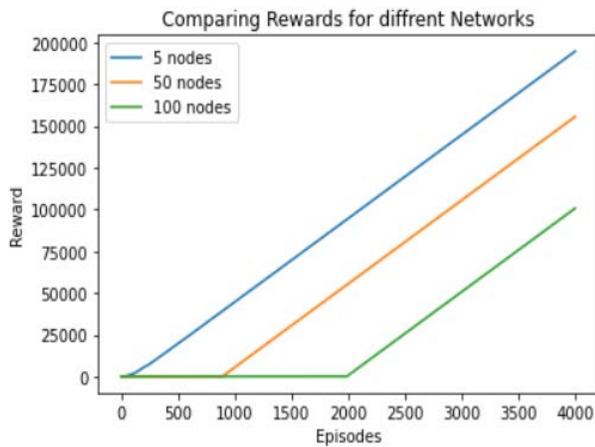


Fig. 17. Comparison Scenario-1, Two and, 3.

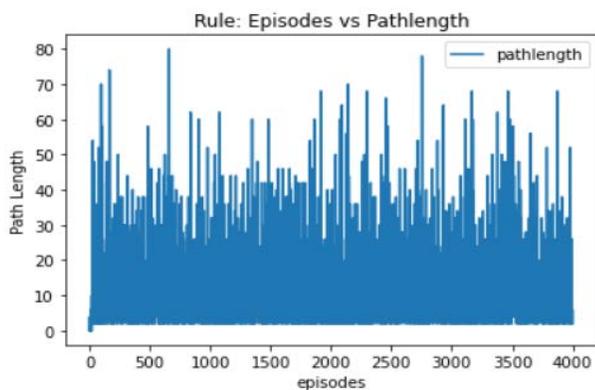


Fig. 18. Rule-Based: Episodes vs. Pathlength.

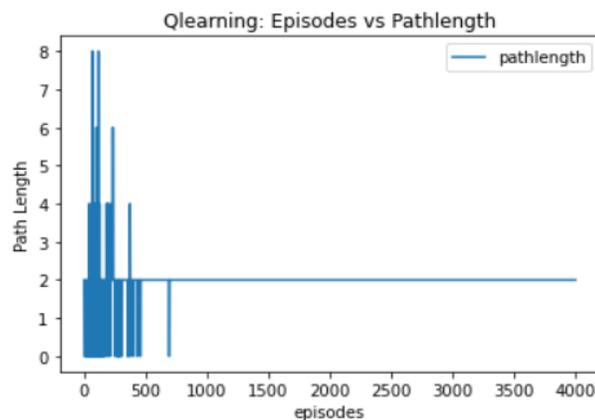


Fig. 19. Q-Learning Based: Episodes vs. Pathlength.

From the comparative analysis in Fig. 20, it can be analyzed that the rule-based method exhibited higher throughput. However, the Q-routing method is highly efficient once it learns the policy. Hence, in the long run, Q learning produces higher throughput compared to the rule-based method. Based on the overall analysis it can be observed that the proposed system offers a good scope for solving network related problem with RL agent algorithm, which can be well evaluated in the proposed customized Net-AI-Gym environment. The proposed system is found to be scalable to

both small network and large network as it perform well in all the three cases of network with different number of nodes. Based on the comparative analysis the proposed Q-learning algorithm outperform the rule-based algorithm in terms of episode vs. reward which shows the stability and efficiency of proposed system to address routing related problem.

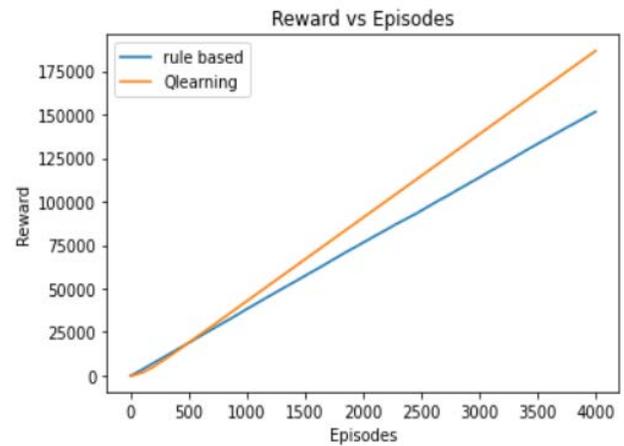


Fig. 20. Rule-Based vs. Q-Learning (Episodes vs. Reward).

VII. CONCLUSION

In this paper, the evaluation of the newly designed environment on the Open-AI Gym, namely Net-AI-Gym, is performed for agent algorithm based on the Q-learning is evaluated. The scalability limitation of network routing protocols using rule-based methods is being overcome with the RL-based Q-routing protocol. The performance validations between rule-based and RL-based Q-learning are carried out, and it is being found that the Q-learning performs better than rule-based routing protocols concerning episodes Vs. Rewards. The self-validation test for scalability for varying nodes provides a consistent result. However, the constraints of space optimization are planned as future research work. This is the first of its kind of work, custom-developed for routing protocol on the newly designed environment. In the future work, the proposed work can be extended towards improving the performance of the system and evaluation of RL algorithm on the customized Net-AI Gym environment with different technique dealing with a complex, high-dimensional state space.

REFERENCES

- [1] Liu, K.R., Sadek, A.K., Su, W. and Kwasinski, A., 2009. Cooperative communications and networking. Cambridge university press.
- [2] Balasubramaniam, Deepa. (2015). Computer Networking: A Survey. International Journal of Trend in Research and Development, 2.
- [3] Jaffali S., Jamoussi S., Khelifi N., Hamadou A.B. (2020) Survey on Social Networks Data Analysis. In: Rautaray S., Eichler G., Erfurth C., Fahrnberger G. (eds) Innovations for Community Services. I4CS 2020. Communications in Computer and Information Science, vol 1139. Springer, Cham.
- [4] Triantafyllou, Anna, Panagiotis Sarigiannidis, and Thomas D. Lagkas. "Network protocols, schemes, and mechanisms for the internet of things (iot): Features, open challenges, and trends." Wireless communications and mobile computing 2018 (2018).
- [5] X. Guo, H. Lin, Z. Li and M. Peng, "Deep-Reinforcement-Learning-Based QoS-Aware Secure Routing for SDN-IoT," in IEEE Internet of

- Things Journal, vol. 7, no. 7, pp. 6242-6251, July 2020, doi: 10.1109/IJOT.2019.2960033.
- [6] Z. Jin, Q. Zhao and Y. Su, "RCAR: A Reinforcement-Learning-Based Routing Protocol for Congestion-Avoided Underwater Acoustic Sensor Networks," *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10881-10891, 15 Nov.15, 2019, doi: 10.1109/JSEN.2019.2932126.
- [7] S. Wang and Y. Shin, "Efficient Routing Protocol Based on Reinforcement Learning for Magnetic Induction Underwater Sensor Networks," in *IEEE Access*, vol. 7, pp. 82027-82037, 2019, doi: 10.1109/ACCESS.2019.2923425.
- [8] P. Nurmi, "Reinforcement learning for routing in ad hoc networks," in *Proc. 5th Int. Symp. Modeling Optim. Mobile, Ad Hoc Wireless Netw., Limassol, Cyprus, 2007*, pp. 1–8.
- [9] B. Debowski, P. Spachos, and S. Areibi, "Q-learning enhanced gradient-based routing for balancing energy consumption in WSNs," in *Proc. 21st IEEE Int. Workshop Comput. Aided Modelling Design Commun. Links Netw. (CAMAD)*, Oct. 2016, pp. 18–23.
- [10] K. Tang, C. Li, H. Xiong, J. Zou, and P. Frossard, "Reinforcement learning-based opportunistic routing for live video streaming over multi-hop wireless networks," in *Proc. IEEE 19th Int. Workshop Multimedia Signal Process.*, Oct. 2017, pp. 1–6.
- [11] X. Shen, Q. Chang, L. Liu, J. Panneerselvam and Z. Zha, "CCLBR: Congestion Control-Based Load Balanced Routing in Unstructured P2P Systems," in *IEEE Systems Journal*, vol. 12, no. 1, pp. 802-813, March 2018, doi: 10.1109/ACCESS.2019.2912996.
- [12] X. Li, X. Hu, R. Zhang, and L. Yang, "Routing Protocol Design for Underwater Optical Wireless Sensor Networks: A Multiagent Reinforcement Learning Approach," in *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9805-9818, Oct. 2020, doi: 10.1109/IJOT.2020.2989924.
- [13] X. He, H. Jiang, Y. Song, C. He, and H. Xiao, "Routing Selection With Reinforcement Learning for Energy Harvesting Multi-Hop CRN," in *IEEE Access*, vol. 7, pp. 54435-54448, 2019, doi: 10.1109/ACCESS.2019.2912996.
- [14] Kashyap, A.; Ghose, D.; Menon, P.P.; Sujit, P.; Das, K. UAV aided dynamic routing of resources in a flood scenario. In *Proceedings of the 2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, Atlanta, GA, USA, 11–14 June 2019; pp. 328–335. 2. Zeng, F.;
- [15] Zhang, R.; Cheng, X.; Yang, L. UAV-assisted data dissemination scheduling in VANETs. In *Proceedings of the 2018 IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, 20–24 May 2018; pp. 1–6.
- [16] Yang, L.; Yao, H.; Wang, J.; Jiang, C.; Benslimane, A.; Liu, Y. Multi-UAV Enabled Load-Balance Mobile Edge Computing for IoT Networks. *IEEE Internet Things J.* 2020, 7, 1
- [17] Roh, B.S., Han, M.H., Ham, J.H. and Kim, K.I., 2020. Q-LBR: Q-Learning Based Load Balancing Routing for UAV-Assisted VANET. *Sensors*, 20(19), p.5685.
- [18] Mili R., Chikhi S. (2019) Reinforcement Learning Based Routing Protocols Analysis for Mobile Ad-Hoc Networks. In: Renault É., Mühlethaler P., Boumerdassi S. (eds) *Machine Learning for Networking. MLN 2018. Lecture Notes in Computer Science*, vol 11407. Springer, Cham.
- [19] Safdar Malik, T. and Hasan, M.H., 2020. Reinforcement Learning-Based Routing Protocol to Minimize Channel Switching and Interference for Cognitive Radio Networks. *Complexity*, 2020.
- [20] Y. Shao, A. Rezaee, S. C. Liew and V. W. S. Chan, "Significant Sampling for Shortest Path Routing: A Deep Reinforcement Learning Solution," in *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2234-2248, Oct. 2020.
- [21] F. Li, X. Song, H. Chen, X. Li and Y. Wang, "Hierarchical Routing for Vehicular Ad Hoc Networks via Reinforcement Learning," in *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1852-1865, Feb. 2019, doi: 10.1109/TVT.2018.2887282.
- [22] M. He, D. Lu, J. Tian, and G. Zhang, "Collaborative Reinforcement Learning Based Route Planning for Cloud Content Delivery Networks," in *IEEE Access*, vol. 9, pp. 30868-30880, 2021, doi: 10.1109/ACCESS.2021.3060440.
- [23] J. Zhang, M. Ye, Z. Guo, C. -Y. Yen and H. J. Chao, "CFR-RL: Traffic Engineering With Reinforcement Learning in SDN," in *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2249-2259, Oct. 2020, doi: 10.1109/JSAC.2020.3000371.
- [24] D. M. Casas-Velasco, O. M. C. Rendon and N. L. S. da Fonseca, "Intelligent Routing Based on Reinforcement Learning for Software-Defined Networking," in *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 870-881, March 2021, doi: 10.1109/TNSM.2020.3036911.
- [25] Xia, Y., Wu, L., Wang, Z., Zheng, X. and Jin, J., 2020. Cluster-Enabled Cooperative Scheduling Based on Reinforcement Learning for High-Mobility Vehicular Networks. *IEEE Transactions on Vehicular Technology*, 69(11), pp.12664-12678.
- [26] Varshini Vidyadhar, Nagaraj R, D.V. Ashoka, "NetAI-Gym: Customized Environment for Network to Evaluate Agent Algorithm using Reinforcement Learning in Open-AI Gym Platform," *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 12, No.4, pp. 169-176, 2021.

Internet of Things (IoT) based ECG System for Rural Health Care

Md. Obaidur Rahman¹

Department of CSE, DUET, EUB, Gabtoli, Dhaka, Bangladesh

Mohammad Abul Kashem²
Department of CSE, DUET
Gazipur, Dhaka, Bangladesh

Fazly Rabbi⁵
Department of Statistics
JNU, Savar, Dhaka, Bangladesh

Al-Akhir Nayan³
Department of CSE, EUB
Gabtoli, Dhaka
Bangladesh

Marzia Ahmed⁶
Department of Software Engineering
Daffodil International University
Dhaka, Bangladesh

Most. Fahmida Akter⁴
Department of CSE, BUP
Mirpur, Dhaka, Bangladesh

Mohammad Asaduzzaman⁷
Department of ME, DUET
Gazipur, Dhaka, Bangladesh

Abstract—Nearly 30% of the people in the rural areas of Bangladesh are below the poverty level. Moreover, due to the unavailability of modernized healthcare-related technology, nursing and diagnosis facilities are limited for rural people. Therefore, rural people are deprived of proper healthcare. In this perspective, modern technology can be facilitated to mitigate their health problems. ECG sensing tools are interfaced with the human chest, and requisite cardiovascular data is collected through an IoT device. These data are stored in the cloud incorporates with the MQTT and HTTP servers. An innovative IoT-based method for ECG monitoring systems on cardiovascular or heart patients has been suggested in this study. The ECG signal parameters P, Q, R, S, T are collected, pre-processed, and predicted to monitor the cardiovascular conditions for further health management. The machine learning algorithm is used to determine the significance of ECG signal parameters and error rate. The logistic regression model fitted the better agreements between the train and test data. The prediction has been performed to determine the variation of PQRST quality and its suitability in the ECG Monitoring System. Considering the values of quality parameters, satisfactory results are obtained. The proposed IoT-based ECG system reduces the health care cost and complexity of cardiovascular diseases in the future.

Keywords—Internet of things (IoT); electrocardiogram (ECG) monitoring system; ECG signal parameters; cardiovascular disease; logistic regression model

I. INTRODUCTION

IoT concept can be utilized in versatile areas such as intelligent health care system, intelligent agriculture, environmental impact predictions, automation industries, etc. [1, 2, 3, 4]. World Health Organization (WHO) [5] mentions

that the uncertainty of health conditions is a widespread problem for aged people. Aged people need to check their health conditions very frequently, especially for senior cardiovascular patients. Existing cardiovascular diagnosis systems need to be improved, including modern technology to detect the heart condition in a low cost, accurate and timely manner [6, 7]. Considering the heart-related issue, electrocardiogram (ECG) monitoring used extensively in rural hospitals and health research centers [8].

Rather than IoT, Cyber-Physical systems (CPS) can be considered as data-centric technology. CPS integrates innovative functionality processes that facilitate communication, computation, and control through IoT [9]. Moreover, it contributes to an advanced intelligence system that significantly affects social life [10]. CPS concept can be activated through Micro Electromechanical Systems for networking in monitoring, computing, and controlling the physical world.

The work concentrates on developing a portable heart monitoring system in the corporation of Electrocardiogram (ECG) technology [11]. Three heart rate detection sensors are utilized to make the device that generates analog data from ECG signals. Moreover, the analog data are converted into CSV format by Arduino microcontroller. The collected data is transmitted to the cloud through a local server, processed into P, Q, R, S, T ECG parameters, and analyzed by the machine learning algorithm. The ECG monitoring process detects ECG signals incorporating non-intrusive sensors, and the signal obtained from sensors transmits through the smartphone by wireless transmission methods, such as Zigbee or Bluetooth [12, 13]. Steady heart rate detection and an immediate heartbeat monitoring system are viable parameters for heart

disease. Experimentation indicated that cardiovascular disease could be treated, controlled, and prevented in a steady data monitoring process using ECG signals [14, 15].

A wearable ECG monitoring gadget-based system has been designed to associate with IoT and cloud service architecture for monitoring heart disease in this study. The sensors placed in the human chest records different ECG data through Arduino, and these data are transmitted to the IoT cloud without any delays. The HTTP and MQTT servers are put in the IoT cloud to provide users with quick and timely access to ECG data. The acquired data is stored in a non-relational database that can continuously improve data storage velocity and flexibility. In addition, a graphical user interface, accessible via the internet, is developed for the availability of cardiovascular-based data for medical experts to analyze the patient's heart conditions. The proposed IoT cloud-related systems in this work ensure the effectiveness, reliability, and accuracy of the data collected.

In addition to that, the measured heartbeat and ECG report of the patient obtained from this intelligent information system can be sent quickly through the text message, web server, and mobile application. The LIVE monitoring option of the webserver can be used by nurses and patient relatives for emergency cases. This process can be used for rural people at an affordable cost.

The features of the system are below:

- It will be a handy tool because it displays all the data and information collected solely from the internet. As a result, it minimizes the stress and pressure of the patient's relatives who work outside the home.
- Doctors can increase diagnostic accuracy by being connected to the health care system via IoT since they have all the relevant patient data at their fingertips. In a nutshell, it allows for continuous and remote patient monitoring.
- Doctors and family members may conduct their jobs without fear because they can track the patient's health status from anywhere. It also gives alerts anytime when a specific health parameter exceeds the optimal limit. Additionally, by getting a warning, doctors and family members can take the appropriate action and saves lives in the emergency.
- Rather than visiting or spending time at hospitals, most older people choose to stay at home with their loved ones. But, on the other hand, people suffer from a variety of diseases because of their stressful lifestyles and become very weak at old age. Furthermore, this project will benefit ICU patients.

II. MATERIALS AND METHODS

A. Proposed System

The following components are used to run the project:

- 1) Arduino MEGA ATmega2560
- 2) Sensors

- a) ECG AD8232
- b) Heartbeat sensor MAX30100
- 3) Wi-Fi module ESP8266 A1 Cloud Inside
- 4) Jumper wires
- 5) Breadboard
- 6) Laptop/ computer.

In the middle of the devices, an ECG monitoring system is used. First, the patient will press their finger against the heartbeat sensor, and the IR sensor's ray will count the beats from the blood flow. Then, the H-Beat button will be pushed and waited for 20 seconds after counting the beats from the blood flow. Finally, the outcome will be transmitted, and the heartbeat rate will be displayed on a laptop screen or a mobile screen utilizing a web page or a mobile app. This complete process is depicted in Fig. 1. Next, the ECG sensor will be affixed to the patient's chest, and the patient will press the 'ECG' button to activate the function. Therefore, the ECG graph, signal, or digits will be generated.

B. Electrical Components Control Unit

An Arduino Mega 2560 microcontroller board [16, 17] is utilized to complete the project. A 16 MHz crystal oscillator, 54 digital input/output pins, 4 UARTs, 16 analog inputs, a USB connection, an ICSP header, a power jack, and a reset button are included on the board. The ATmega2560 board is powered through a USB or an external power supply. The source of energy is automatically selected. Two power adapters are used. One is a 9v battery, and the other one is a 5v laptop USB. Heartbeats, ECG, and Wi-Fi sensors are getting 3.3v from those power sources. If more power is driven in the sensors, they may get damaged. Registers are used to minimize the source voltage and avoid damages.

C. Hardware and Software Implementation

Arduino is the system's central control unit. A heartbeat sensor, an ECG sensor, and various manual buttons are available on the input side. The Arduino com port displays the output. The Wi-Fi Module allows to transfer of data to the cloud, and once the data is uploaded, the results can be reviewed by signing in to the server using a computer or smartphone. The block diagram (Fig. 2) describes the complete hardware process.

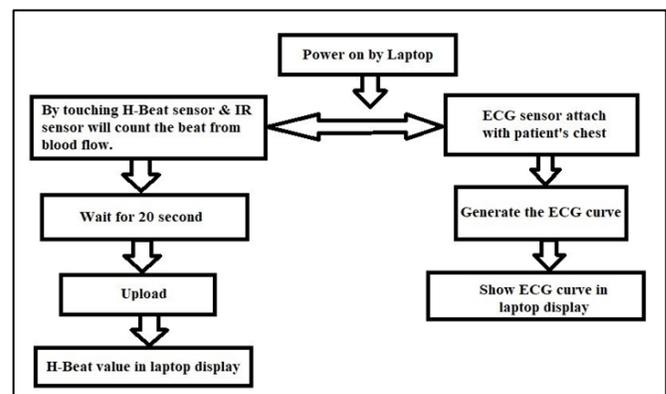


Fig. 1. Working Flow Diagram.



Fig. 2. Design of the IoT Device-based ECG System.

The software and hardware parts are merged in software implementation. Necessary code was written to command the hardware. In the coding section, the Arduino pin 2 (SDA and SDL) is used to initialize the cardiac bit sensor pin. Two other buttons are initialized in pin GND and 3.3v. ECG sensors LO+ and LO- are initialized in Arduino pin 10, 11. Additional controls are initialized in pin OUTPUT to A0, GND to GND, and 3.3v to 3.3v. The ECG and Heartbeat values are both zero at the beginning. The Wi-Fi cloud module contains an auto-configuration method. Wi-Fi modules RX and TX are initialized in Arduino pin 9 and 8, CH_EN and VCC are connected in Arduino pin 3.3v. Another pin named GPIO_0 is used for uploading the program to the Wi-Fi module.

In the beginning, combining two different sensors, programs are launched simultaneously and get the output result in the COM3 port of the Arduino. In the heartbeat approach, the hardware is employed to read the heartbeat. First, the heartbeat is counted for 20 seconds. At which point the heart bit pin raises high, the value is calculated. While taking data, the heartbeat is estimated for 60 seconds or 1 minute. Lastly, the heartbeat data is uploaded to the cloud. The command line for the ECG method is the same as the heartbeat method. If the ECG count is less than 50, then count the ECGs and send analog or digital data to the cloud, which is subsequently sent to the MQTT web server and displays the ECG result. The analog ECG data is uploaded and converted to digital data using the loop approach. ECG data is uploaded when it scores greater than 80, and for a lower score, an ERROR message is shown.

Wi-Fi module ESP8266 is connected to Arduino for uploading the MQTT program [18]. ESP8266 GPI 03 pin is connected to GND. After uploading the program, the RX, TX, and GPIO_0, the pin is unplugged. The RX and Tx pin of ESP8266 is connected to TX pin 8 and RX pin 9 of Arduino. After doing all these procedures, the project runs, and it transmits the sensors data to the cloud. The cloud visualizes the result on the MQTT box.

D. MQTT Box Site Implementation

REST has proven to be one of the most important advancements for Web applications [19]. REST stands for Representational State Transfer, and it is a technique of composing hypermedia applications. Structures for constructing RESTful Web administrations are currently

included in every significant development dialect. The hardware of the project is related to a program to implement and visualize the website. On the webpage, it shows the current value. A web front interface was used in the MQTT box. For this purpose, we downloaded the MQTT box and installed the software. When the heartbeat is measured, data is uploaded into the cloud server, and then it transmits the result to the MQTT box to show real value on the website according to date and time. The ECG result is shown using a web service that calls data from the server and generates it using Arduino code.

To design and test the MQTT communication protocol, a developer's helper application was used. MQTT Box for ECG Monitoring System boosts the MQTT process. Apps for Mac, Linux, and Windows are also available for MQTT Box. In the ISM database, the cyber-level serves as the significant information hub [20]. To deploy cyberspace, information is extracted from each source and compiled. With massive data, a specific analyzer is employed to extract data that gives a better sense of each patient's ECG data monitoring status.

III. EXPERIMENTAL ANALYSIS AND RESULTS

A. Dataset Configuration and Analysis

We created a dataset to train the Gradient Boosting Model (GBM) [21, 22, 23]. Approximately 2000 samples were collected from volunteers. In addition, related information like age, 'P', 'Q', 'R', 'S' and 'T' were also collected. Table 1 shows a tiny part of our dataset.

Exploring the relationship among different variables, some concise analyses were done to measure the mean, median, count, minimum, maximum, standard deviation, and quartile value of the dataset, which have been described in Table 2.

TABLE I. DATASET CONFIGURATION

Record No	Age	P	Q	R	S	T
1	21	91.6	100.0	100.0	100.0	90.0
2	23	100.0	100.0	100.0	100.0	100.0
3	30	100.0	100.0	100.0	100.0	100.0
4	25	100.0	100.0	100.0	100.0	100.0
5	20	78.5	100.0	80.0	100.0	100.0

TABLE II. DATASET MEASUREMENTS

	R. N	Age	P	Q	R	S	T
Count	20.00	20.00	20.00	20.00	20.00	20.00	20.00
Mean	10.50	29.85	97.06	96.26	95.26	96.51	97.66
Std	5.92	8.86	5.88	7.57	8.33	7.15	4.50
Min	1.00	18.00	78.50	76.19	76.19	76.19	85.00
25%	5.75	22.75	98.43	98.43	93.53	98.43	98.63
50%	10.50	28.50	100.00	100.00	100.00	100.00	100.00
75%	15.25	36.25	100.00	100.00	100.00	100.00	100.00
Max	20.00	45.00	100.00	100.00	100.00	100.00	100.00

B. Exploring Attributes type Cleaning the Dataset

The model used the parameters described in Table 3.

Drawing Box Plot, many attributes have been found containing outliers or extreme values (Fig. 3). All the observations along with outliers were deleted. After that, some robust statistics were utilized to calculate the different features of this dataset. However, the entire dataset was considered to estimate various measures. The dataset contained no null values and needed not to treat with any null values.

C. Dependent and Independent Variables

Quality is dependent on the target variable, and all other variables are independent variables. From these target values, a heart condition was measured (Table 4). There are six distinct quality values through analysis, and these are 14, 1, 1, 1, 2, 1. Grouping the quality, we found that most of the heart is in 14, 2 & 1, indicating the condition as good. We found two lousy quality hearts in our dataset, and three people were close to a tragic situation.

TABLE III. DATASET PARAMETERS

S.N.	Column	Non-Null Count	Data Type
1.	Record No	20 not_null	int64
2.	Age	20 not_null	int64
3.	P	20 not_null	float64
4.	Q	20 not_null	float64
5.	R	20 not_null	float64
6.	S	20 not_null	float64
7.	T	20 not_null	float64

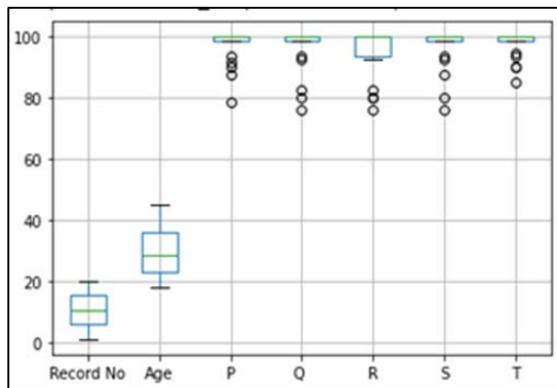


Fig. 3. BOX Plot of Dataset.

TABLE IV. QUALITY VALUES FOR GOOD OR BAD HEART CONDITION

Mean value	Quality value
76.13	1
80.00	2
82.35	1
92.85	1
93.75	1
100.00	14

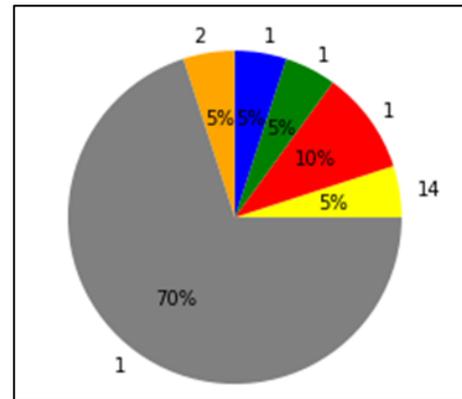


Fig. 4. Quality Wise Pie Chart.

Fig. 5. Investigation of Variables Correlation and Covariance

D. Dataset Visualization

A quality-wise Pie-Chart was drawn that describes the percentage ratio of the six qualities. In the chart, yellow indicates quality 14, red, green, blue, and orange indicate quality 1 and 2. According to the values mentioned in Fig. 4, it can be concluded that 70.00% sample's heart condition was excellent when collecting data.

Another measurement was run to investigate the correlation and covariance of different variables with the target variables. The outcomes identify variables responsible for various heart conditions. The correlation result is mentioned in Table 5 and covariance in Table 6.

After sorting the significant correlations, we found that 'Q' (ECG, Q - parameter), 'S' (S - parameter), 'T' (T - parameter) & 'P' (P - parameter) have some moderate correlation, comparing with other attributes. On the other hand, 'R' (R - parameter included angina) has a significant positive correlation with Q (Q - parameter included angina).

TABLE V. CORRELATION VARIABLES

	R. N	Age	P	Q	R	S	T
R. N.	1.0000	0.4431	0.1777	0.0657	0.1772	0.0408	0.2775
Age	0.4431	1.0000	0.4338	0.1623	0.2878	0.1383	0.1095
P	0.1777	0.4338	1.0000	-0.2135	0.2052	-0.2072	0.1713
Q	0.0657	0.1623	-0.2135	1.0000	0.8460	0.9893	0.4546
R	0.1772	0.2878	0.2052	0.8460	1.0000	0.8372	0.3474
S	0.0408	0.1383	-0.2072	0.9893	0.8372	1.0000	0.5011
T	0.2775	0.1095	0.1713	0.4546	0.3474	0.5011	1.0000

TABLE VI. COVARIANCE VARIABLE

	R. N	Age	P	Q	R	S	T
R. N.	35.000	23.236	6.1802	2.945	8.735	1.726	7.389
Age	3.236	78.555	22.594	10.888	21.256	8.760	4.371
P	6.180	22.594	34.533	-9.498	10.046	-8.703	4.530
Q	2.945	10.888	-9.498	57.285	53.345	53.515	15.485
R	8.735	21.256	10.046	53.345	69.405	49.846	13.027
S	1.726	8.7607	-8.703	53.515	49.846	51.072	16.118
T	7.389	4.3712	4.530	15.485	13.027	16.118	20.251

E. Comparison among Significant Variables and Target Variables

Comparing among significant variables and 'R' variables, the relationship between the variables can be virtualized. The correlation of the 'R' variable with other variables is shown in Table 7. 'Q' type makes an essential factor for having heart disease because most cases with 'R' I have chest pain. Two 'R' variables raise to level 120 – 150 among the age range 40 – 60. Fig. 5 represents the change of various parameters. From the findings, it can be concluded that people who have heart disease have maximum high blood pressure, and cholesterol is very high. Par's plot shows the distribution of single variables depicted in Fig. 6.

TABLE VII. CORRELATION WITH OTHER VARIABLES

R	1.000000
Q	0.846016
S	0.837237
T	0.347479
Age	0.287883
P	0.205213

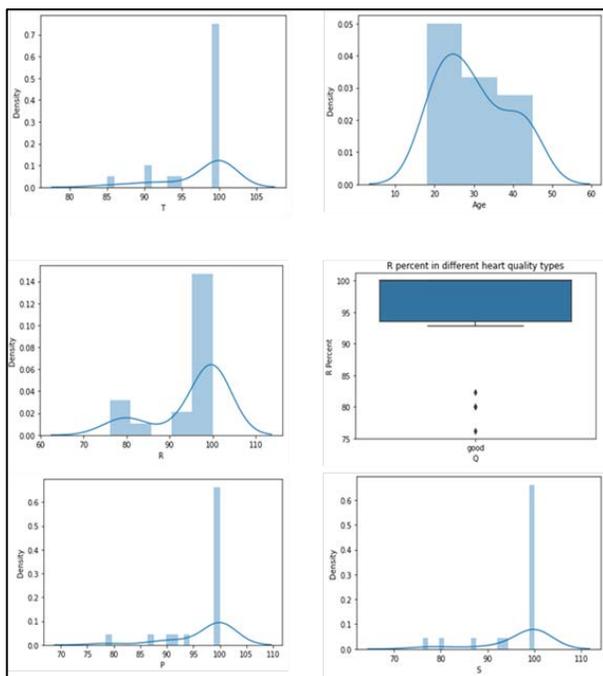


Fig. 6. Change of Various Parameters.

F. Heartbeat Result Analysis

The heartbeat result was analyzed by an automatic blood pressure system to see if the heartbeat sensor is functional or not (Fig. 7). For further processing, data were collected from five different persons of a specified age range. The information was listed along with a particular day and time.

G. ECG Report Analysis

Three electrodes are inserted on the patient's chest at first. The red electrode is implanted on the right side of the chest. The green electrode is situated on the left side of the chest, whereas the yellow electrode is located under the green electrode. Then the ECG push button is pressed. The value is converted into a curve and uploaded to the webserver and virtualized through mobile app and website. Arduino com port result is shown in Fig. 8. The key feature of the measured ECG is depicted in Fig. 9. ECG_PQRST data from the device is mentioned in Table 8.

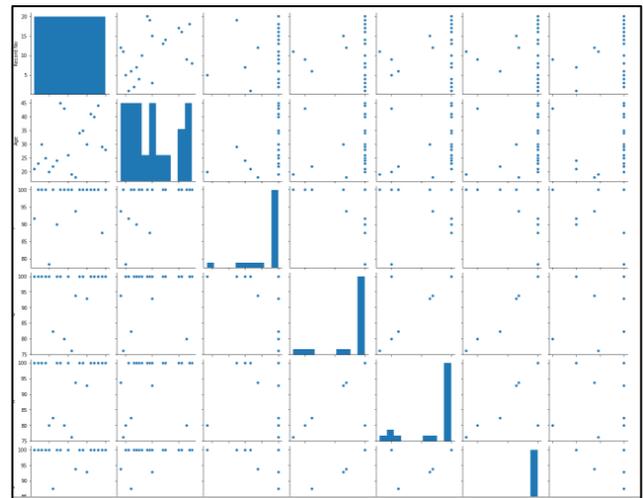


Fig. 7. Pair Plots among the Variables.



Fig. 8. Heartbeat Result Analysis.

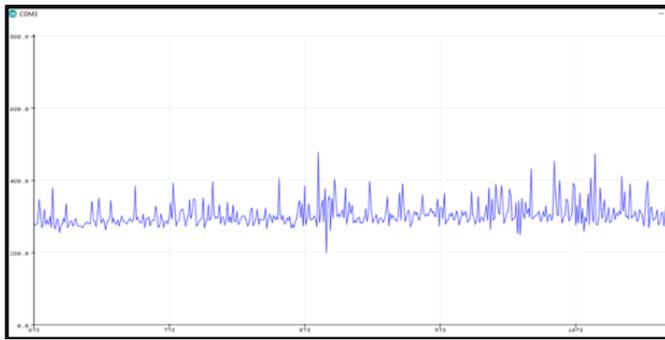


Fig. 9. Arduino com Port Result.

TABLE VIII. ECG_PQRST DATA FROM THE DEVICE

Record No	Age	P	Q	R	S	T
1	21	91.6	100	100	100	90
2	23	100	100	100	100	100
3	30	100	100	100	100	100
4	25	100	100	100	100	100
5	20	78.5	100	80	100	100
6	22	100	82.35	82.35	87.5	100
7	24	90	100	100	100	90
8	45	100	100	100	100	100
9	43	100	80	80	80	85
10	26	100	100	100	100	100
11	19	100	76.19	76.19	76.19	94.54
12	18	93.75	93.75	93.75	93.75	93.75
13	34	100	100	100	100	100
14	35	100	100	100	100	100
15	30	100	92.85	92.85	92.85	100
16	41	100	100	100	100	100
17	40	100	100	100	100	100
18	44	100	100	100	100	100
19	29	87.5	100	100	100	100
20	28	100	100	100	100	100

H. ML (Machine Learning) Intercept and Coefficients

A statistical tool known as linear regression predicts the future value of heart condition measurement parameters (P, Q, R, S, T) by analyzing the past. The technique uses the least square method to draw a straight line that minimizes the difference between current values and resulting values. The best fit line associated with the n points (S1, T1), (S2, T2), (Sn, Tn) has the form $y = mx + b$. Using the formula and we have got the intercept value which is 14.319164821638992.

From the analysis, the coefficient value for S is 0.962445, for T is -0.173491. For age, the value was 0.162937, where a positive sign implies that the response variable increases when the predictor variable rises. In contrast, a negative sign suggests that the response variable drops when the predictor variable increases.

I. Prediction

The score of one variable is predicted based on the scores of a second variable. The criterion variable is the variable that forecasts and is denoted as Y. The variable on which the predictions are based is the predictor variable and abbreviated as X. The project used only one predictor variable: Y. The projections of Y are evaluated when it is plotted as a function of X from a straight line. This complete process of prediction is listed in the Table 9.

In Fig. 10 indicates a positive association between X and Y. If Y is forecasting based on X, the greater the value of X may provide an accurate prediction of Y. The regression line in the figure is made up of the expected score on Y for each possible value of X.

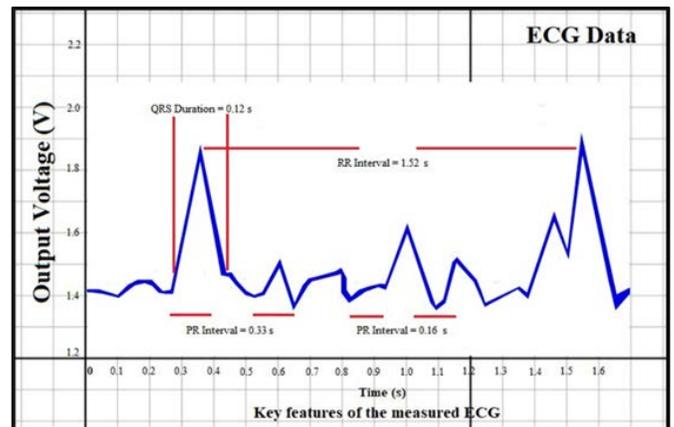


Fig. 10. The Key Feature of the Measured ECG.

TABLE IX. PREDICTION

S. N	Actual	Predicted
2	100	98.10267
5	82.35	84.76861
17	100	100.3838
19	100	97.7768
12	100	98.75442

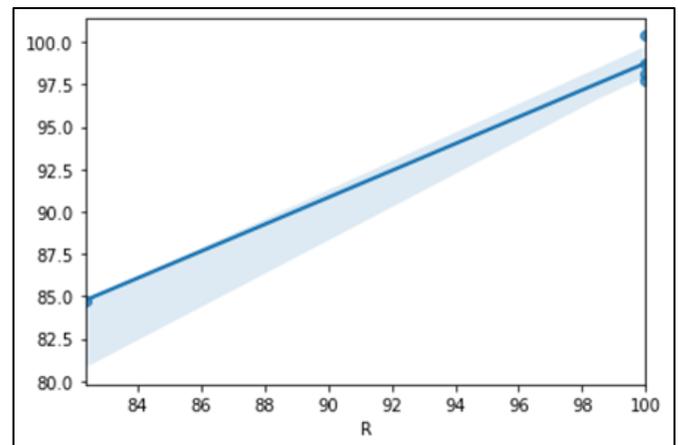


Fig. 11. Prediction Plot.

The error for a point is the difference between the actual value and the predicted value. From our analysis, the mean absolute error value is 1.633702533697678, and the mean squared error is 3.2181888486775514. Thus, the accuracy percentage is 93.5434261396096.

The bar diagram below (Fig. 12) shows the relationship between actual and predicted plots. For example, in Fig. 11, the blue indicates true, and the orange represents the expected value.



Fig. 12. Relationship between Actual and Predicted Data.

J. Cost Analysis

In developing countries, hospitals buy a lot of new medical equipment. Famous hospitals import medical equipment from other countries. As a result, hospital administrators pay a significant amount of money to bring the equipment to provide the best possible service to their patients.

The project described in this article is not costly. Every component is inexpensive and of good quality. As a result, it is affordable to everyone without causing financial hardship. The list of devices selected for ECG monitoring system purposes is listed in Table 10.

TABLE X. COST ANALYSIS

Components	Unit Price
Arduino Mega 2560	800/-
Heartbeat/Pulse sensor (MAX30100)	750/-
ECG sensor (AD8232)	1250/-
Esp8266 A1 cloud Wi-Fi module	180/-
CP2102 USB 2.0 to TTL UART Module	180/-
Male-Female Jumper Wires	100/-
400 Tie Points Breadboard White	100/-
Total	3360/-

IV. CONCLUSION

We have created and executed an ECG monitoring system that is entirely based on current IoT technologies. The IoT-based ECG monitoring system is constructed based on the proposed design. IoT-based healthcare platform links with smart sensors affixed to the human body for health monitoring. We talked about IoT-based patient monitoring systems in this article. Smartphones or gadgets use intelligent technologies, and we have discussed the advantages, disadvantages, and opportunities. Continuous remote monitoring is required for observing the medical patient. Our research work provides the ability to monitor patients via web app services and mobile massage services continuously. This research also contrasted the early medical system to modern health monitoring. The work will bring change in medical science and be a blessing for rural areas. The research work has proved its benefits already. We are planning for the further

development of the project by promising that one day every people of our country will get immediate medical treatment with the help of our project.

REFERENCES

- [1] E. Park, Y. Cho, J. Han, and S. J. Kwon, "Comprehensive Approaches to User Acceptance of Internet of Things in a Smart Home Environment," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2342–2350, 2017.
- [2] S. K. Vishwakarma, P. Upadhyaya, B. Kumari, and A. K. Mishra, "Smart Energy Efficient Home Automation System Using IoT," 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), 2019.
- [3] E. Husni, G. B. Hertantyo, D. W. Wicaksono, F. C. Hasibuan, A. U. Rahayu, and M. A. Triawan, "Applied Internet of Things (IoT): Car monitoring system using IBM BlueMix," 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA), 2016.
- [4] D. Balsamo and S. Das, "Health Monitoring Based on Internet of Things (IoT)," *Health Monitoring Systems*, pp. 99–120, 2019.
- [5] Ageing, In: World Health Organization. <http://www.who.int/topics/ageing/en/>, 2016.
- [6] A. Banerjee and S. K. Gupta, "Analysis of Smart Mobile Applications for Healthcare under Dynamic Context Changes," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 904–919, 2015.
- [7] Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, "Health-CPS: Healthcare Cyber-Physical System Assisted by Cloud and Big Data," *IEEE Systems Journal*, vol. 11, no. 1, pp. 88–95, 2017.
- [8] M. Bansal and B. Gandhi, "IoT & Big Data in Smart Healthcare (ECG Monitoring)," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019.
- [9] M. Bsoul, H. Minn, and L. Tamil, "Apnea MedAssist: Real-time Sleep Apnea Monitor Using Single-Lead ECG," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp. 416–427, 2011.
- [10] M. F. U. Hassan, D. Lai, and Y. Bu, "Characterization of Single Lead Continuous ECG Recording with Various Dry Electrodes," *Proceedings of the 2019 3rd International Conference on Computational Biology and Bioinformatics - ICCBB 19*, 2019.
- [11] L. Strachan, "Monitoring the Critically Ill Patient - Second edition Monitoring the Critically Ill Patient - Second edition," *Nursing Standard*, vol. 22, no. 17, pp. 30–30, 2008.
- [12] H. Kim, S. Kim, N. V. Helleputte, A. Artes, M. Konijnenburg, J. Huisken, C. V. Hoof, and R. F. Yazicioglu, "A Configurable and Low-Power Mixed Signal SoC for Portable ECG Monitoring Applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 8, no. 2, pp. 257–267, 2014.
- [13] C. H. Tseng, "Coordinator Traffic Diffusion for Data-Intensive Zigbee Transmission in Real-time Electrocardiography Monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3340–3346, 2013.
- [14] S. P. Preejith, R. Dhinesh, J. Joseph, and M. Sivaprakasam, "Wearable ECG platform for continuous cardiac monitoring," 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016.
- [15] R. Thanuja and R. Balakrishnan, "Real time sleep apnea monitor using ECG," 2013 Ieee Conference On Information And Communication Technologies, 2013.
- [16] S. Sarwito, I. R. Kusuma, and F. A. Cahyono, "Automatic Stacking Crane Prototype using Microcontroller Arduino Mega 2560," *International Journal of Marine Engineering Innovation and Research*, vol. 1, no. 1, 2016.
- [17] A. Ghosh, "Colour based Product Organization Tool using Arduino," *International Journal for Research in Applied Science and Engineering Technology*, 2019, 7(5), pp. 3547–3551, DOI: 10.22214/ijras.2019.5581.
- [18] A. Rashkovska, M. Depolli, I. Tomašič, V. Avbelj, and R. Trobec, "Medical-Grade ECG Sensor for Long-Term Monitoring," *Sensors*, vol. 20, no. 6, p. 1695, Mar. 2020, DOI: 10.3390/s20061695.

- [19] M. Melnichuk, Y. Kornienko and O. Boytsova, "WEB-SERVICE. RESTFUL ARCHITECTURE", Automation of technological and business processes, vol. 10, no. 1, 2018. Available: 10.15673/atbp.v10i1.876.
- [20] R Santos et al., "What is MQTT and How It Works | Random Nerd Tutorials", Random Nerd Tutorials, 2020. [Online]. Available: <https://randomnerdtutorials.com/what-is-mqtt-and-how-it-works/>. [Accessed: 11- Dec- 2020].
- [21] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Frontiers in Neurobotics*, vol. 7, 2013, DOI: 10.3389/fnbot.2013.00021.
- [22] A. A. Nayan, J. Saha, A. N. Mozumder, K. R. Mahmud, A. K. A. Azad, M. G. Kibria "Early Detection of Fish Diseases by Analyzing Water Quality Using Machine Learning Algorithm", *Walailak Journal of Science and Technology*, vol. 18, 2021.
- [23] A. A. Nayan, M. G. Kibria, M. O. Rahman and J. Saha, "River Water Quality Analysis and Prediction Using GBM," 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), 2020, pp. 219-224.

Design of a Plastic Shredding Machine to Obtain Small Plastic Waste

Witman Alvarado-Diaz¹, Jason Chicoma-Moreno², Brian Meneses-Claudio³, Luis Nuñez-Tapia⁴

Image Processing Research Laboratory (INTI-Lab)
Universidad de Ciencias y Humanidades
Lima, Perú

Abstract—One of the biggest environmental problems in the world is the excess of plastic waste. Although it is true, around the world there are companies that are responsible for recycling plastic waste, since 42% of plastics that are generated worldwide have a single utility. In Peru, the culture of recycling is not promoted, each year approximately each person uses 30 kilos of plastic and that only in Metropolitan Lima and Callao 46% of plastic waste is generated nationwide. In view of this problem, this article will design a plastic shredding machine to obtain small plastic waste to help people to be dedicated to the recycling industry in an automated way, it would also generate jobs since it requires a staff in charge of the machine and it will also be extremely useful to reduce plastic pollution in the environment, which due to COVID-19 is increasing. Through the design of the plastic waste shredding machine, the plastic will be selected by color and type of plastic composition, whether it is Polyethylene Terephthalate (PET), High Density Polyethylene (HDPE), Low Density Polyethylene (LDPE), Polychloride vinyl (PVC) or Others (Plastic Mix), then it will go through the shredding process to become small plastic waste, which could be turned into filament for 3D printer.

Keywords—Automation; pollution; filaments; recycling; plastic waste

I. INTRODUCTION

One million plastic bottles are bought every minute and 500 billion bags are used per year. Eight million tons end up in the oceans every year, threatening marine life [1]. The great problem of plastic is inherent to its usefulness: most of the products made of this material have a very short useful life and are usually easily disposed of [2], for this reason there are many companies that are in charge of recycling plastic to give them different applications. 42% of the plastic used in the world is destined to the packaging of food and manufactured products. In other words, single-use plastics that barely spend a few minutes in the hands of consumers [3]. If no action is taken, by 2050 there will be about 12 billion tons of plastic waste distributed in landfills and the ocean [4].

Plastic in Peru represents 10% of all the waste generated in the country, since 2015 we have had a huge growth in plastic [5]. On average in Peru, approximately 30 kilos of plastic are used per year per citizen [6]. Each year, there are about 3 billion plastic bags, almost 6 thousand bags per minute. In Metropolitan Lima and Callao, 886 tons of plastic waste are generated per day, representing 46% at the national level [6] due to the fact that the population in Peru grows towards the cities, 75% live in urban areas, and this means more garbage

production [7], causing, in recent years, several districts such as Comas and San Martín de Porres were declared in sanitary emergency because their garbage collection systems collapsed [8]. According to the ex-Minister of the Environment Muñoz Fabiola, she indicated that in other countries there are no sanitary landfills and that absolutely everything is valued [9].

The quarantine generated by COVID-19 in the world presented positive effects at the environmental level: improvements in water and air quality and less pollution of the ozone layer [10]. The sanitary measures that are now being faced by the COVID-19 causes the index of plastic material to increase the figure spontaneously. The plastics that stand out the most for their use are gloves, masks, alcohol gel containers and plastic bags; There are also the Polyethylene due to the increase in delivery. There are no exact figures, but organizations estimate that household waste grew by 30% to 40% [11].

In Peru, a biodegradable plastic is one that degrades to carbon dioxide (CO₂), methane (CH₄), water and biomass by the action of microorganisms, contains a minimum of 50% volatile solids, has limited concentrations of dangerous chemicals and its degradation is carried out in a reasonable time: 90% degradation in 6 months in the presence of oxygen (O₂) and 2 months in the absence of O₂ [6]. These are the measures that are being taken in Peru to reduce plastic pollution, but even so the culture of recycling is not a common measure that the Peruvian citizen has in mind, it is that only 3% of Peruvians recycle the garbage they generated daily [12].

The objective of the research work is to design a plastic Plastic Shredding Machine to obtain small plastic waste, in such a way that it allows people dedicated to the recycling industry to obtain waste in an automated way and will also be very useful to reduce plastic pollution in the environment. For the development of the design, SolidWorks software was used for the mechanical part that makes up the Shredder machine and an Arduino UNO programmable board that will allow the machine to work in an automated way.

In section II, the methodology will indicate the mechanical, electronic and design of the Plastic Shredding Machine. In Section III, the mathematical calculations, the formulas that were considered for the design of the Plastic Shredding Machine will be placed. In Section IV, the results that are generated according to the tests carried out on the design of the Shredder will be presented. In Section V the discussion will be presented, where the importance of this work with respect to

other works carried out will be indicated. Finally, in Section VI, the conclusion and recommendation of the design of the Plastic Shredding Machine will be presented.

II. RELATED WORKS

Plastic waste shredding machines are of great importance as mentioned below in some studies. For example: In [13], the authors identified that when shredding disposable plastic, the small pieces can be used to make new plastic products, that is why they proposed the development of a plastic waste shredding machine, likewise, the shredding machine was designed with a traditional method of use scissors to cut materials in a small way and from the scraping used by rabbits when digging, some of the blades have sharp curved edges to attract the plastic towards the teeth of the cutting blades. The performance of the machine is 27.3 kg/h and the efficiency is 53% for all types of plastic and 95% for the type of polyvinyl chloride plastic, concluding that the machine could be very useful in a situation in which considerable plastics have to be crushed and also efficient in the crushing of large sizes.

In [14], the author identifies that the PUCP (Pontificia Universidad Católica del Perú) works with 3D printing technology in various areas of the University and that printed material is increasing since it cannot be discarded because they are highly polluting, that is why they proposed the design of a recycling machine oriented to the production of ABS plastic filaments for 3D printing in the PUCP, thus using an Arduino UNO programmable board to control the DC motors, a servo motor and a touch screen, it also uses the Autodesk Inventor 2017 software to design the mechanical part of the system. The result was that the designed system has the capacity to produce at least 0.5kg of ABS per hour, concluding that, by having an interchangeable nozzle extrusion system, a system is obtained capable of manufacturing not only filaments of 1.75, 2.85 and 3.00 mm in diameter as it was determined in the beginning, but also other variants that are in this range of values.

Finally, in [15], the authors identified that most of the PET bottles produced are not recycled, on the contrary, new bottles are produced every time, thus increasing waste, which is why they proposed to carry out the design and construction of a plastic Shredder for the recycling and management of plastic waste, generated in the area of the Petroleum Training Institute, Effurun, Nigeria, thus using a crushing chamber made of a rigid and thick 0.610mm steel plate, also a 0.22mm steel hopper where plastic waste accumulates while shredding is carried out with a base of the shredder casing that has a wire mesh used to regulate the type and size of the shredded plastic waste. Obtaining as a result 98.44% of crushing efficiency, with a crushing rate of 0.575kg/s, concluding that the crushed plastic waste has a size that ranges between 10mm and 20mm and the results obtained reveal that the performance of the machine it is satisfactory.

III. METHODOLOGY

A. Mechanical Part

Since the shredding machine will be in constant friction from the cuts of the plastic material, it must be made of a durable material. For this, suitable materials were chosen for the mechanical assembly of the Shredder.

- Aluminum
- Stainless steel
- Close type pulley
- Bearings

Aluminum will be used for the case of the Plastic Shredding Machine, which is also a lightweight and durable material. The blades of the plastic waste shredding machine will be made of stainless steel to maintain the durability of the friction when making the cuts, the shaft that will hold the blades will also be made of stainless steel, which will be supported by rolling bearings for mobility shredding machine.

B. Electronic Part

The electronic part is composed of a DC motor, which will be an important factor through of its force it will allow to crush the plastic waste that is placed in the machine, for this, the power of the motor must be considered. For this project, the DC motor that we will be using has GGM geared motors of the helical coaxial type in DC with powers from 120W M6 with 3.0 N.m (30kg.cm).

It will also have an Arduino UNO board to work together with a driver module and be able to control the DC motor, in such a way that the shredder has an automatic function.

C. Design

To arrive at the result of the design that will be presented in the article, a preliminary study was made which consists of the durability of the product, its efficiency, and its sustainability over time. Mention will be made of the design of the different parts that make up the crushing machine.

The most important evaluation criteria for the "Crushing and Classification" module were selected in the conceptual phase, which were:

- Ease of manufacture
- Ease of assembly and disassembly
- Reliability
- Ease of maintenance
- Low noise and vibration level
- Inexpensive and affordable spare parts
- Recirculation of material
- Under weight
- Little wear on the cutting elements

1) *Metal sieve or sieve in the shape of a half moon*: The classification of the material already crushed has the appropriate size. To fulfill this function, a metal sieve or a crescent-shaped sieve is used as shown in Fig. 1. Since, thanks to its geometric shape, it does not allow the accumulation or waste of plastic material at any point. Likewise, thanks to its reduced distance to the cutting elements, it allows the

recirculation of material not conforming to the appropriate size to reintegrate it in the crushing process.

2) *Storage module*: The most important evaluation criteria for the "Storage" module, as shown in Fig. 2, were selected in the conceptual phase, which are:

- High storage capacity
- Under weight
- Ease of Manufacturing
- Ease of downloading material
- Ease of Installation
- Safety for the operator

3) *Shredding module*: This module allows the size reduction process to be carried out through the cutting elements, which reduce the plastic material that enters the feed hopper to the appropriate size.

Off-center blades provide the optimal solution to "push" parts into the cutting chamber in the same way that they allow for return. This configuration allows for even distribution of cutting force and simple scissor cutting for greater energy savings, as well as quiet operation and high cutting performance. The geometry of the cutter blades as shown in Fig. 3, allows them to be changed without the need for adjustments, in addition to reducing downtime due to blade change problems.



Fig. 1. Half-Moon-Shaped Sieve for the "Crushing and Classification" Module".

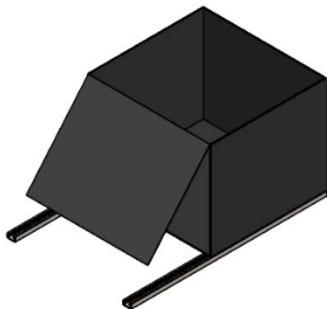


Fig. 2. Storage Module.

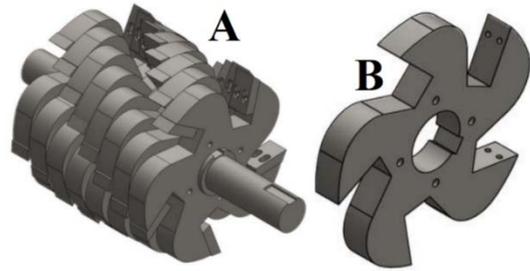


Fig. 3. A) Staggered Blade Design B) Blade Design.

4) *Crushing shaft design*: In this section, the construction of the crushed sleeper is carried out, which is subjected to bending and torsion loads created by the power transmission of the pulley, the cutting forces, and the weight of the elements it contains, as shown in Fig. 4.

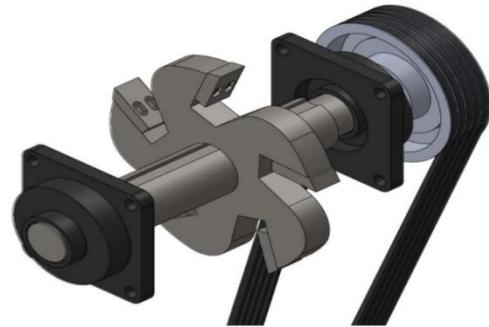


Fig. 4. Isometric View of the Crushing Shaft.

The mechanical part is an important factor, because the machine will be in constant friction that is why durable materials should be considered as mentioned above, which together with the electronic part manage to crush the plastic in such a way that the machine is efficient and not present complications in its operation.

IV. MATHEMATICAL CALCULATIONS

A. Mathematical Calculations

To start the calculation and selection of the different components of the crusher, it is necessary to know the resistance to cutting of the plastics most used in industries. Table I shows the tensile strength values of different materials. It is also necessary to know the resistance to shear of the selected materials, which through of a mathematical formula is the values as shown in Table II.

TABLE I. TENSILE STRENGTH VALUES

Material	kg/cm ²
Polyethylene (PE)	110 – 375
Polypropylene (PP)	500
Polystyrene (PS)	350 – 600
PVC	350 - 630

TABLE II. CUT RESISTANCE

Material	kg/cm ²
Polyethylene (PE)	88 – 330
Polypropylene (PP)	400
Polystyrene (PS)	280 – 480
PVC	280 - 504

With the values already exposed, the resistance to shear cutting will be calculated, through of the following formula [16]:

$$\tau = \frac{4}{5} \times \sigma \text{ (kg/cm}^2\text{)} \quad (1)$$

Where:

τ = Resistance to shear cutting (kg/cm²).

σ = Tensile strength (kg/cm²).

In Table III, the values of charpy impact resistance and melting temperature are shown.

TABLE III. IMPACT RESISTANCE AND MELTING TEMPERATURE

Material	Charpy impact resistance (kg/cm ²)	Melting temperature (°C)
PET	3.6	170 – 270
PVC	15	80
PP	4 – 20	160 – 270
PS	3 – 12	105 – 270
PE	18	70 – 120

As mentioned above, polypropylene was chosen, which has higher impact resistance than other thermoplastic polymers. To calculate the force required to generate the failure of the polypropylene, the work required to perform the failure of the material must first be calculated, which is calculated using Equation 2 [16].

$$\partial w = \partial W^{el} + G_c \times l \times \partial a \quad (2)$$

Where:

∂W = Work required to break the material (Joule).

∂W^{el} = Elastic energy change (Joule).

G_c = Energy absorbed per unit area (Joule/m²).

l = Length of the material cut by the blade (length of the cutting edge) (m).

∂a = Advancement of the fracture during impact (m/s).

Also, Fig. 5 shows how the material will be cut, the so-called scissor cut, which is essentially a shear cut that offers greater energy savings.

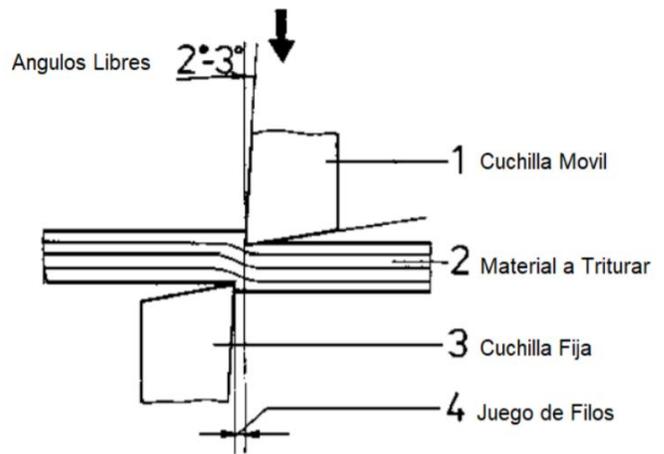


Fig. 5. General Structure of Shear or Scissor-Type Cutting.

Based on Fig. 6, Newton's second law is implemented, which allows us to relate the movement of the body with the forces acting on it through equations 3 and 4, which are written in their tangential and normal components.

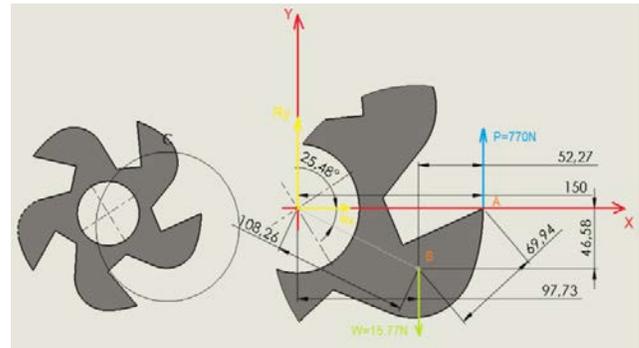


Fig. 6. Center of Mass of the Blade Holder.

$$\sum F_t = m \times a_t \quad (3)$$

$$\sum F_n = m \times a_n \quad (4)$$

Where:

m = Mass of the element (kg).

F = Forces on the element (N).

a_t = Tangential acceleration (m/s²).

a_n = Normal Acceleration (m/s²).

Tangential acceleration and normal acceleration are defined by equations 5 and 6.

$$a_t = r \times \alpha \quad (5)$$

$$a_n = r \times \omega^2 \quad (6)$$

Where:

r = Radius (m).

α = Angular Acceleration (rad/s^2).

ω = Angular speed (rad/s).

And finally, Equation 7, which refers to the sum of the moments around a point, will be applied to determine the reactions and the angular speed of the body; it should be considered that this is a non-centroid movement, since the axis of rotation does not coincide with the center of mass of the body, for this reason the system of external forces is not reduced to a couple $I\alpha$, in the same way we consider that the rotation is constant, therefore the angular speed (ω) is constant and the angular acceleration (α) is zero.

$$\sum M_0 = I \times \alpha \quad (7)$$

Where:

M_0 = Moments of a point (Nm).

I = Moment of inertia ($kg \times m^2$)

Through equations 3 and 4 the force summations will be carried out.

V. RESULTS

In the article, a complete study was elaborated on the efficiency of the plastic crushing machine that will have a sustainability in time with the role that it is exercising. With this, the purpose of reducing pollution by plastic waste would be fulfilled, since it would fulfill its main role, which is to obtain small plastic waste. Inside the shredding module, the shaft formed by the blades is placed as shown in Fig. 7. The container where the small parts of the plastic shredded by the blades accumulate is shown in Fig. 8.

The design of the plastic waste shredding machine would have an approximate efficiency of 85%, because it does fulfill the main job that is shredding plastic, but that missing percentage is since the blades are not strong enough for the plastic bottles additionally.

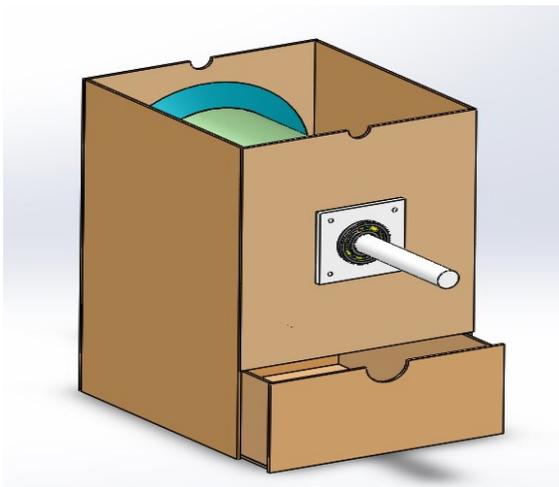


Fig. 7. Shredder Module.

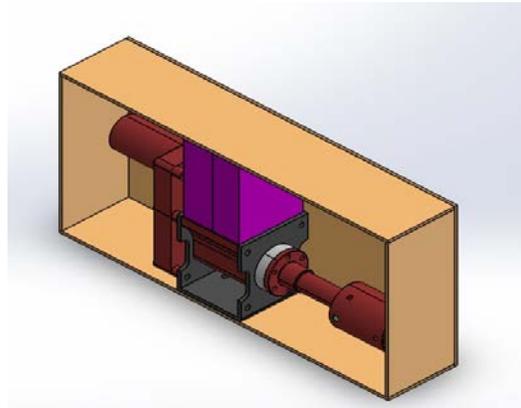


Fig. 8. Storage Module.

In Fig. 9, the implementation of a plastic shredder machine that generates filaments for 3D printers is observed, where the motor will help in the rotation of the blades and thus crush the plastic bottles with the help of an extruder generate filaments that are used a lot in 3D printers. In addition, the power source to feed the system is also identified in the lower part of the motor, so this would be a more complete project.



Fig. 9. Plastic Shredding Machine that Generates Filaments for 3D Printers.

VI. DISCUSSION

The articles made about shredding machines share the same purpose, which is to reduce plastic waste that pollutes the environment over the years. The design of a shredding machine does not require a microcontroller as mentioned [13], since to carry out the crushing function only a good mechanical part is needed, that is why it adapts to the work carried out by [15], since the machine when shredding plastic does not require any programming. On the other hand [14], it indicates that it is necessary to use an Arduino UNO since it will give one more function to the shredding machine, therefore, it is necessary to indicate what function it will perform the shredding. Although it is true, each job has a specific efficiency, and they are of great importance for those people who are dedicated to the recycling business. It should be noted that these works serve as a basis for other projects that are intended to be carried out.

VII. CONCLUSIONS

It is concluded that the plastic crushing machine works efficiently, this would confirm that the use of this system will help reduce the pollution generated by the plastic, since by

crushing the plastic bottles it allows them to be recycled faster than traditional recycling.

It is concluded that the crushing machine has a structure that is not strong enough to push and bring the plastic towards the blades, due to this it has difficulty in crushing some plastic bottles. Although, it is capable of crushing Technopor, it is a more porous and structured alloy of small balls.

It is concluded that the use of metal blades is strong enough to crush plastic bottles, initially plastic blades were considered for the crusher, but they could not crush the plastic bottles.

It is concluded that the use of the Arduino UNO board helps a lot to control the motors that are part of the shredder machine, and its programming is more user-friendly unlike other microcontrollers.

As work in the future, a heating system will be placed as shown in Fig. 9, which will be of great help to obtain filaments for 3D printers that are used by various professionals, for example, in the biomedical field. This would be making the most of the plastic shredded by the machine.

It is recommended to consider that the blades to be incorporated into the structure of the crushing machine can withstand the force generated by the motor when the machine is in operation. It is also necessary that the blades are staggered on their shredding axis as it allows the cutting to be more efficient.

REFERENCES

- [1] C. Garcia Nuñez, C. Ruiz Chacon, E. Uchofen Diaz, A. Velasquez Kolich, and R. Villalta Herrera, "Producción y comercialización de colgadores de prendas de vestir elaborados a base de botellas de plástico recicladas," Universidad San Ignacio de Loyola, Lima -Perú, 2020.
- [2] E. Murrieta Flores, "Implementación de un sistema de conteo de botellas reciclables para delimitar el tiempo de carga de celular," Quito - Ecuador, Jun. 2019.
- [3] M. Jaén, P. Esteve, and I. Banos-González, "Los futuros maestros ante el problema de la contaminación de los mares por plásticos y el consumo," Rev. Eureka sobre Enseñanza y Divulg. las Ciencias, vol. 16, no. 1, pp. 1–17, Aug. 2019, doi: 10.25267/reveurekaensendivulgcienc.2019.v16.i1.1501.
- [4] J. Bolaños Zea, "Reciclado de Plástico PET," Universidad Católica San Pablo, Arequipa - Perú, 2019.
- [5] Y. Flores Icochea, A. Huaraca Huaman, C. Lujan Espinoza, and O. Retamozo Calle, "Efectividad de la bacteria Ideonella Sakaiensis para la biodegradación de materiales de Polietileno Tereftalato (PET), en el periodo 2019 II," Universidad Cesar Vallejo, Lima -Perú, 2019.
- [6] P. Mendoza Osorio, "Diseño persuasivo en la concientización sobre el uso de bolsas plásticas en mercados de Lima," Universidad Peruana de Ciencias Aplicadas, Lima -Perú, 2019.
- [7] C. Huamaní Montesinos, J. Tudela Mamani, and A. Huamaní Peralta, "Gestión de residuos sólidos de la ciudad de Juliaca - Puno - Perú," Rev. Investig. Altoandinas - J. High Andean Res., vol. 22, no. 1, pp. 106–115, Jan. 2020, doi: 10.18271/ria.2020.541.
- [8] C. Murillo Salinas and N. Ochante Enriquez, "Manejo de residuos sólidos en defensa del derecho al ambiente frente al Covid-19 en el Mercado Santa Luzmila, Comas 2020," Lima - Perú, 2020.
- [9] R. García Delgado and M. Silva Tarrillo, "Evaluación de adoquines que contienen agregados de canteras y vidrio reciclado," Rev. Nor@ndina, vol. 3, no. 2, pp. 123–132, Dec. 2020, doi: 10.37518/2663-6360x2020v3n2p123.
- [10] F. Ormaza González and D. Castro Rodas, "Covid-19 Impacts on Beaches and Coastal Water Pollution: Management Proposals Post-pandemic," Preprints, Jun. 2020, doi: 10.20944/preprints202006.0186.v1.
- [11] L. Visitación Figueroa, C. Nieto Aravena, and A. Visitación Figueroa, "Experiencias y avances en la política sobre la gestión de residuos sólidos en el Perú," Policy Advances in Solid Waste Management in Peru," Rev. LIDER, vol. 19, pp. 9–21, Jun. 2017.
- [12] A. Remaycuna Vasquez, "Construcción y Procesos Psicométricos de una Escala de Actitudes Frente al Reciclaje en Universitarios Peruanos 2020," PAIAN, vol. 11, no. 2, pp. 30–42, Dec. 2020, doi: 10.26495/rcp.v11i2.1514.
- [13] A. Waleola Ayo, O. Olukunle, and D. Adelabu, "Development of a Waste Plastic Shredding Machine," Int. J. Waste Resour., vol. 07, no. 02, 2017, doi: 10.4172/2252-5211.1000281.
- [14] J. P. Porras Solorzano, "Diseño de una máquina recicladora orientada a la producción de filamentos de plástico ABS para la Impresora 3D en la PUCP," Pontificia Universidad Católica del Perú, Lima-Perú, 2018.
- [15] A. David and O. Oluwayomi Joel, "Design and Construction of a Plastic Shredder Machine for Recycling and Management of Plastic Wastes," Int. J. Sci. Eng. Res., vol. 9, no. 5, May 2018.
- [16] D. Esteban and F. Julián, "Mecanizado básico. Novedad 2017", Editex, 2017.

An HC-CSO Algorithm for Workflow Scheduling in Heterogeneous Cloud Computing System

Jai Bhagwan¹, Sanjeev Kumar²

Department of Computer Science & Engineering
Guru Jambheshwar University of Science & Technology
Hisar, India

Abstract—Many scientists are using meta-heuristic techniques for dynamic workflow task scheduling in the area of cloud computing systems to get optimum solutions. Many swarm intelligent algorithms have been designed so far which are having many limitations as some get trapped in local optima, a few are having low convergence speed, some are having poor global search facilities, etc. Still, there is a requirement of designing a new algorithm or modification of existing algorithms to overcome the limitations of the existing techniques. A new Hybrid Cat Swarm Optimization algorithm named H-CSO was designed inspired by the HEFT algorithm and the initialization problem of the Cat Swarm Optimization was overcome. Still, that algorithm has a limitation of getting stuck in local minima. To overcome this algorithm a part of the Crow Search Algorithm has been integrated into H-CSO and described in this paper. After simulation, it was found that the new hybrid algorithm named HC-CSO outperforms CSO and H-CSO.

Keywords—Cloud computing; Crow Search Algorithm (CSA); Cat Swarm Optimization (CSO); H-CSO; HC-CSO; HEFT; Self-Motivated Inertia Weight (SMIW); Virtual Machines (VMs)

I. INTRODUCTION

Information Technology has been reshaped by the evolution of cloud computing technology via big storage facilities, high-performance computing, and other hardware and software services. The current technology includes the evolution of computing eras where computers were connected via the internet that took the form of distributed computing. This further transformed into cluster computing, cluster to grid computing, and then cloud computing [1]. The major aim of cloud computing technology is to provide high-performance computing services at the minimum cost. The cloud technology shifted the users' data from client machines to network-abled machines which are having high-powered processors and hardware parts. Cloud computing provides services in the form of Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [15]. Any end-user can pick up the services as per his requirements. The main benefit of the cloud is that it doesn't include geographical boundaries to provide the services to the end-users [2]. This means that the users need not to know the physical locations of the service providers and computing datacenters. Cloud technology is flexible because a user can increase the number of services and drop whenever it is required due to the pay-per-usage policy. Various cloud service providers are Amazon Web Services, Microsoft, Google, Rackspace, salesforce.com, etc. The cloud system

can be classified into four categories such as private cloud, public cloud, community cloud, and hybrid cloud [3].

The performance of the cloud can be improved at various levels such as network level, scheduling level, database level, etc. After studying various research papers, it can be seen that numerous algorithms have been designed for workflow task scheduling, load balancing, energy consumption management, etc. in cloud computing. Workflow task scheduling is one of the major areas where a lot of improvement is required. Task scheduling can be of two types static and dynamic scheduling. In static scheduling, the execution times of the tasks are pre-estimated or known but in the case of dynamic scheduling, it is not known. This era is of dynamic task scheduling. So, dynamic task scheduling algorithms are required to be improved. The optimization techniques play an important role to improve the cloud scheduling problem nowadays. A few famous techniques are Ant Colony Optimization, Particle Swarm Optimization, and Cat Swarm Optimization, etc. Cat Swarm Optimization belongs to Swarm Intelligent (SI) family [4]. The algorithms used in this paper are as under:

1) *Cat swarm optimization*: This is an intelligent algorithm originally developed in the year of 2006. This algorithm is inspired by the behaviour of the original cat which is having to two modes known as seeking (resting mode) and tracing (attacking mode). Here, N numbers of cats are generated randomly and each cat denotes a solution, its position, a flag, and fitness value. M dimensions in the search space represent the position, and each dimension in the search space is having its self-velocity. The flag is used to identify that the cat is in either seeking mode or in tracing mode and this flag is set by a parameter known as Mixing Ratio (MR). After finding the fitness, the best cat is stored in the memory at each iteration and finally, the best solution or cat is identified at the end of the final iteration. The two modes of Cat Swarm Optimization are described below [1][21]:

i) *Seeking mode*: This mode represents the resting mode of a cat and four parameters play an important role in it: SMP (Seeking Memory Pool), SRD (Seeking range of the Selected Dimension), CDC (Counts of Dimension to change), and SPC (Self Position Consideration). In SMP, one position is selected by a cat randomly for moving to the next position. Let's say, the SMP is set to 10, now for every cat 10 new random positions will be generated and one among them will be

measurement. The proposed technique worked better as compared to Multi-Objective Ant Colony Optimization, Multi-objective Genetic Algorithm, and Multi-Objective Particle Swarm Optimization. The authors [2] proposed a Multi-Objective Cat Swarm Optimization algorithm and after comparison with the existing Multi-Objective PSO technique, the proposed approach was found better in the account of energy consumption, execution cost, and execution time. The researchers in [3] introduced a Cat Swarm Optimization-based technique for workflow scheduling. The results were compared with PSO (Particle Swarm Optimization) and found that the CSO reduces the processing cost, made a good load balance, and gave the optimum results in less iteration. In [5] a Hybrid Particle Swarm Optimization algorithm using the Hill-Climbing technique was proposed by the scientists. The proposed algorithm was found effective in terms of makespan after experiments. The scientists [6] introduced a Binary Hybrid Particle Swarm Optimization and Gravitational Search algorithm for load balancing of virtual machines (VMs). The experimental results showed that the proposed algorithms worked better than the existing Binary PSO load balancing algorithm in terms of load balancing. In [7] the researchers offered a new HPSOGWO algorithm that is a combination of Particle Swarm Optimization and Grey Wolf Optimization. The idea behind this algorithm was to improve the exploitation of PSO and exploration of GWO for making the better strength of the proposed algorithm. The results concluded that the HPSOGWO is better than standard PSO and GWO variants concerning solution stability, convergence speed, and quality. The authors in [8] proposed an IPSO (Improved Particle Swarm Optimization) algorithm to improve the allocation of large-length tasks. The introduced algorithm outperformed despite Ant Colony, Honey Bee, and Round-Robin algorithm in accounts of load balance, makespan, and degree of imbalance. In [9] authors presented a Particle Swarm Optimization based technique for workflow scheduling and found that the Particle Swarm Optimization based technique saved the cost equal to 3 times as compared to BRS (Best Resource Selection) algorithm. Also, the presented technique balanced the load efficiently. The authors of the paper [10] proposed a PSO (Particle Swarm Optimization) based scheduling technique for independent tasks. The proposed technique was improved using a load balancing strategy. The newly introduced method was compared with Improved PSO, Round-Robin, and existing load balancing techniques. The experimental results showed that the proposed technique is better than the above said three algorithms for resource utilization and makespan. The scientist [11] introduced the MPSO (Modified Particle Swarm Optimization) algorithm for the reduction of cost as compared to the existing Particle Swarm Optimization algorithm. The simulation results showed that the proposed technique worked better. In paper [12] the authors modified Particle Swarm Optimization by modifying the parameters like MIPS and Bandwidth for effective load balancing. The simulation results showed that the proposed model resulted in a reduction of execution time. The researchers of the paper [13] compared a

task scheduling strategy based on Ant Colony Optimization with FCFS and RR (Round-Robin). The simulation results proved that the ACO outperforms traditional FCFS and Round-Robin algorithms. The authors [14] developed a hybrid optimization technique using Flower Pollination and Grey Wolf Optimization. The PEFT algorithm was used for the initialization of the proposed method for workflows. The simulation resulted that the proposed method is effective as compared to Flower Pollination with Genetic Algorithm in terms of cost and reliability. In [16] the researchers designed an Improved Social Learning Optimization Algorithm by introducing the Small Vector Position method for task scheduling problems. After simulation, it was found that the proposed approach worked well as compared to GA and PSO-based techniques. In the paper [17], the researchers proposed an Adaptive Cost-based Task Scheduling technique to scheduling the tasks between virtual machines at minimum cost. The simulation results concluded that the proposed technique is performing well in terms of communication cost, execution time, CPU utilization and execution cost rather than cost-efficient task scheduling. In [18] research paper, the scientists implemented a Dynamic Adaptive Particle Swarm Optimization algorithm to enhance the efficiency of Particle Swarm Optimization for better makespan, resource utilization. The authors also proposed an algorithm named MDAPSO. DAPSO and MDPSO worked better than the original Particle Swarm Optimization. The authors of [19] research paper proposed an Improved Particle Swarm Optimization based technique to solve workflow scheduling problems in cloud systems. The proposed method worked better as compared to existing state-of-art methods. In [20], the scientists proposed a cloud scheduling model named Task Scheduling System. The proposed Genetic Algorithm-Chaos Ant Colony Optimization worked better than ACO and GA algorithms in respect to cost and convergence speed.

From the above study, it is found that the ACO, PSO, and CSO algorithms are optimizing the scheduling of independent as well as workflows tasks. The ACO algorithm is performing well for local searching; the PSO and CSO algorithms are good for global searching and get stuck in local optima easily. Various researchers integrated a few techniques and formulae in these algorithms to improve the performance of the cloud. But, these techniques are old. The CSO algorithm is a good performer among the ACO and PSO; the H-CSO is better than the CSO.

So, a new method is required to integrate into the H-CSO so that its limitation of getting trapped in local optima can be overcome. Hence, the local searching part of the Crow Search Algorithm has been integrated into the H-CSO algorithm. The details are indicated in the proposed methodology section.

III. PROPOSED METHODOLOGY

From the related work, it is learned that a single algorithm could not be able to give better results for workflow task scheduling.

Proposed Algorithm

Input (Tasks (T₁, T₂, T₃ ... T_n), Virtual Machines (VM₁, VM₂, VM₃ ... VM_m))

Output (Optimal Makespan and Cost of n Tasks on m VMs)

BEGIN PROCEDURE

1. Initialize fl (flight length), r_k, velocity factor V_k, c, γ_{max}, Coefficient c1, MR flag, and no. of iterations
/* Calculate Rank of Workflows Tasks in DAG using HEFT Algorithm */
2. Feed workflows in HEFT
3. **For** Each Task in DAG **Do**
4. Calculate average execution time of all VMs
5. **If** Task t_i is the last Task **Then**
6. Rank value of t_i = its average execution time
7. **Else**
8. rank_u(t_i) = WAv_g_i + Max t_j ∈ succ (t_i) (CAvg_{i,j} + rank_u(t_j))
 Where WAv_g_i is average execution cost
 Succ (t_i) is set of immediate successor of task t_i
 CAvg_{i,j} is average communication cost
9. **End If**
10. **End For**
11. Assign Tasks to VMs according to HEFT Rank
12. **If** Solution not Optimized **Then**
13. Generate a set of Crows by the Population generated by HEFT of Size N
14. **While** No. of Iterations not Exceeded **Do**
15. **For** K=1 to N
16. Update the positions by the following equation : // Do local search
17. $X_{K,D} = X_{K,D} + r_K * fl_{K,D} * (M_{L,D} - X_{K,D})$
 Where, X_{K,D} is current position of Crow_k, r_k is uniformly distributed random number [0, 1]
 fl_{K,D} is flight length (less than 1 i.e. 0.5) of the Crow_k at current iteration
 M_{L,D} is present best location of Crow_k in Dimension D
18. Feed the population generated by Local CSA in H-CSO // Do local and global search as given below
19. Assign the velocity V_k to each Cat
20. According to Mixing Ratio (MR) flag Distribute Cats to Seeking and Tracing Modes
21. **If** current Cat_k is in Seeking Mode **Then**
22. Generate S (SMP) Copies of Cat_k and Spread them in D Dimensions where each Cat has a velocity (V_{k,D})
23. Evaluate the Fitness value of all Copies and Discover Best Cats (X_{BEST,D})
24. Replace Original Cat_k with the Copy of Best Cats (X_{BEST,D})
25. **Else If** current Cat_k is in Tracing Mode **Then**
26. Compute and Update Cat_k velocity by following equations:
27. $\gamma = \gamma_{max} * \exp\left(-c * \left(\frac{itr}{itr_{max}}\right)^c\right)$
 Where, γ is a weight factor calculated by Self-Motivated Inertia Weight method
 γ_{max} and c are constant factors greater than 1, both are set as 2.
28. $V_{K,D} = \gamma * V_{K,D} + \left(c1 * r1 * (X_{BEST,D} - X_{K,D})\right)$
 Where, D = 1, 2, 3, ..., M.
 c1 is acceleration coefficient, r1 is random number in the range of [0, 1]
29. Update the position of every dimension of Cat_k by using following equation:
30. $X_{K,D} = X_{K,D} + V_{K,D}$
31. Evaluate Fitness of all Cats and find out Best Cats (X_{BEST,D}) with Best Fitness
32. **End If**
33. Update Best Cats (X_{BEST,D}) in Memory
34. **End For**
35. **End While**
36. **End If**
37. return (Optimal Solution)

END PROCEDURE

Fig. 2. Pseudo-Code of Proposed HC-CSO Algorithm.

Many scientists advised improving the existing algorithms. The H-CSO algorithm is a combination of the HEFT, Cat Swarm Optimization algorithm, and Self-Motivated Inertia Weight useful in overcoming the velocity outrange problem of the standard CSO. As said earlier, this H-CSO algorithm gets trapped in local minima in the case of a large search space and complex workflows' tasks environment. A few limitations of the H-CSO algorithm are described below in short:

1) The H-CSO algorithm is having only a good global searching capacity.

2) The H-CSO algorithm gets trapped in local minima due to a large number of cats is always residing in seeking Mode as compared to tracing mode.

3) Due to a lack of balance between seeking and tracing modes, optimal results could not be got using the H-CSO algorithm.

In the proposed algorithm named HC-CSO, a local search part of a well-known Crow Search Algorithm is integrated into the H-CSO for avoiding it getting trapped in the local optima. The working of the HC-CSO algorithm is described as:

First of all, the initialization of various parameters takes place then pre-processing of workflows tasks is executed with the HEFT method as shown in the pseudo-code of the proposed algorithm. After the pre-processing of workflows, the tasks are assigned to available VMs. If this solution is getting optimized at the very first stage then the algorithm is stopped and returns the optimized schedule otherwise the initial solution generated by the HEFT algorithm is fed to the CSA algorithm. Here, a number of N Crows are generated using the population obtained by the local Crow Search Algorithm. Then, the population got from the local CSA is fed to the H-CSO algorithm. Now, N numbers of Cats are generated and velocity value V_k is assigned to each cat for further processing. In the next phase, the cats are randomly distributed into the seeking and tracing mode as per the Mixing Ratio (MR) rate and flags are set to each cat. If the current Cat_k is found in the seeking mode as per flag value, this mode gets executed otherwise the tracing mode gets executed. The position and velocities of the cats in the tracing mode are updated using the equations given in the pseudo-code in Fig. 2 and the best cats get stored in the memory in the form of solution. This process is continued until the termination condition is not matched and finally an optimal solution is returned in the form of better makespan and cost. The proposed algorithm has both the best local and global searching capacity, so the results could be reached at an optimum level. The proposed algorithm is also useful to avoid the premature convergence of the H-CSO method. The pseudo-code of the proposed algorithm named HC-CSO has been described step by step in Fig. 2.

IV. SIMULATION SETUP

The simulation environment has been created in the CloudSim tool for simulating the different workflows used in this paper. Various experiments were carried out over a computing machine having the configuration as Processor – Intel® Core™ i3-5005U at a speed of 2.0 GHz, RAM – 4.0 GB, HDD – 1 TB, and OS – Windows 10.

A. Parameters

For simulation, a PowerDatacenter has been designed having the configuration as RAM – 25 GB, MIPS per VM – 1000 MIPS, Storage – 1 TB, Bandwidth – 50000 bps. Rest configurations of the heterogeneous cloud environment are depicted in Table I. The scheduling policy was set as Time Shared. The parameters of the CSO, the H-CSO and, the proposed HC-CSO are also summarized in Table I.

TABLE I. SIMULATION PARAMETERS

PowerDatacenter	
Parameters	Values
Number of Hosts	1
System Architecture	x86
VMM	Xen
OS	Linux
Number of Cloudlets	1000 in Each Workflow
Cloudlets Length Type	Random
Numbers of VMs	10, 20 and 30
CPU (PEs Number)	1
RAM per VM	512-1024 MB
Bandwidth	1000-1500 bps
Processing Elements per VM	500 – 1000 MIPS
Image Size	10000 MB
Policy Type	Time Shared
Standard CSO and H-CSO	
No. of Cats	100
Iterations	300
Weights (C1)	1.5
r1 and r _k (Random Variables)	[0, 1]
Mixed Ratio Percentage	Random Range [0, 1] i.e. 0.2-0.3
HC-CSO (Proposed Algorithm)	
No. of Cats	100
Iterations	300
Weights (C1)	1.5
r1(Random Variable)	[0, 1]
Mixed Ratio Percentage	Random Range [0, 1]
fl (Flight Length)	0.5

B. Cost Plan

The cost plan (in Indian Rupees) of workflow scheduling is depicted in Table II.

TABLE II. COST PLAN

Resource	Processor	RAM	Storage	Bandwidth
Size	500-1000 MIPS	512 MB	Unlimited	1000 bps
Cost	Rs. 3.0 per processor	Rs. 0.05 per MB	Rs. 0.1	Rs. 0.1 per MB

C. Cloudlets

Cloudlets are called tasks to be submitted for execution on virtual machines. In this paper, the scientific workflows named CyberShake, Montage, Inspiral, and Sipt have been used to test the proposed algorithm along with others. Each workflow is having 1000 tasks.

D. Performance Metrics

There are several performance metrics like Makespan, Processing Cost, Waiting Time, Response Time, Energy Consumptions, and Resources Utilization, etc. to test the performance of the algorithms. In this paper, makespan and cost are used to measure the performance of the proposed algorithm, and these parameters are given in the coming topics.

1) *Makespan*: Makespan [23] is referred to the maximum time taken for finishing the last task in a group. It is the most widely used metric to measure the performance of a scheduling algorithm. A lesser makespan decides that the algorithm is efficient enough. The makespan is computed by the given equation 5.

$$\text{Makespan} = \max (CT_i)_{ti \in \text{tasks}} \quad (5)$$

Where, CT_i is the completion time of task $_i$

2) *Processing cost*: Processing Cost is along with makespan is another important performance metric because cloud service providers want to give efficient services at the minimum costs in this competitive environment. The processing cost can be measured by equation 6.

$$\text{Total Cost} = \frac{MF+CF}{2} \quad (6)$$

Where, MF is Movement Factor

CF is Cost Factor

$$MF = \frac{1}{\text{No. of Hosts/Datacenters}} \left[\sum_{x=1}^{VMx} \left(\frac{\text{Number of Migrations}}{\text{Used VM}} \right) \right] \quad (7)$$

$$CF = \sum_{x=1}^{VMx} \left(\frac{\text{Processing Cost} * \text{Memory of Tasks}}{VM * \text{Datacenter}} \right) \quad (8)$$

3) *Fitness function*: Equation 9 is depicting the fitness function which is used by meta-heuristic techniques to check that the solution is optimized or not at various levels in the form of makespan along with computation cost.

$$F_X = \frac{1}{\text{Datacenter} * VMj} \left[\sum_{i=1}^{DCi} \sum_{j=1}^{VMj} \frac{1}{VM} \frac{CPU \text{ Utilized}}{CPUij} + \frac{\text{Memory Utilized}}{\text{Memoryij}} + \frac{\text{Makespan Utilized}}{\text{Makespanij}} + \frac{\text{Bandwidth Utilized}}{\text{Bandwidthij}} \right] \quad (9)$$

Equations 7 and 8 help in the calculation of the total processing cost represented in equation 6.

V. SIMULATION RESULTS AND DISCUSSION

Four scenarios and a set of 10, 20, and 30 virtual machines have been set in CloudSim for simulation. To evaluate the performance of the proposed algorithm HC-CSO, the simulation results were compared with standard CSO and H-CSO. The makespan results concerning all scenarios are displayed in Table III. The results shown in Table III are the average of the results retrieved by the proposed algorithm HC-CSO along with other algorithms which were executed several times.

Simulation results of the proposed algorithm HC-CSO, H-CSO, and standard CSO for the scientific dataset

Cybershake_1000 are displayed in Fig. 3. Southern California Centre collected data and made the CybeShake workflow to analyze seismic hazards. For execution, this workflow requires almost near to lower CPU power and memory. It can be seen that at the x-axis 10, 20, and 30 VMs are being displayed and makespan at the y-axis. The results express that the HC-CSO outperforms other algorithms better local search after adding the Crow Search Algorithm and a better balance between seeking and tracing mode.

It is seen in Fig. 4 that the experiment carried out with the Montage_1000 dataset needs less memory and CPU power as compared to other workflows. The Montage dataset is having astronomical images collected and stored by NASA. The graph tells that the group of 10, 20, and 30, VMs as well as makespan, are shown at the x-axis and y-axis correspondingly. The results depiction tells the proposed algorithm HC-CSO performed efficiently as compare to other algorithms due to better convergence in less number of iterations.

TABLE III. MAKESPAN EVALUATION (IN SEC)

Scenarios	VMs	CSO	H-CSO	HC-CSO
Scenario - 1 CyberShake_1000	10	4603.23	4520.45	4413.65
	20	3693.52	3500.27	3425.29
	30	2401.29	2350.15	2220.15
Scenario - 2 Montage_1000	10	2310.70	2270.09	2210.43
	20	2021.13	1982.05	1813.58
	30	1308.28	1219.01	1187.67
Scenario - 3 Inspirial_1000	10	52300.41	50940.98	49859.35
	20	40350.03	39241.18	37174.27
	30	23975.57	22548.23	20140.03
Scenario - 4 Sipht_1000	10	32820.02	31107.81	30740.41
	20	25203.13	23293.24	22980.37
	30	19113.29	18279.02	18033.13

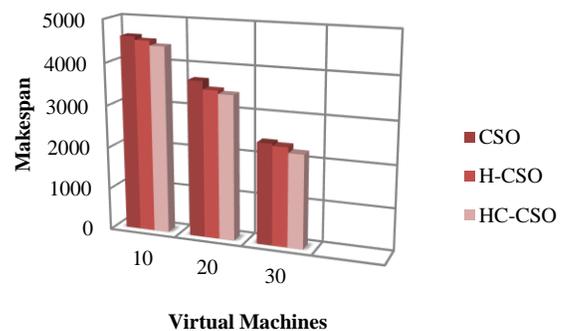


Fig. 3. Makespan Comparisons for CyberShake_1000 Tasks.

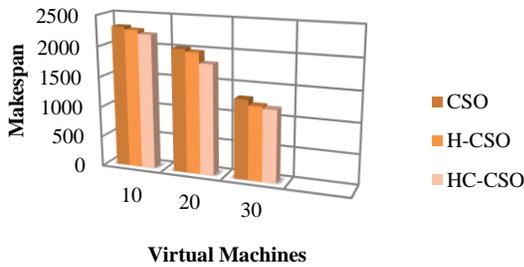


Fig. 4. Makespan Comparisons for Montage_1000 Tasks.

Fig. 5 is demonstrated the experimental results of the proposed algorithm named HC-CSO along with other algorithms like CSO and H-CSO for the dataset Inspiral_1000 which is related to the physics field and used to analyze the gravitational waves. This dataset needs high powered CPU and a large amount of memory for execution. The results concluding here that the proposed algorithm performed better than other algorithms because the HC-CSO algorithm has both capabilities of global and local searching as well as the algorithm also manage the velocities outrange. This capability of the proposed algorithm manages under-loaded and overload machines effectively by task migration. In this graph, the numbers of VMs are being displayed at the x-axis and makespan at the y-axis.

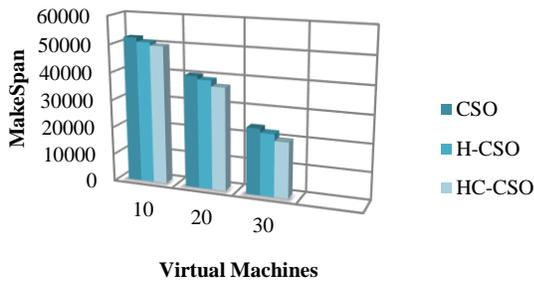


Fig. 5. Makespan Comparisons Inspiral_1000 Tasks.

Fig. 6 is displaying the results of the Sipt_1000 workflow processed with the proposed HC-CSO algorithm and others.

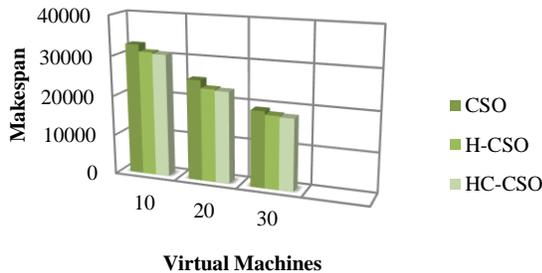


Fig. 6. Makespan Comparisons for Sipt_1000 Tasks.

For displaying the 10, 20, and 30 sets of VMs and makespan, the x-axis and y-axis are being used respectively. The Sipt dataset is used to represent sRNA-encoding genes

of several bacteria and it has been released by the HIB (Harvard International Bioinformatics) Centre. This dataset requires huge memory and a high computational CPU. For the processing of this dataset, the proposed algorithm HC-CSO again worked well as compared to standard CSO and the H-CSO because the proposed algorithm chooses the most appropriate VM instead of high or low power due to better properties of local as well as global searching and the HC-CSO also avoids unnecessary diversity. The proposed algorithm could not trap in local minima due to Crow Search Algorithm local searching property.

In the last, it is concluded that the HC-CSO algorithm is giving better results in the form of a better makespan for all scenarios in comparison to standard CSO, and H-CSO. For all the scenarios, HC-CSO works better than CSO and, H-CSO because of the better pre-processing of tasks by HEFT, a good balance between seeking and tracing modes due to the Crow Search algorithm. The HC-CSO algorithm along with the Crow Search Algorithm searches the VMs at a local and global level very carefully to optimize the results in the minimum number of iterations. The SMIW method restricts the Cats to go outside the search space.

Table IV is summarizing the evaluation of the processing cost for the execution of all scenarios designed for simulation.

TABLE IV. COST CONSUMPTIONS (IN INDIAN RUPEES)

Scenarios	VMs	CSO	H-CSO	HC-CSO
Scenario - 1 CyberShake_1000	10	484.53	459.43	447.89
	20	601.43	590.29	567.78
	30	703.47	668.13	640.25
Scenario - 2 Montage_1000	10	181.83	167.53	159.13
	20	191.19	184.19	173.19
	30	198.13	190.89	169.17
Scenario - 3 Inspiral_1000	10	3663.48	3629.55	3413.42
	20	3759.37	3685.87	3543.65
	30	5430.17	5217.58	5130.71
Scenario - 4 Sipt_1000	10	2811.93	2759.19	2645.13
	20	2973.87	2918.13	2703.29
	30	3353.29	3109.24	2999.43

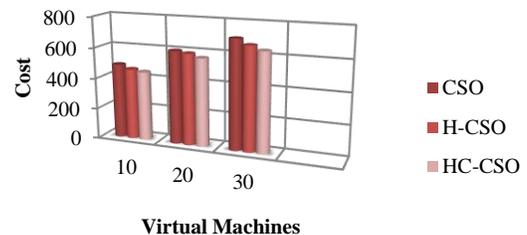


Fig. 7. Cost Comparisons for CyberShake_1000 Tasks.

For the first scenario having CybeShake_1000 dataset, a group of 10, 20, and 30 VMs along with processing cost are demonstrated at the x-axis and y-axis respectively in Fig. 7. The graph is expressing that the proposed HC-CSO algorithm is consuming minimal costs as compared to other algorithms. This is because the proposed algorithm has faster convergence and the positions of the workflows tasks on various VMs are updated smartly.

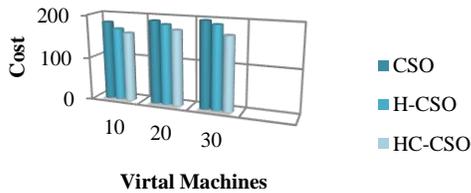


Fig. 8. Cost Comparison for Montage_1000 Tasks.

Fig. 8 is summarizing the processing cost consumption for the proposed algorithm HC-CSO, H-CSO, and standard CSO algorithms while working on 10, 20, and 30 VMs for the Montage_1000 dataset. The x-axis and y-axis of the graph are demonstrating the VMs and computing costs respectively. The proposed algorithm works better than H-CSO and CSO because the workflow tasks are migrated effectively for under-loaded virtual machines irrespective of their computing capacity.

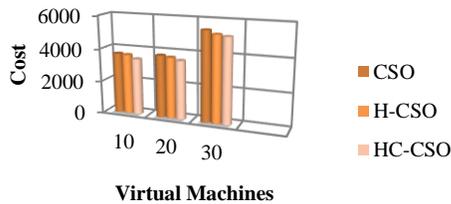


Fig. 9. Cost Comparison for Inspiral_1000 Tasks.

The processing cost results are being depicted in Fig. 9 for the Inspiral_1000 dataset. It can be seen VMs and Processing at the x-axis and y-axis respectively in the graph. The HC-CSO algorithm consumes less cost while executing the Inspiral_1000 dataset with all sets of VMs in comparison to other algorithms, this is because; the proposed algorithm makes an effective balance between seeking and tracing modes due to CSA integration. The VMs were picked up for execution of tasks irrespective of their MIPS, RAM, and bandwidth.

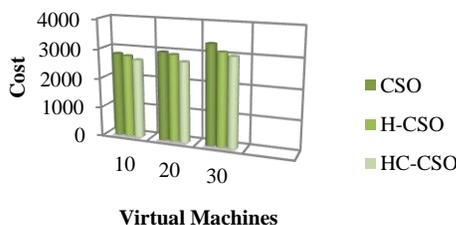


Fig. 10. Cost Comparisons for Sipt_1000 Tasks.

Fig. 10 is describing the results of processing cost obtained by the proposed HC-CSO, CSO, and H-CSO algorithms with various groups of virtual machines for the Sipt_1000 dataset. The VMs and processing cost can be seen at the x-axis and y-axis correspondingly in Fig. 10. Again, for the Sipt_1000 workflow; the proposed algorithm outperforms CSO and H-CSO in terms of computing cost. It is because the global and local searching properties of the proposed algorithm are balanced and the overloaded VMs loads are migrated to other VMs very smartly in the minimum time.

With these results, it can be specified that the proposed HC-CSO algorithm is better than other algorithms like CSO and H-CSO in respect of makespan and processing cost. The reason behind this is the good combination of global and local due to CSA. The Self-Motivated Inertia Weight factor integration overcomes the velocity outrange problem of Cats at tracing mode. The proposed algorithm chooses the virtual machines which are idle, under-loaded, or overloaded for workflow tasks migration among different VMs irrespective of their computing power, RAM, and Bandwidth.

VI. CONCLUSION

In the current technological era, cloud computing is one of the important emerging technologies used to store a large volume of data and other computing services in various science and technological fields. For computing facilities, various heuristic and meta-heuristic techniques have been developed. In this paper, an intelligent workflow scheduling algorithm named HC-CSO has been proposed to solve the workflow task scheduling problem. The proposed algorithm is a combination of H-CSO and the Crow Search algorithm. The H-CSO algorithm is an integration of HEFT and SMIW methods. The HEFT algorithm pre-processed the workflows tasks and initialized the proposed HC-CSO algorithm with these tasks. This process saves time and optimizes the results in a fewer number of iterations. The proposed HC-CSO algorithm didn't get trapped in local optima due to the good local searching capacity of the CSA. Velocity outranges cause to push the Cats outside the search space and affects the performance of the algorithm. The Self-Motivated Inertia Method overcame this problem.

The proposed HC-CSO algorithm outperformed CSO and H-CSO in terms of makespan and computing cost with all four scenarios having four scientific datasets CyberShake, Montage, Inspiral, Sipt, on a group of 10, 20, and 30 VMs. This is because the proposed algorithm chose the best VM among a group of VMs using its perfect global and local searching capacity for workflows task scheduling. The proposed approach is a generalized algorithm and will perform well with all types of scientific datasets despite a particular one.

In the future, the proposed algorithm can be tested at a wide scale to reduce makespan, cost and, other parameters in the cloud system and many other fields. The efficiency of the proposed algorithm can also be tested for independent tasks. A new technique can also be developed to enhance cloud performance.

REFERENCES

- [1] G. Danlami, S. I. Abdul, Z. Anazida, Z. Zalmiyah, and A. Ahmad, "Hybrid Cat Swarm Optimization and Simulated Annealing for Dynamic Task Scheduling on Cloud Computing Environment," *Journal of Information and Communication Technology*, vol. 17, no. 3. pp. 435-467, July 2018.
- [2] B. Saurabh, S. Santwana, and D. Madhabanda, "A Multi-Objective Cat Swarm Optimization Algorithm for Workflow Scheduling in Cloud Computing Environment," *International Journal of Soft Computing*, vol. 10, pp 17-45, 2015.
- [3] B. Saurabh, S. Santwana, and D. Madhabanda, "Workflow Scheduling in Cloud Computing Environment using Cat Swarm Optimization," *IEEE International Advance Computing Conference*, pp. 680-685, 2014.
- [4] G. Danlami, S. I. Abdul, Z. Anazida, Z. Zalmiyah, and A. Ahmad, "Cloud Scalable Multi-Objective Task Scheduling Algorithm for Cloud Computing using Cat Swarm Optimization and Simulated Annealing," *IEEE International Conference on Information Technology*, pp. 1007-1012, 2017.
- [5] D. Negar, and J. N. Nima, "A Hybrid Particle Swarm Optimization and Hill Climbing Algorithm for Task Scheduling in the Cloud Environments," *ICT Express*, vol. 4, pp. 199-202, 2017.
- [6] S. A. Thanna, A. S. Ashraf, and D. Yassine, "Binary PSO for Load Balancing Task Scheduling in Cloud Environment," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 5, pp. 255-264, 2018.
- [7] S. Narinder, and S. B. Singh, "Hybrid Algorithm of Particle Swarm Optimization and Grey Wolf Optimizer for Improving Convergence Performance," *Journal of Applied Mathematics*, Hindawi, vol. 2017, pp. 1-15, 2017.
- [8] S. Heba, N. Heba, S. Walla, and M. H. Hany, "IPSO Task Scheduling Algorithm for Large Scale Data in Cloud Computing Environment," *IEEE Access*, vol. 7, pp. 5412-5420, 2019.
- [9] P. Suraj, W. Linlin, M. G. Siddeswara, and B. Rajkumar, "A Particale Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments," *IEEE International Conference on Advanced Information Networking and Applications*, pp. 400-407, 2010.
- [10] E. Fatemeh, and M. B. Seyed, "A PSO Based Task Scheduling Algorithm Improved using a Load-Balancing Technique for the Cloud Computing Environment," *Concurrency and Computation Practice and Experience*, Wiley, Special Issue, pp. 1-16, Dec. 2017.
- [11] Z. Zhou, C. Jian, H. Zhigang, Y. Junyang, and L. Fangmin, "A Modified PSO Algorithm for Task Scheduling Optimization in Cloud Computing," *Concurrency and Computation Practice and Experience*, Wiley, Special Issue, pp. 1-11, Sept. 2018.
- [12] M. Neha, and S. Gaurav, "Modified Particale Swarm Optimization based upon Task Categorization in Cloud Environment," *International Journal of Engineering and Advanced Technology*, vol. 8, pp. 67-72, 2019.
- [13] A. T. Medhat, E. Ashraf, E. K. Arabi, and A. T. Fawzy, "Cloud Task Scheduling Based on Ant Colony Optimization," *International Conference on Computer Engineering and Systems*, IEEE, pp. 64-69, 2013.
- [14] S. Khurana, and R. K. Singh, "Workflow Scheduling and Reliability Improvement by Hybrid Intelligence Optimization Approach with Task Ranking," *EAI Endorsed Transactions on Scalable Information Systems*, Nov. 2019.
- [15] M. S. Raja, P. Sanchita, and K. Abhishek, "Task Scheduling in Cloud Computing: Review," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 6, pp. 7940-7944, 2014.
- [16] L. Zhizhong, Q. Jingxuan, P. Weiping, and C. Hao, "Effective Task Scheduling in Cloud Computing Based on Improved Social Learning Optimization Algorithm," *International Journal of Online and Biomedical Engineering*, vol. 13, no. 6, pp. 4-21, 2017.
- [17] A. S. M. Mohammed, G. Radhamani, A. G. H. Mohamed, and H. H. Syed, "Adaptive Cost-Based Task Scheduling in Cloud Environment," *Scientific Programming*, Hindawi, vol. 2016, pp. 1-9, 2016.
- [18] A. Ali, and A. O. Fatma, "Task Scheduling Using PSO Algorithm in Cloud Computing Environments," *International Journal of Grid Distribution Computing*, vol. 18, no. 5, pp. 245-256, 2015.
- [19] P. Guang, and W. Katinka, "Efficient Task Scheduling in Cloud Computing using an Improved Particle Swarm Optimization Algorithm," *International Conference on Cloud Computing and Service*, pp. 58-67, 2019.
- [20] C. Hongyan, L. Xiaofei, Y. Tao, Z. Honggang, F. Yajun, and X. Zongguo, "Cloud Service Scheduling Algorithm Research and Optimization," *Security and Communication Networks*, vol. 2017, pp. 1-7, 2017.
- [21] M. A. Aram, A. R. Tarik, and A. M. S. Soran, "Cat Swarm optimization Algorithm: A Survey and Performance Evaluation," *Computational Intelligence and Neuroscience*, vol. 2020, pp. 1-17, 2020.
- [22] A. Alireja, "A Novel Metaheuristic Method for Solving Constrained Engineering Optimization Problems: Crow Search Algorithm," *Computer and Structures*, vol. 169, pp. 1-12, 2016.
- [23] K. Mala, and S. Sarabjeet, "A Review of Metaheuristic Scheduling Techniques in Cloud Computing," *Egyptian Informatics Journal*, vol. 16, pp. 275-295, 2015.

Improving Imbalanced Data Classification in Auto Insurance by the Data Level Approaches

Mohamed Hanafy¹, Ruixing Ming²

School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou, China^{1,2}
Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Assuit University, Assut, Egypt¹

Abstract—Predicting the frequency of insurance claims has become a significant challenge due to the imbalanced datasets since the number of occurring claims is usually significantly lower than the number of non-occurring claims. As a result, classification models tend to have a limited ability to predict the occurrence of claims. So, in this paper, we'll use various data level approaches to try to solve the imbalanced data problem in the insurance industry. We developed 32 machine learning models for predicting insurance claims occurrence {(under-sampling, over-sampling, the combination of over-and under-sampling (hybrid), and SMOTE) × (three Decision tree models, three boosting models, and two bagging models) = 32}, and we compared the models' accuracies, sensitivities, and specificities to comprehend the prediction performance of the built models. The dataset contains 81628 claims, each of which is a car insurance claim. There were 5714 claims that occurred and 75914 claims that didn't occur. According to the findings, the AdaBoost classifier with oversampling and the hybrid method had the most accurate predictions, with a sensitivity of 92.94%, a specificity of 99.82%, and an accuracy of 99.4%. And with a sensitivity of 92.48%, a specificity of 99.63%, and an accuracy of 99.1%, respectively. This paper confirmed that when analyzing imbalanced data, the AdaBoost classifier, whether using oversampling or the hybrid process, could generate more accurate models than other boosting models, Decision tree models, and bagging models.

Keywords—Machine learning; classification; insurance; imbalanced data problem; resampling methods

I. INTRODUCTION

The use of machine learning techniques and the transformation of the insurance market into a new level of digital applications are the insurance industry's current challenges. There are two types of insurance: life insurance and non-life insurance. Non-life insurance, specifically auto insurance, is the subject of this study.

A variety of variables influence automobile insurance pricing [1]. And these factors would affect the cost of a client's insurance policy. Credit history is an example of one of these factors; studies indicate that people with poor credit are more likely to file claims, commit fraud, or miss payments, putting the insurance company in a financial bind. Another factor to consider is the client's location; studies have shown that densely populated areas with heavy traffic have a higher rate of accidents, resulting in a higher number of claims. This would result in a significant rise in the customer's insurance premium. However, it is unjust for a good client to pay more simply because of where they live; this creates a problem for the

consumer because if the insurance premium is raised, he will be unable to afford it, resulting in the insurance provider losing these clients. So, necessitating the creation of an appropriate method for evaluating the risk each client poses to insurers.

As a result, insurance rates should be adjusted based on a client's skill and other personal information, making car insurance more accessible to consumers. Where insurance companies should customize a custom premium for each customer because this will help the insurers to adjust to any situation and manage any loss. Since It would be unreasonable to expect a client with a good driving record to pay the same insurance premium as a client with a poor driving record; as a result, the model should classify which clients are unlikely to file claims, lower their insurance costs, and raise insurance costs for those who are likely to file claims.

Data imbalanced problem are more likely to arise in the case of insurance data since the number of occurring claims is usually significantly lower than the number of non-occurring claims. And one of the major problems with machine learning techniques is that they are affected by the data set's unequal binary class distribution. In other words, when the data is unbalanced, certain machine learning techniques will simply ignore the small class and allocate the majority of the unseen cases to the common class, resulting in high overall model accuracy. Nonetheless, the performance of the prediction models for the small class will be substantially reduced. To solve this problem, we will use resampling techniques, such as Over-sampling, under-sampling, the combination of over-and under-sampling (hybrid), and the synthetic minority over-sampling technique (SMOTE), to improve the classification efficiency for imbalanced data.

We used a large dataset given by a large automotive company based in Egypt. In this study, we apply data-level approaches that could reduce overfitting caused by data imbalance. And we built 32 machine learning models for predicting the occurrence of auto insurance claims ((under-sampling, over-sampling, hybrid of over-and under-sampling, and SMOTE) × (three Decision tree models, three boosting models, and two bagging models) = 32). And we compared the models' accuracy, sensitivities, and specificities to better understand the built models' prediction efficiency.

The following is the structure of this paper: Section II presents the previous studies. Section III explain the data collection, machine learning models, and data-level approaches. Section IV compared the results of the built thirty-

two prediction models. Section V presents concludes. Section VI presents the future work.

II. RELATED WORK

Over the last decade, many researchers have used machine learning algorithm to forecast the occurrence of auto insurance claims. And while machine learning models are efficient at predicting. But when the data is unbalanced, machine learning techniques will simply ignore the small class and allocate the majority of the cases to the common class, resulting in high overall model accuracy. Nonetheless, the performance of the prediction models for the small class will be substantially reduced. The following studies show a lack of using the resampling methods to solve the unbalanced data problem except for the study of [1] that only used the oversampling method.

In the study of machine learning approaches for auto insurance big data [1], they built eight classifiers to predict the occurrence of the claims using big insurance data, including XGBoost, J48, RF, C5.0, CART, K-NN, logistic regression, and naïve Bayes algorithms, and they handled the heavy imbalanced data using the over-sampler method. The RF model performed the best among the eight models. And [2] used two competing methods, XGBoost, and logistic regression, to predict the frequency of motor insurance claims. This study shows that the XGBoost model is slightly better than logistic regression. Furthermore, a model for predicting insurance claims was developed by [3]; they built four classifiers to predict the claims, including XGBoost, J48, ANN, and naïve Bayes algorithms. The XGBoost model performed the best among the four models. Another example of a similar and satisfactory solution to the same problem is the thesis "Research on Probability-based Learning Application on Car Insurance Data" by [4], which used only a Bayesian network to classify either a claim or no claim. And the [5] research also aims to look at data mining techniques for creating a predictive model for auto insurance claim prediction. And they compared three ML methods for predicting claims. Their findings showed that the best predictor was the neural networks.

In summary, despite the relevance of the imbalanced data problem in the insurance industry, there is a lack of comprehensive comparison among the prominent resampling approaches as a strategy to deal with it. The purpose of this study is to investigate the impact of the unbalanced data problem on the performance of machine learning models. This paper solves the unbalanced data problem with several resampling approaches and compares them while utilizing various machine learning classifiers to fill in the gaps in the literature.

In comparison to prior studies, the following are the novel innovations and vital procedures of this study:

- Applying and comparing several resampling algorithms, including the Random Over Sampler, Random Under Sampler, SMOTE, and the hybrid.
- Applying several machine learning models, such as three Decision tree models, three boosting models, and two

bagging models, to compare the performance of resampling methods.

- Measuring the effectiveness of implemented machine learning models utilizing various performance measures such as Accuracy, Sensitivity, and Specificity.
- Demonstrating how resampling affects the performance of classifiers.

III. METHODS

A. Data Collection

Our dataset is progressive record keeping, which usually updates over time to reflect the updated status of a particular customer, which means that provided data is the snapshot of some particular date, where all the records show the updated status of each customer. This dataset is updated on changes in circumstances of the customer, such as marital status, age, etc. The sample auto insurance claims dataset was collected between 2014 and 2018.

The data used in this study is real-life data obtained from an Egyptian car insurance firm; we end up with 81628 claims in the dataset, each of which is a car insurance claim. In total, there are 5714 claims that occurred, and 75914 non- occurred claims, suggesting that the data is heavily unbalanced. And as we mention, the performance of classification algorithms is greatly affected by imbalanced data. So, we apply four resampling techniques to solve the problem of data imbalance. Each claim comprises 17 features. Table I provides Attributes of the data.

B. Data Preprocessing

Numeric values are allocated to categorical variables. For example, instead of male or female as the gender of the insured, the "Male" component would be (1), and "female" would be (0). After this phase, we can apply our data to all machine learning models.

C. Machine-Learning for Auto Insurance Claims Occurrence:

1) *Decision trees*: A decision tree D , in more formal terms, is made up of two kinds of nodes:

- A leaf node that represents the response variable's given class/region.
- A decision node that defines a test on a single attribute (predictor variable) with one branch and subtree for each test outcome.

Using a recursive divide and conquer method, a decision tree can be used to classify an observation by beginning at the top decision node (called the root node) and going down through the other decision nodes until a leaf is encountered.

The following models are the most well-known methodology for constructing decision trees [6,7,8].

- a) CART algorithm.
- b) C5.0 algorithm.
- c) C4.5 algorithm.

TABLE I. ATTRIBUTES OF THE DATA

No.	Description	Descriptive statistics
1	Representing the age of the insured.	Min. :18.00 Mean :44.84 Max. :81.00
2	Represent the Value of the car in Egyptian pounds.	Min. :75000 Mean :304003.5 Max. :6610100
3	Represent the car age in years.	Min. :0.000 Mean :10.298 Max. :28.000
4	Represent the car type.	Such as Mercedes, BMW, and Toyota, etc.
5	Represent the car use.	Commercial:30148 Private: 51480
6	Represent the insured education level.	High School :12244 Bachelors :32856 Masters :16325 PhD :7347 Other :12856
7	Represent the annual Income of the insured In Egyptian pounds.	Min. :0 Mean :61572 Max. :367030
8	Represent the number of dependants for the insured.	Min. :0.0000 Mean :3.000 Max. :5.0000
9	Represent the Marital status	Married:48977 not married:32651
10	Represent the insured's occupation.	The job of the insured
11	Represent the claim frequency.	Min. :0.0000 Mean :0.8007 Max. :5.0000
12	Represent if the insured license was revoked before.	No:71880 Yes:9748
13	Represent The insured gender.	MALE:38365 FAMLE:43263
14	Represent the Distance to work in km	Min. :5.00 Mean :33.42 Max. :142.00
15	Represent where the insured live urban vs. rural area.	Urban:66118 Rural:15510
16	Represent the number of Years on job for the insured(yoj).	Min. :0.00 Mean: 10.47 Max. :23.00
17	Claim occurred or not. (the target variable)	non occurred: 75914 occurred: 5714

2) *Bagging trees*: Bagging, or bootstrap aggregation, is an ensemble meta-algorithm. This algorithm increases the model's consistency and accuracy while also reducing overfitting. In classification, it weights the output to ensemble into a single output. Leo Breiman suggested bagging in 1996 [9] as a way to improve classification results.

The following are the two most common bagging algorithms:

- a) Bagged CART.
- b) Random forest.

3) *Boosting*: Boosting is an ensemble technique that, like training, creates several individual models sequentially. Each new model attempts to correct the errors of the previous group of models. Boosting, like bagging, can be used to improve any supervised machine learning algorithm. Boosting, on the other hand, is most effective when weak learners are used as sub-models. As a result, boosting has historically been used on shallow decision trees. By shallow, I mean a decision tree with a limited number of levels of depth or a single split. Boosting's aim is to bring together a group of weak learners to form a strong ensemble learner [10,11,12].

The following are the two most common bagging algorithms:

- a) AdaBoost.
- b) XGBoost.
- c) Stochastic Gradient Boosting.

D. A Data-level Approach and Imbalanced Data

The imbalanced data problem exists in many datasets; as a result, classifiers models are biased against the minority class and are unable to predict it accurately [13]. In contrast, most machine learning models perform better when applied with balanced datasets [14,15,16,17].

Analysis of our database shows that they are extremely imbalanced, and the two forms of insurance claims are not balanced, with 93% (n=75914) of the auto insurance claims occurred, and those non-occurred were 7% (n=5714). As a consequence, the imbalanced data problem must be addressed. Many techniques have been developed to resolve the problem of unbalanced data. One of the most successful approaches for addressing unbalanced data is using a sampling-based approach, either Random Over Sampler [18], Random Under Sampler [19], and SMOTE [20].

We will use the ROSE and the ovun.sample function incorporates more conventional class inequality solutions, such as over-sampling the minority class, under-sampling the majority class, or a combination of over-and under-sampling. And also, we will use the DMwR package to apply SMOTE as a resampling method.

1) *Over-sampling technique*: This technique increases the weight of the minority class. It's important to note that the technique of over-sampling is typically used more than other methods.

a) *Random over sampler*: Random Over-Sampling is a technique based on bootstrap that supports the binary classification task in the presence of unbalanced classes by generating synthetic examples from a conditional density estimation of the two classes [21]. It handles both continuous and categorical data by randomly replicating samples from the minor class [22]. As a result of this process, the dataset grows in size. The argument is that no new samples are generated by a random over-sampler, and the variety of samples remains constant. Since the sample size grows, the oversampling technique takes longer to construct a model and can cause

overfitting because it duplicates samples from a minor class. [23,24].

b) *SMOTE*: SMOTE is similar to random oversampling. However, it does not regenerate the same instance. It creates a new instance by appropriately combining existing instances, thus making it possible to avoid the disadvantage of overfitting to a certain degree. Moreover, SMOTE is an oversampling technique that produces new minority samples by combining two minorities and one of their K nearest neighbours [25]. This approach is a statistical technique for creating new instances to increase the number of minority samples in a dataset. This algorithm takes characteristic features for the target class and its closest neighbours, then produces new samples by combining the characteristics of a specific case with those of its neighbours.

2) *Random under sampler*: Under-sampling is one of the simplest techniques to dealing with the problem of unbalanced data. It balances the majority and minority classes. The process of under-sampling includes arbitrarily deleting examples from the majority class in the training dataset, referred to as random under-sampling [26]. By reducing the amount of data, under-sampling can save time when building a model, but it comes at the cost of losing information [27,28].

3) *Hybrid methods*: There are several benefits and drawbacks of over-sampling and under-sampling. Combining these two strategies will add the strengths of these two methods to a new method [26].

E. Development of Prediction Models and Prediction Performance Evaluation

This study built thirty-two prediction models ((under-sampling, oversampling, the combination of over-and under-sampling (hybrid), and SMOTE) × (three Decision tree models, three boosting models, and two bagging models) = 32). The accuracy, sensitivity, and specificity of each model are used to compare the prediction performance of the established models.

This study randomly divided the data into a training dataset and a test dataset at a ratio of 7:3. The hyperparameters are tuned using a 10-fold cross-validation that was performed only on the training dataset to get the best performance for each machine learning model. And the test dataset was used to evaluate the prediction performance. The Data-level Approaches must only be applied to the training set while the test set still unbalanced. We used R version 4.0.2 to conduct all analyses.

F. Evaluation Methods

Evaluation methods are essential in comparing and selecting the best model because they are assessing the efficiency of classifiers [1].

Table II shows the Evaluation methods used in this study. Where TP is the number of true positives, the number of false positives is FP, the number of true negatives is TN, and the number of false negatives is FN.

TABLE II. EVALUATION METHODS

Accuracy	Referred to the overall correctly prediction	$\frac{(TP + TN)}{(TP + FP + TN + FN)}$
Sensitivity	Referred to the correct rate of predicting the occur claims	$\frac{TP}{(TP + FN)}$
Specificity	Referred to the correct rate of predicting the non-occur claims.	$\frac{TN}{(FP + TN)}$

IV. RESULTS AND DISCUSSION

To show the difference between the ability of the machine learning classifiers to predict the insurance claims occurrence before and after handling the unbalanced data problem, we compared all applied models on the unbalanced data and also on the balanced data created by different resampling techniques. We measure the performance of models on the testing data using accuracy, sensitivity, and specificity.

A. Comparing the Performance of the Built Machine Learning Models

Tables III, IV, and V show the accuracy, sensitivity, and specificity respectively of the thirty-two prediction models, as well as Fig. 1.

Table III shows the Accuracy of each machine learning technique on unbalanced data and balanced datasets generated by four different resampling models. And we must consider that only if the data is balanced will Accuracy be a valuable metric, while when the data is unbalanced, the Accuracy would be meaningless. Because when the data is unbalanced, most machine learning techniques will simply ignore the small class and allocate most of the unseen cases to the majority class, resulting in high overall model accuracy and high specificity, while the sensitivity of the models will substantially be reduced. The AdaBoost classifier achieved 99.4 % accuracy by using the oversampling, which is the highest of all other classifiers. And the lowest accuracy outcome goes to the C5.0 with 74.84% by using the under-sampling.

Table IV refers to the Sensitivity of the machine learning models. Sensitivity relates to the ability to predict the occurrence of claims. We can note that the Sensitivity for all ML models with the unbalanced data is lower than the Sensitivity for balanced data created by different resampling methods because the occurred claims represent a small class with only 7% in our data. Therefore, before solving the unbalanced data problem, machine learning techniques will simply ignore the small class (occurred claims). Thus, resulting in very low Sensitivity in the case of the unbalanced data. While the Sensitivity is improved after applied the resampling methods. This refers to the effectiveness of using the resampling methods to handle the unbalanced data problem in the insurance industry. And the highest Sensitivity goes to the AdaBoost classifier with 92.94% using the oversampling, while the lowest one goes to the AdaBoost model with 0.46% using the unbalanced data.

TABLE III. THE ACCURACY OF THE DEVELOPED MODELS

Models	Unbalanced	OVER	UNDER	HYBRID	SMOTE
Decision trees models					
C5.0	0.9393	0.9426	0.7484	0.9822	0.8441
C4.5	0.9432	0.96	0.794	0.9544	0.8117
CART	0.9441	0.7791	0.7586	0.8076	0.8316
Bagging models					
Bagged CART	0.939	0.928	0.7487	0.9044	0.8266
Random forest	0.94	0.978	0.7807	0.9806	0.8533
Boosting models					
AdaBoost	0.9385	0.994	0.7516	0.9919	0.8689
XGBoost	0.9386	0.8786	0.7951	0.8715	0.8641
Stochastic Gradient Boosting	0.9396	0.7865	0.7648	0.796	0.8461

TABLE IV. THE SENSITIVITY OF THE DEVELOPED MODELS

Models	unbalanced	OVER	UNDER	HYBRID	SMOTE
Decision trees models					
C5.0	0.1422	0.8415	0.7768	0.8878	0.8098
C4.5	0.1268	0.713	0.6333	0.8633	0.8064
CART	0.1463	0.7098	0.7024	0.6707	0.8
Bagging models					
Bagged CART	0.161	0.1904	0.7309	0.3873	0.6499
Random forest	0.0456	0.8223	0.7244	0.8428	0.8064
Boosting models					
AdaBoost	0.0046	0.9294	0.7175	0.9248	0.8109
XGBoost	0.1139	0.7744	0.7153	0.795	0.8337
Stochastic Gradient Boosting	0.1162	0.7358	0.7585	0.7289	0.7722

Table V refer to the Specificity of the machine learning models. Specificity refers to the ability to predict non-occurred claims. We can note that the Specificity for all models with unbalanced data is highest than the Specificity for balanced data created by different resampling methods because the non-

occurred claims represent the majority class with 93% in our data. Therefore, before solving the unbalanced data problem, machine learning techniques will allocate the most unseen cases to the majority class (non-occurred claims). This is resulting in very high overall model Specificity in the case of the unbalanced data. But our objective is to detect MINOR class more accurately than MAJOR class; therefore, we are interested in Sensitivity more than Specificity. SO, we need to resample the dataset to force algorithms to identify both classes with equal importance. And the highest Specificity in the dataset belongs to AdaBoost classifiers with 99.93% using the unbalanced data, and the lowest one goes to the C5.0 model with 74.65 % using the under-sampling.

Last but not least, from Tables III, IV, V, and Fig. 1, we can conclude that using the resampling methods is very effective for handle the unbalanced data problem in the insurance industry, because the best results are achieved after applied the data-level approaches.

And the best models are AdaBoost with the over and hybrid methods, then the C5.0 model with the hybrid method, and then the random forest model with the hybrid method. Where AdaBoost with the over and hybrid methods achieved a sensitivity of 92.94%, a specificity of 99.82%, and an accuracy of 99.4%. And a sensitivity of 92.48%, a specificity of 99.63%, and an accuracy of 99.19%, respectively. And the C5.0 model with the hybrid method has a sensitivity of 88.78%, a specificity of 98.79%, and an accuracy of 98.22%. Then there's the random forest model with the hybrid method, which has a sensitivity of 84.28%, a specificity of 98.96%, and an accuracy of 98.06%.

TABLE V. THE SPECIFICITY OF THE DEVELOPED MODELS

	unbalanced	OVER	UNDER	HYBRID	SMOTE
Decision trees models					
C5.0	0.9935	0.9487	0.7465	0.9879	0.8477
C4.5	0.9924	0.9761	0.8045	0.9604	0.812
CART	0.9922	0.7832	0.7619	0.8159	0.8335
Bagging models					
Bagged CART	0.9859	0.9781	0.7499	0.9396	0.8386
Random forest	0.9982	0.9881	0.7843	0.9896	0.8564
Boosting models					
AdaBoost	0.9993	0.9982	0.7538	0.9963	0.8726
XGBoost	0.9923	0.8854	0.8003	0.8913	0.8654
Stochastic Gradient Boosting	0.9932	0.7898	0.7652	0.8003	0.8509

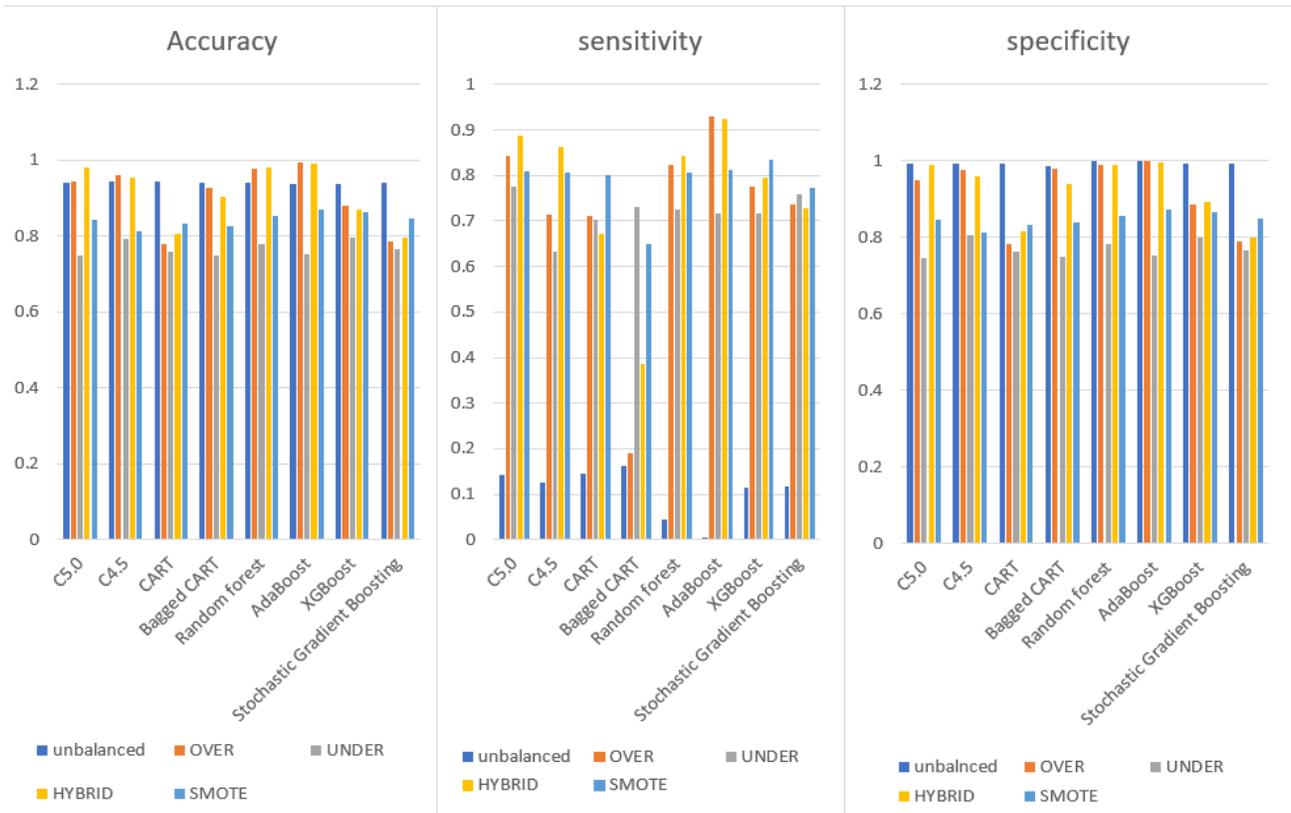


Fig. 1. Comparison between the Developed Models based on the Accuracy, Sensitivity and Specificity.

B. Variables Importance for Auto Insurance Claims Classification in the AdaBoost with the Oversampling

The importance of the variables of the final model (AdaBoost with the oversampling) is presented in Fig. 2.

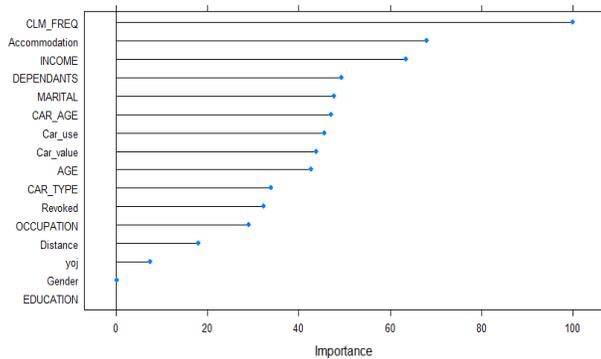


Fig. 2. The Importance of Variables in the AdaBoost Classifier with Oversampling-based Claims Occurrence Prediction Model.

V. CONCLUSION

This study specifically established models for improving the classification efficiency of imbalanced data by using oversampling, under-sampling, the combination of over-and under-sampling, and SMOTE as resampling approaches. ((under-sampling, oversampling, a combination of over-and under-sampling (hybrid), and SMOTE) × (three Decision tree models, three boosting models, and two bagging models) =32)

for predicting auto insurance claims occurrence. According to the findings of this analysis, the AdaBoost model with over and hybrid could generate more accurate models than other boosting models, Decision tree models, and bagging models, then the C5.0 model with the hybrid method, and then the random forest model with the hybrid method.

VI. FUTURE WORK

Further research is required to compare the accuracy using various datasets from various fields to prove the prediction efficiency of an AdaBoost classifier with resampling methods to solve the imbalanced data problem. And Future work may be done in the following directions: Using hybrid machine learning classifiers to improve comparison and performance. And also, use different feature selection approaches to enhance model results and gain a deeper understanding of the important features.

REFERENCES

- [1] Hanafy, Mohamed, and Ruixing Ming. 2021. Machine learning approaches for auto insurance big data. *Risks* 9: 42.
- [2] Pesantez-Narvaez, Jessica, Montserrat Guillen, and Manuela Alcañiz. 2019. Predicting motor insurance claims using telematics data—XGBoost versus logistic regression. *Risks* 7: 70.
- [3] ABDELHADI, SHADY, KHALED ELBAHNASY, and MOHAMED ABDELSALAM. 2020. A PROPOSED MACHINE MODEL TO PREDICT AUTO INSURANCE CLAIMS USING LEARNING TECHNIQUES. *Journal of Theoretical and Applied Information Technology* 98.
- [4] Jing, Longhao, Wenjing Zhao, Karthik Sharma, and Runhua Feng. Year. Research on Probability-based Learning Application on Car Insurance

- Data. Paper presented at the 2017 4th International Conference on Machinery, Materials and Computer (MACMC 2017).
- [5] Weerasinghe, KPMLP, and MC Wijegunasekara. 2016. A comparative study of data mining algorithms in the prediction of auto insurance claims. *European International Journal of Science and Technology* 5: 47-54.
- [6] Breiman, L. (1984). *Classification and Regression Trees*. Routledge.
- [7] Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.
- [8] Quinlan, J. R. et al. (1996). Bagging, boosting, and c4. 5. In *AAAI/IAAI*, Vol. 1, pages 725–730.
- [9] Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2):123–140.
- [10] Kégl, Balázs. 2013. The return of AdaBoost. MH: multi-class Hamming trees. *arXiv preprint arXiv:1312.6086*.
- [11] Chen, Tianqi, and Carlos Guestrin. Year. Xgboost: A scalable tree boosting system. Paper presented at the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining.
- [12] Friedman, Jerome H. 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics*: 1189-232.
- [13] Kotsiantis, Sotiris, Dimitris Kanellopoulos, and Panayiotis Pintelas. 2006. Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering* 30: 25-36.
- [14] Hussain, Lal, Kashif Javed Lone, Intiaz Ahmed Awan, Adeel Ahmed Abbasi, and Jawad-ur-Rehman Pirzada. 2020. Detecting congestive heart failure by extracting multimodal features with synthetic minority oversampling technique (SMOTE) for imbalanced data using robust machine learning techniques. *Waves in Random and Complex Media*: 1-24.
- [15] H. Byeon, Development of a physical impairment prediction model for Korean elderly people using synthetic minority over-sampling technique and XGBoost. *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, pp. 36-41, 2021.
- [16] R. Mohammadi, R. Javidan, M. Keshtgari, and N. Rikhtegar, SMOTE: an intelligent SDN-based multi-objective traffic engineering technique for telesurgery. *IETE Journal of Research*, vol. 1-11, 2021.
- [17] H. Byeon, Predicting the depression of the South Korean elderly using SMOTE and an imbalanced binary dataset. *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, pp. 74-79, 2021.
- [18] L. Abdi, and S. Hashemi, To combat multi-class imbalanced problems by means of over-sampling techniques. *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 1, 238-251, 2015.
- [19] S. J. Yen, and Y. S. Lee, Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications*, vol. 36, no. 3, pp. 5718-5727, 2009.
- [20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.
- [21] Menardi, Giovanna, and Nicola Torelli. 2014. Training and assessing classification rules with imbalanced data. *Data mining and knowledge discovery* 28: 92-122.
- [22] M. Wang, X. Yao, and Y. Chen, An imbalanced-data processing algorithm for the prediction of heart attack in stroke patients. *IEEE Access*, vol. 9, pp. 25394-25404, 2021.
- [23] H. Byeon, Predicting the depression of the South Korean elderly using SMOTE and an imbalanced binary dataset. *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, pp. 74-79, 2021.
- [24] M. T. García-Ordás, C. Benavides, J. A. Benítez-Andrades, H. AlaizMoretón, and I. García-Rodríguez, Diabetes detection using deep learning techniques with oversampling and feature augmentation. *Computer Methods and Programs in Biomedicine*, vol. 202, pp. 105968, 2021.
- [25] Bunkhumpornpat, Chumphol, Krung Sinapiromsaran, and Chidchanok Lursinsap. Year. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. Paper presented at the Pacific-Asia conference on knowledge discovery and data mining.
- [26] Lunardon, Nicola, Giovanna Menardi, and Nicola Torelli. 2014. ROSE: A Package for Binary Imbalanced Learning. *R journal* 6.
- [27] H. Byeon, Development of a physical impairment prediction model for Korean elderly people using synthetic minority over-sampling technique and XGBoost. *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, pp. 36-41, 2021.
- [28] H. Guan, Y. Zhang, M. Xian, H. D. Cheng, and X. Tang, SMOTEWENN: Solving class imbalance and small sample problems by oversampling and distance scaling. *Applied Intelligence*, vol. 51, no. 3, pp. 1394-1409, 2021.

A New Feature Filtering Approach by Integrating IG and T-Test Evaluation Metrics for Text Classification

Abubakar Ado¹, Mustafa Mat Deris²
Noor Azah Samsudin³

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Malaysia

Aliyu Ahmed⁴

Department of Computer Science
Bauchi State University
Gadau, Nigeria

Abstract—High dimensionality is one of the main issues associated with text classification, such as selecting the most discrepant features subset for classifier's effective utilization is a difficult task. This significant preprocessing stage of selecting the relevant features is often called feature selection or feature filtering. Eliminating the non-relevant and noise features from the original feature set will drastically reduce the size of the feature set and the time complexity of the classification models and also improve or maintain their performance. Most of the existing filtering method produced a subset with relatively high number of features without much significant impact on running time, or produced subset with lesser number of features but results in performance degradation. In this paper, we proposed a new bi-strategy filtering approach that integrates Information Gain with t-test that selects a subset of informative features by considering both the score and ranking of respective features. Our approach considers the results' disparity produced by the benchmark metrics used in order to maximized and lessen their advantage and disadvantage. The approach set a new threshold parameter by computing V-score of the features with minimum scores present in both the two subsets and further refined the selected features. Hence, it reduces the size of the features subset without losing much informative features. Experiment results conducted on three different text datasets have shown that the proposed method is able to select features that are highly discrepant and at the same time achieves a significant improvement in terms of classification accuracy and F-score at the cost of a minimum running time.

Keywords—Dimensional reduction; feature filtering; feature selection; t-test; information gain; V-score

I. INTRODUCTION

In this emerging era of computing and internet technology, especially the emerging of social media, text analytic becomes more cumbersome[1]. As a result, both the size of features and instances of a textual dataset has been increasing rapidly. The increasing size of the text data results in diverse research problems to text analytic tools, such as machine learning. Text classification is one of the pronounce problem associated with text analytics [2][3], and currently is becoming one of the most vital research direction in the field of machine learning.

Text classification or documents classification is the problem of assigning unlabeled text instances to one or more predefined labelled classes or categories [4] [5][6][7]. Text classification has been utilized in various application domains [8], e.g. spam filtering [9], Sentiment Analysis [10], Natural

Language Processing [11][12] Information Retrieval, Text Mining and so on.

One of the most crucial steps of the preprocessing of text data is the presentation of text documents into vector space via Bag_of_Word (BOW) [13][14][15]. The final product of this task is associated with two main issues, a vast number of features representation, and the presence of irrelevant and noisy features which general termed high dimensionality[15]. These issues can cause a lot of problems for the Text classification task, which is known to be intrinsically high dimensional [4] [5]. Classification in a situation that involves high number of features or high-dimensional space can become infeasible or very difficult due to computational complexity expensiveness [16][17]. However, feature reduction approach is considered as a dimensional reduction problem. The huge features generated introduces the so-called "dimensionality curse" with thousands of features that increase the computational complexity of a classifier [18][19][20][21][22]. Curse of dimensionality is a popular known problem for machine learning models [23]. When it arises in text classification, it seriously worsens the performance of the classifier in terms of classification accuracy and running time [5][24].

The main goal of dimensionality reduction is to reduce the number of features without worsening the performance of the classifier [14] [19]. As the key way to overcome this problem, feature selection (FS) technique can be applied to filter out irrelevant, redundant and noisy features and selects the most informative subset of features from the original features set [1][19]. This task will aggressively reduce the original vectors space representation of features into lower-dimensional vector representation [25][26][27]. Moreover, the properties of the informative features in the original feature set would be unaltered in the processes of dimensionality reduction. Feature selection (FS) approach ranks the original features according to some criterion evaluation (scores) and selects the top-ranked features to form an informative subset [27], which retains a good degree of discriminating capability in separating documents of various categories [28][29]. In contrast to the feature selection, feature extraction approach transforms the text documents on to a new lower-dimensional space from their original high dimensional feature instead of selecting a features subset from the original features set [30][31] [15].

Generally, feature selection methods [32] are broadly grouped into filter methods, wrapper methods[33], and embedded methods [34][35][27]. Filter methods [36] are independent that they do not interact with classifier when constructing an informative features subset. They rely on metrics for evaluating and ranking the importance of a feature prior to the classification. The methods can attain quick feature sorting to effectively filter out a high number of non-relevant or noise features [27]. They select features subset by considering the usefulness of a feature according to evaluation metrics [35][28][27][37]. Filter methods usually have good computational efficiency but sacrifice classification accuracy to some extent. Information Gain [38], Chi-Square [39], Fisher Score [40], ReliefF [41], *t*-test [4] are among the few filter based methods. Wrapper methods are dependent on classifiers that they frequently interact with the classification algorithm in order to construct a subset of informative features [13][35][27]. They evaluate a particular feature subset by training and testing a given classifier. The methods are tailored to a particular classifier [42]. These methods have bad computational efficiency but result in high classification accuracy, and they are not usually favoured in text classification task [43]. Heuristic Search Algorithms (HSA) and Sequential Selection Algorithms (SSA) [44][45][46] are common examples of classical wrapper methods. Embedded Methods integrate classifiers with feature selection technique during the training phase and optimally search feature subset by designing an optimization function [35][44][47]. Like wrapper methods, embedded methods frequently interact with the classifier but have computational efficiency better than wrapper methods, and are also tailored to a specific classifier [43]. Selection-Perceptron (FS-P)[48], Support Vector Machines (SVM-RFE) [49], Lasso (L1) and Elastic Net (L1+L2) based models [50][51] are some few examples of embedded based methods.

This paper is based on filter FS approach, and goal of this research work is to propose a new approach that selects more informative features from the original features set which help classification model to achieve good performance with regard to both time complexity and classification accuracy. The main point of view is on dimensional reduction, to reduce the number of features and processing time without sacrificing the classification accuracy. The features are exposed to double filter-based evaluation metrics (IG and *t-test*), in which at the final output, are obtained, only the discriminate features that highly contribute to the classification task, and produce a lower dimensionality subset base on features' respective rank and score. The approach blends the concepts of intersection and vector magnitude to select a subset of refined informative features by considering both the score and ranking of respective features. An experiment conducted with three distinct text datasets has shown that the proposed approach produces acceptable results by achieving a recorded performance of 67.65%, 54.74%, and 80.16%, and running time of 7464ms, 4689ms, and 29806ms on 20NewsGroups, NewsCategory, and Reuters-21784, respectively. This shows that the method retains most of the informative features when compared with other chosen methods.

The remaining body of this paper is systematically partitioned as follows: In Section 2, related works are presented. The proposed approach and the Filter-based feature selection methods employed explicitly by the approach, namely IG and *t*-test are discussed in Section 3. Properties of the datasets used and experimental set up are devoted to Section 4. Experiment results and discussion are systematically placed in Section 5. Finally, the study ends with a conclusion and highlights of possible future work which are given in Section 6.

II. RELATED WORKS

There are large number of research works on filter-based feature selection metrics to remove irrelevant and noisy features in text classification problem. The primary aim is often to reduce the feature dimensionality so as to minimize the processing time without sacrificing or improving the classification accuracy. In an effort to reduce the computational complexity, some numerous current works hybridized multiple scoring metrics to select most informative features. Results discrepancy is among the top challenges in hybridization approach as different results would be obtained when applying different evaluation metrics on the same dataset [38], and this issue can result in selecting noncontributory features. In this section, we will briefly present some review of those works, and lastly, we will summarize the drawbacks of the existing methods.

Lewis [52] uses mutual information (MI) to measure the importance of a feature, thus proposed a new scoring metric known as Mutual Information Maximization (MIM) that computes the relevancy between *n* features and classes. Liu and Setiono [39] proposed an algorithm that computes the score of each feature and selects relevant features based on chi-square score. The algorithm calculates the numeric attribute intervals and selects features according to the statistical data characteristics. A comparative study by Mladenic and Grobelnik [53] on a different dataset was conducted, and only for the Multinomial Naïve Bayes (NB) model upheld Odds Ratio over a wide variety of evaluation metrics been compared. For feature filtering, Bi-Normal Separation (BNS) has previously been described to be outstanding in ranking terms. Forman [54] improve an existing scoring metric for features by substituting IDF with BNS. The new method, TF-BNS scales the magnitude values and rank features by computing the BNS score of every feature. Empirical evaluation of text classification tasks using Support Vector Machine (SVM) shown significantly better performance in terms of F-measure and accuracy. Uguz [55] applies IG to ranked terms in a given document according to their importance in the initial stage of his proposed framework. Vinh et al. [56] proposed a new approach for selecting feature by normalizing well known MI (Mutual Information) measurement and used it to assess the potentiality of the features. Despite the competitive results achieved, the proposed approach could not conceal the highly correlated features influence the classification outcomes. Azhagusundari and Thanamani [57] developed a feature selection method based on IG for selecting the discriminant features from a give original set. The authors used IG to build a discernibility matrix which could be used to select the optimal subset of

features from the set of original data. Experimentally they showed their method obtained comparative classification accuracy on comparison with the original dimensionality. A greedy feature selection method using mutual information is introduced by Hoque and et al. [58]. The method blends feature–feature and feature–class MI to select the optimal feature subset. Wang et al. [4] use the concept of term frequency and developed a new feature scoring metric approach based on t-test, the method measures the diversity of the distributions of a feature between the particular category and the entire dataset. Experiment results indicate that the proposed method is marginally better than IG and chi-square method in terms of micro-F1 and macro-F1. Rehman et al. [5] proposed a novel function metric for feature ranking named Normalized Differences Measure (NDM), which evaluate the rank of a term by considering the term's relative document frequencies in both positive and negative classes. Zhou et al. [37] proposed a feature selection algorithm that uses segmented term frequency to compute the frequency of a document. Moreover, the impact of the same feature term to the classification under the dissimilar frequency of term is deeply considered. The algorithm uses the resultant terms' frequencies to give scores to each available feature and selects those features that are above a defined threshold. When Compared with six different FS methods, the empirical result demonstrated that the proposed method could able to increase classification accuracy on a textual dataset.

All the works mentioned earlier are single FS methods that consider only a single strategy for the selection of an informative subset of features. Consideration of multiple strategies altogether is impossible with a single feature selection method. In view of that, the hybridization approach has received significant attention in the field of dimensional reduction currently. The methods combined different FS methods considering various aspects of the features into single. Tsai and Hsiao [59] combine multiple methods for dimensional reduction to figure out more informative features for stock prices prediction task. The method integrates decision tree, PCA, and genetic algorithm as search methods, and utilizes the concept of an intersection, union, and multi-intersection approaches to filter out irrelevant variables. An intermediary method of union (OR) and Intersection (AND) approach named modified union is presented by Bharti and Singh [60]. The authors applied union (OR) and intersection (AND) on k -top selected ranked features, and on remaining unselected features subset, this merges the feature subsets into a single subset and further select the most relevant features. The feature filtering methods used in the study are document frequency (DF) together with term variance (TV). To exploit the advantages of two different FS methods, a hybridization of cluster-based and the frequency-based approach is presented by Nguyen and Bao [13]. The proposed method termed FCFS on comparison with its counterpart achieved the best performance in terms of micro-F1. To tackle the problem of results discrepancies, a new feature selection approach that combines the computed scores from multiple FS methods into one is proposed by Rajab [61]. The proposed method normalizes and computes vector score (V-Score) magnitude of each feature using the scores produced based on IG and Chi-square function metrics, and selects the top-ranked features.

Kamalov and Thabtah [62] proposed a method that selects optimal features from sets with ranking features produced by three different ranking strategies. The authors used vector scores (V-Scores) to stabilize the scores obtained from three methods (IG, Chi-square, and inter-correlation) and assign a new rank to each feature. To further remove non-relevant and noise features from feature subsets produced by two different evaluation functions, Li et al. [27] consider the application of union approach on the lowest rank feature subset produce by Fisher score and IG methods.

Many studies have investigated the strength of several filtering methods and their combination in the literature. Forman [32] empirically studied and compared twelve different evaluation metrics for feature selection on a text classification problem, and they finally revealed that BNS with IG has the minimum correlated failure so as mark best backup choice. The impact of integrating five methods for FS was investigated by Thubaity et al. [63]. The study employed IG, Chi-square, NGL, GSS, and RS methods on Arabic textual dataset. Union (OR) and intersection (AND) approach were utilized to integrate the scores produced from various FS methods employed to a single sorted feature set. Results Analysis showed there was no any improvement recorded in terms of classification accuracy when more than three FS metrics were integrated, while a small improvement was noticed for integrating two to three FS metrics. Vora and Yang [64] present a comparative study on ten different filtering methods namely Fisher Score, Chi-square, Gini Index, Laplacian Score, IG, mRmR, CFS, FCBF, Kruskal-Wallis, and RELIEFF. Experimented on five different text dataset, the authors found that combination of Kruskal-Wallis, Gini Index with SVM classifier lead the race as it achieved competitive classification performance but takes longer processing time, while IG and Chi-2 are projected as methods with a large number of similar features have been selected.

III. MATERIAL AND METHOD

This section presents a brief discuss on the information gain and t-test algorithm since both are useful for the proposed approach that will be explained in sub-section C.

A. Information Gain Algorithm

Gain (IG) [38][65], is an information theoretical and entropy-based method which is widely used in the field of dimensional reduction [43] [37]. IG is previously used to determine attribute use in splitting instances in decision tree-based models [66] and currently is applied to select the informative features subset in a given set of features. The method computes and assigns score to each feature considering the variation between entropy obtained based on presence or absence of term in a given category [37]. High information gain or high score indicates the discriminating capability of a feature and ranked top. The entropy of discrete random variable X is formulated as:

$$H(X) = -\sum_{x_i \in X} P(x_i) \log(P(x_i)) \quad (1)$$

x_i denotes a specific event of the variable X , $P(x_i)$ denotes the probability of an event (x_i). The general formula for computing IG of a given feature t is given as:

$$IG(t, c) = \sum_{c \in \{c_i, \bar{c}_i\}} \sum_{t \in \{t_k, \bar{t}_k\}} P(t, c) * \log \left(\frac{P(t, c)}{P(t) * P(c)} \right) \quad (2)$$

where $P(t, c)$ is the probability of class c and occurrence of the feature t . $P(t)$ is the probability of class containing feature t , $P(c)$ is the probability of class c . \bar{t}_k and \bar{c}_k denote feature not present, and class not present, respectively. Let N represent the total number of documents in a given dataset, and N_s with indicated subscripts values represents counts of documents. Using Maximum Likelihood estimates (MLEs) of probabilities, equation (2) can be expressed as:

$$I(t, c) = \frac{N_{11}}{N} \log_2 \left(\frac{NN_{11}}{N_1 N_1} \right) + \frac{N_{01}}{N} \log_2 \left(\frac{NN_{01}}{N_0 N_1} \right) + \frac{N_{10}}{N} \log_2 \left(\frac{NN_{10}}{N_1 N_0} \right) + \frac{N_{00}}{N} \log_2 \left(\frac{NN_{00}}{N_0 N_0} \right) \quad (3)$$

In information theory logic, a term/feature contains about the class, if the distribution of a term is equivalent in the class as it is in the whole collection, then $I(t, c) = 0$. IG attains its optimal value if the term is a perfect discriminator for class membership if the term exists in a document if only the document is in the class.

B. Student Statistical Test Algorithm

Statistical Test (t -test) is a statistical-based method which is commonly used to evaluate if the means of two groups are statistically different from each other by computing a ratio between the mean difference of two groups and the variability of the two groups [4][67]. Presently, t -test is widely used as an evaluation function to select significant features that contribute to classifying instances. The method computes score of feature by measuring the distinct distributions of the term in relevant category and documents collection [68]. The formula for calculating t -test is given as:

$$t - test(t_i, c_k) = \frac{|tf_{ki} - \bar{t}_i|}{m_k \times s_i} \quad (4)$$

$$S_i^2 = \frac{1}{N-K} \sum_{k=1}^k \sum_{j \in C_k} (tf_{ij} - \bar{t}_i)^2 \quad (5)$$

$$m_k = \sqrt{\frac{1}{N_k} - \frac{1}{N}} \quad (6)$$

Each class's specific scores obtained from (4) are combined to find the final score as follows:

$$t - test_{avg}(t_i) = \sum_{k=1}^k t - test(t_i, C_k) \quad (7)$$

where S_i denotes the standard deviation within a category, C_k denotes the k^{th} category, N_k is the number of documents in k^{th} category, k is the total number of categories, tf_{ki} denotes the average TF of term t_i in category k , \bar{t}_i denotes average TF of term t_i in the corpus. N is the total number of documents.

However, when the score is less than the defined threshold, it indicates that the feature has lower discrimination ability; otherwise, the feature will contribute in the classifying instances and will be selected.

C. Proposed Approach

Considering the problem of result discrepancy produced when two filtering methods are combined, and the risk of

losing informative features, an approach is proposed named new bi-strategy feature filtering approach which hybridizes IG with t -test to remove indiscriminate features by taking into consideration both feature ranking and vector score magnitude (V -score). The approach applies IG and t -test metrics independently to compute scores and assign the computed scores to each feature in the original features set, let say D_1 and D_2 . The top-ranked features that are greater than a predefined threshold K_1 are considered as significant features and are selected, new subsets of features S_1 and S_2 , which are based on IG and t -test are generated independently. Next, a feature with minimum IG score from S_1 and a feature with minimum t -test score from S_2 that are present in both S_1 and S_2 are selected and their V -scores are computed. The minimum V -score among the two computed V -scores is set as the new threshold K_2 . The approach further refines the features subsets by selecting a feature only if it is present in S_1 or S_2 and its V -score is greater than the new defined threshold K_2 otherwise it is an indiscriminate feature and will be neglected.

V -score of a given feature is computed using the concept of vector magnitude proposed in [61] that is, summing the squares of a vector's coordinates and taking the square root of the summation, it is formulated as:

$$V_{score} = \sqrt{(IG_{score})^2 + (t - test_{score})^2} \quad (8)$$

NB: The values of the scores produced by IG is different from that of t -test. So, we have to normalize the scores first before computing V -score so as to uniformly transform them into equivalent scale.

Let us consider Table I below, which contains few samples extracted from 20NewsGroups. It shows generated ranking of each feature based on the chosen filter methods. It can be seen that there are presence of discrepancies in the output. This issue arises due to the different theoretical strategy used by distinct filter methods to compute the score of each feature in the given dataset. IG ranked "Thanks" the lowest while t -test ranked it the highest, there is high assurance for IG method to eliminate this particular feature when a threshold is defined despite it has been selected by t -test method as the most informative feature. Therefore, both methods fall into the problem of losing informative features, likewise the existing hybrid filtering methods. Nevertheless, the proposed approach mitigates such issue by considering both the ranking and score of each feature. The approach sets a new thresholds base on computed V -score and further refines the features subset.

TABLE I. RANKING PRODUCED BY IG AND T-TEST FILTER METHODS

Features	Ranking	
	IG	t-test
space	1	3
god	2	2
orbit	3	5
religion	4	6
people	5	4
thanks	6	1

Framework of the proposed approach is based on the following algorithm 1:

ALGORITHM 1: PROPOSED BI-STRATEGY FILTERING APPROACH “MIN_MAX_V-SCORE”

INPUT:

- D : Set of documents with N features
- c : Set of label classes
- K_j : Initial threshold

OUTPUT:

- SL : Subset of selected features

FEATURE_SCORING(D, c)

```

1.  $L_1 \leftarrow [], L_2 \leftarrow []$ 
2.  $T \leftarrow$  EXTRACT TREMS IN DOCUMENTS ( $D$ )
3. For each  $t_i$  in  $T$  do:
4.    $A(t_i, c) \leftarrow$  COMPUTE(SCORE( $t_i, c$ )) using
     eqe(1) through eqe(3)
5.   APPEND( $L_1(A(t_i, c), t_i)$ )
6.    $A(t_i, c) \leftarrow$  COMPUTE(SCORE( $t_i, c$ )) using
     eqe(4) through eqe(7)
7.   APPEND( $L_2(A(t_i, c), t_i)$ )
8. End For
9. SORT( $L_1$ ), SORT ( $L_2$ )
10. Return  $L_1, L_2$ 

```

Begin

```

1.  $L_1, L_2 \leftarrow$  FEATURE_SCORING( $D, c$ )
2.  $FS1 \leftarrow k_1 \% \{L_1\} = \{t_1, \dots, t_q\}$ 
3.  $FS2 \leftarrow k_1 \% \{L_2\} = \{t_1, \dots, t_n\}$ 
4.  $j \leftarrow q$ 
5. For  $t_i$  in [SORT.descend( $FS1$ )]
6.   Normalise ( $t_i$ )
7.   If  $t_i \in \{FS1\} \cap \{FS2\}$ 
8.      $V_{score}(1) \leftarrow$  COMPUTE( $V_{score\_of\_t_i}$  using
       eqe (8))
9.     Break
10.  For  $t_j$  in [SORT.descend( $FS2$ )]
11.    Normalise ( $t_j$ )
12.    If  $t_j \in \{FS1\} \cap \{FS2\}$ 
13.       $V_{score}(2) \leftarrow$  COMPUTE( $V_{score\_of\_t_j}$  using
        eqe(8))
14.      Break
15.   $k_2 \leftarrow$  MIN( $V_{score}(1), V_{score}(2)$ )
16.  APPEND [ $SL, \{FS1 \cap FS2\}$ ]
17.  For  $t_i$  in [ $\{L_1 \cup L_2\} - \{SL\}$ ] do
18.    If  $V_{score}(t_i) \geq k_2$ 
19.      APPEND [ $SL, (t_i)$ ]
20.  End For
21. Return  $SL$ 

```

End

A summarized flowchart of the proposed methodology is depicted in Fig. 1. The process begins with the raw datasets as input. After relatively balancing the all unbalanced datasets, then original features set is constructed using TF-IDF. Next step is the initial features subsets formation using IG and t-test filter methods to compute and assign a score to all features and a sequence of high ranked features will be selected. Next,

we employed the proposed BI-strategy filtering approach to further refined the initial features subsets and generate the new informative features subset. Lastly, we validate the new approach by recording the Classification accuracy and f-score of the selected classifiers.

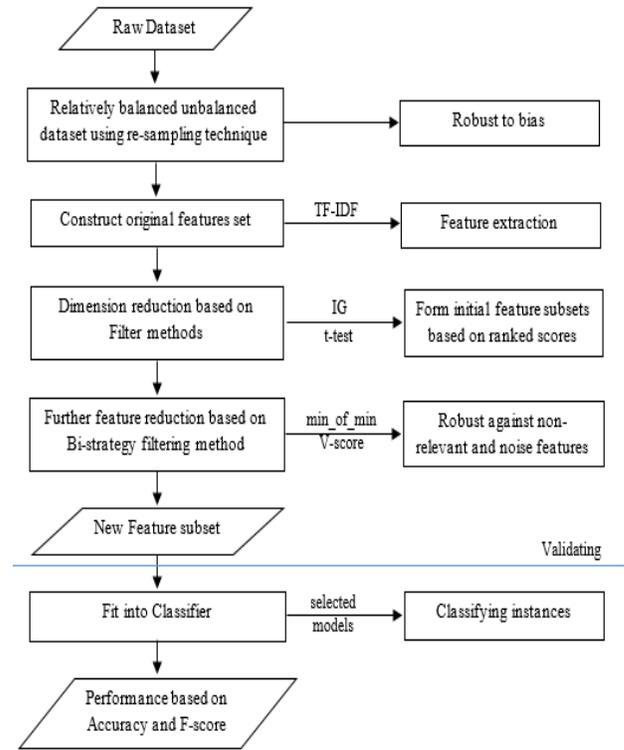


Fig. 1. Flowchart of the Proposed Methodology.

D. Experiment and Datasets

In this sub-section, summary of the datasets used and the experimental process adapted are briefly explained. The classification algorithms employed are also presented. Lastly, the section ends with a discussion on classifiers and implementation requirements.

1) *Dataset*: To evaluate the proposed approach in this experiment, the well-known three text benchmark datasets widely used for multi-class classification task is selected. Two of the datasets (Reuters 21578 and News Category) are unbalanced while the other one (20newsGroups) is balanced. We believed that both the datasets are highly dimensional with large number of samples, and also diversity amount of classes is considered. The summary information of the datasets is display in Table II.

TABLE II. SUMMARY OF THE DATASETS USED

Dataset	#Instances	#Features	#Classes
20news Groups	18846	173451	20
News Category	140597	1268350	36
Reuters-21578	11367	16578	90

The 20NewsGroups approximately comprises of 20,000 documents gathered from the collection of Usenet Newsgroups [69], and it consists of relatively balanced 20 distinct categories, each category contains around 1000 documents. The Reuters-21578 comprises of 21578 documents gathered from Reuters newswire, and it consists of unbalanced 135 categories with each document is associated with at least one categories (multi-label) [70][60]. Before importing the dataset into our experiment, we assigned only one category label to each document by stripping out all country names on the list and selecting the first topic left. Moreover, any document that is not associated with any topic was also eliminated from the dataset. This significantly reduced the number of categories and documents to 90 and 11367. The News Category comprises of around 150, 000 samples gathered from Short News Category, and it consists of unbalanced 41 categories with each document is associated with at least one categories (multi-label). We combined some few categories that can be naturally merged together, such as 'CULTURE & ARTS', ARTS & CULTURE', and 'ART'. We finally reduced the number of categories to 36.

2) *Experiment settings*: In the initial phase of the experiment, all English letters are converted into lowercase, stop words are removed, and words having non-characters are filtered. After then, roots of English words are found by applying porter stemmer algorithm [71]. And lastly, feature extraction is performed using TF-IDF weighting [72]. NB: all the three datasets are randomly divided into 60% training and 40% testing.

To validate the proposed filtering approach and its effectiveness on classification models, two existing benchmark methods for feature filtering are selected for comparison, namely IG and t-test. The selection is based on the fact that the proposed method is a hybrid of the selected methods. Besides the new approach, three other existing hybrid filtering approaches include Union (OR) approach, Intersection (AND) approach and Vector Magnitude (V-score) approach proposed in [61] are also selected. The initial threshold value K_l is based on the number of features been ranked in the original set and was set as 60% for both IG and t-test, any feature below the predefined threshold is low scored feature and will be disregarded otherwise will be qualified for further selection evaluation. Jaccard Similarity Coefficients (JCC) is used in this study to measure the similarity of features been selected by different benchmark filtering methods.

Five different well-known classification methods are used for validation purpose in this study. The selection is based on the positive recommendation of the methods in terms of text multi-class classification. The selected methods including Support Vector Machine (SVM) [18][22], Naïve Bayes (NB) [73], Decision Tree (DT) [38], Random Forest (RF)[74][18], and Ridge Regression (RR) [75][76]. All these models will be used to record the classification accuracy and performance of the stated filtering methods. Default values of most of the parameters associated with the classification methods are retained. For SVM and NB, multi-class SVC with kernel function and MultinomialNB are adapted while for RF number

of estimation was set to 100 when executed on 20News groups and Reuters 21578 datasets and set to 20 on News Category dataset, respectively.

Because of space limit, the performance of the classification methods will be reported using two standard recognized metrics widely used for text classification in literature, namely, Accuracy and F1-score. Accuracy is the percentage of the documents that are classified correctly in the given entire documents dataset. F1-score is the representation of harmonic mean of precision and recall. Accuracy and F1-score were computed using the following equations.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (9)$$

$$f - score = \frac{2 \times P \times R}{P+R} \quad (10)$$

Where, 'TP' = True Positive (documents correctly classified as positive), 'TN' = True Negative (documents correctly classified as negative), 'FP' = False Positive (documents incorrectly classified as Positive), 'FN' = False Negative (documents incorrectly classified as Negative), and 'P' and 'R' are precision and recall values and are computed using the following equations.

$$precision = \frac{TP}{TP+FP} \quad (11)$$

$$recall = \frac{TP}{TP+FN} \quad (12)$$

In this study, all the implementations for the experiment are conducted on Python (V3.8.2) environment, which is installed on a computer with Windows 8 (OS). Other minimum required conditions for the experiment include Intel(R) Core™ i5 processor4300m@2.60GHz/8GRAM/64 GB.

IV. DISCUSSION OF RESULTS

Table III shows a brief description of the chosen filtering methods based on the formulation and strategy adapted. As it can be seen that all the selected hybrid filtering method were formulated by integrating information theory and statistical theoretical based benchmark methods (IG and t-test), the reason behind the selection of this two benchmark methods is by considering the Jaccard Similarity Coefficients between them which is very low compared to other methods. The average percentages of features reduced by different methods in all the three datasets are displayed in Fig. 2. From the figure, it can be seen that the percentage reduction differences in terms of feature dimensions between the proposed method (PM) and existing methods (IG, TS, UA, IA, VS). PM, UA, and IA reduced the number of features by 52.07%, 19.2% and 60% on 20NewsGroups Dataset, where as 48.08%, 26.53%, and 53.48% on NewsCategory Dataset and finally 45.25%, 28.86% and 51.15% on Reuters-21578 Dataset respectively. While IG, TS and VS reduced the features by 40% in all the three datasets, this is because a fixed threshold K_l was defined in all the experiments. The figure reveals in all the three datasets, the proposed method comparatively reduces the feature dimensions, with IA and UA achieved the highest and lowest percentage of features been reduced in all the datasets.

TABLE III. A BRIEF DESCRIPTION OF THE EXISTING AND PROPOSED FILTERING METHOD(S) EMPLOYED FOR THE EXPERIMENT

Method	Acronym	Description	Strategy
Information Gain	IG	Selects top-ranked k features based on IG scores: {K%(IG)}	Single
t-test	TS	Selects top-ranked features based on t-test score: {K%(TS)}	Single
Union Approach	UA	Selects features from hybrid of IG and TS based on Union (OR) approach: {K%(IG)} U {K%(TS)}	Hybrid
Intersection Approach	IA	Selects features from hybrid of IG and TS based on intersection (AND) approach : {K%(IG)} ∩ {K%(TS)}	Hybrid
V-Score	VS	Selects features from hybrid of IG and TS based on V-score: {K%(VS)}	Hybrid
Proposed Approach	PM	Selects features from hybrid of IG and TS based on modified V-score: {K ₂ %VS({K ₁ (IG) U K ₁ (TS)})}	Hybrid

The proposed approach could not beat IA method in terms of feature reduction because we seriously take into

TABLE IV. PERFORMANCE OF THE EXISTING AND PROPOSED FILTERING METHOD(S) IN TERMS OF ACCURACY AND RUNNING TIME ON NEWS CATEGORY DATASET

Classifier	Metrics	PM	IG	TS	UA	IA	VS
Ridge	Accuracy	0.59657	0.59413	0.59604	0.59553	0.54383	0.59455
	Running Time (ms)	7340	7568	7818	8239	7234	7756
Multinomial NB	Accuracy	0.54583	0.53742	0.54318	0.53990	0.52575	0.53918
	Running Time (ms)	1223	1249	1287	1302	1215	1242
LinearSVC	Accuracy	0.60230	0.60001	0.60226	0.59994	0.54972	0.59913
	Running Time (ms)	14259	15527	15621	15855	14117	15497
Decision Tree	Accuracy	0.45740	0.45415	0.45214	0.45655	0.42503	0.45668
	Running Time (ms)	91011	96414	96382	102909	90715	97462
Random forest	Accuracy	0.53541	0.52881	0.52657	0.53331	0.51271	0.53361
	Running Time (ms)	120613	125019	124444	131862	120578	126953

TABLE V. PERFORMANCE OF THE EXISTING AND PROPOSED FILTERING METHOD(S) IN TERMS OF ACCURACY AND RUNNING TIME ON 20NEWSGROUPS DATASET

Classifier	Metrics	PM	IG	TS	UA	IA	VS
Ridge	Accuracy	0.75362	0.75203	0.75216	0.75266	0.73381	0.74973
	Running Time (ms)	0859	1064	1054	1173	0791	1071
Multinomial NB	Accuracy	0.74770	0.74000	0.75008	0.74177	0.72638	0.73134
	Running Time (ms)	0.035	0047	0055	0095	0034	0046
LinearSVC	Accuracy	0.74923	0.74832	0.7538	0.74460	0.72621	0.74159
	Running Time (ms)	0757	0977	1001	1063	0742	0931
Decision Tree	Accuracy	0.47285	0.47930	0.47054	0.47064	0.44798	0.47170
	Running Time (ms)	8032	9552	10765	8268	7958	8797
Random forest	Accuracy	0.65909	0.65759	0.66289	0.65015	0.64113	0.65051
	Running Time (ms)	27641	29985	30004	31452	27602	30043

consideration the risk of avoiding losing informative features which will suffer the performance of classifier as discovered with IA and related methods. In particular, our proposed approach saves as an intermediary between IA and the other methods.

Tables IV, V and VI show the classifiers' performance including Ridge, MNB, SVC, DT and FR based on classification accuracy and running time after applying the existing and proposed filtering methods on the text datasets. Best results are face bolded. The impact of filtering methods on the classifier performance in both the five classifiers results is noticeable.

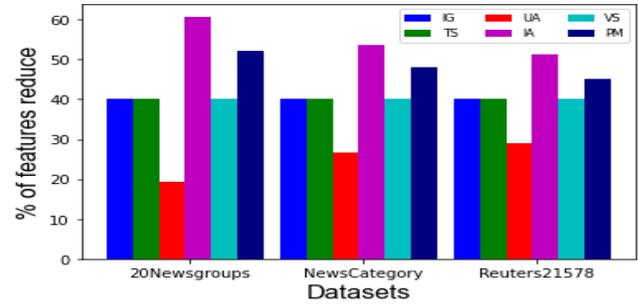


Fig. 2. Percentage of Features Reduced for 20 New Groups, News Category, and Reuters-21578 Datasets.

TABLE VI. PERFORMANCE OF THE EXISTING AND PROPOSED FILTERING METHOD(S) IN TERMS OF ACCURACY AND RUNNING TIME ON NEWS CATEGORY DATASET FOR DIFFERENT CHOSEN FILTERING METHODS ON 20 NEWSGROUPS DATASET

Classifier	Metrics	PM	IG	TS	UA	IA	VS
Ridge	Accuracy	0.84774	0.84697	0.84667	0.84579	0.83403	0.84579
	Running Time (ms)	2788	3114	3158	3160	2615	3023
Multinomial NB	Accuracy	0.80123	0.79420	0.79683	0.80299	0.74947	0.80064
	Running Time (ms)	0074	0.089	0092	0084	0074	0084
LinearSVC	Accuracy	0.84902	0.84667	0.84755	0.84667	0.80872	0.84667
	Running Time (ms)	1180	1324	1348	1356	1144	1311
Decision Tree	Accuracy	0.72619	0.72032	0.72618	0.72527	0.67178	0.72589
	Running Time (ms)	2822	2935	2974	3045	2820	2995
Random forest	Accuracy	0.78511	0.78100	0.77602	0.77631	0.72157	0.78130
	Running Time (ms)	7676	7901	7880	7942	7676	7899

In Table IV, accuracy results and running times are summarized for different chosen filtering methods on 20 Newsgroups dataset. Both the methods showed good performance with all the classifiers except with DT. The average classification accuracy, when filtering methods (IG, TS, UA, IA, VS and PM) were applied are 67.54%, 68.09%, 67.20%, 65.51%, 66.90%, and 67.65%. However, from the average score, we notice that the accuracy by different filtering methods is basically at the same level across all the classifiers with TS achieved the highest accuracy score followed by our method with no much significant difference. As can be seen from the table, in terms of running time, the proposed approach and IA marked the lowest as they achieved an average running time of 7464ms and 7425ms, thus supersede TS and the other methods. In particular, our method shows competitive performance on 20NewsGroups dataset.

The classification performance on NewsCategory dataset is shown in Table V. the average classification accuracy for the filter methods are 54.29%, 54.40%, 54.50%, 51.14%, 54.46%, and 54.75%. We notice the average accuracy of the proposed approach is comparatively little bit higher than that of the other filter methods compared. From the table, it can be seen that in most cases, our approach achieved a lower running time (4689ms averagely) but a little bit higher than IA (4635ms averagely). However, the overall comparison on NewsCategory dataset shows our method achieved significant performance.

Table VI reports the classification performance on Reuters-21578 dataset. It shows that the accuracy by different filtering methods is roughly similar. The average accuracy for the filter methods are 79.78%, 79.87%, 75.72%, 80.00%, and 80.16%. Compared with the other filter methods, the average accuracy of the proposed method is comparatively higher. The lowest average running time is achieved by IA as 29865ms and then followed by our method as 29806ms upon all the filter methods.

In general, the performance of the filter methods reported on each dataset is roughly at the same level across all the five classifiers. On 20NewsGroups dataset, we observed that TS recorded the highest classification accuracy with a slice

difference than that of our method, but the running of our method is significantly lower than that of TS. While on NewsCategory and Reuters-21576 datasets, our method recorded the highest classification accuracy with lower running time. IA generally recorded the lowest running time upon all the classifiers but sacrificed their performance. This indicates that the method filters out some informative features, thus reducing classification capability. On the other hand, IG, TS, UA, and VS achieved competitive performance but the running time is significantly high, and this indicates the presence of noise and irrelevant features in the final subset produced which need to filter out so as to reduce the time complexity. We also observed that the best accuracy results were obtained with the Ridge classifier and SVC classifier, whereas the results that are obtained with DT are comparatively bad. Generally, the proposed method achieves acceptable performance on all datasets, which indicates that this kind of filter method can not only reduce the size of the features set but also ensure that informative features are retained so that the performance of a classifier is not sacrificing.

Despite work done to balance the two unbalanced datasets used in this experiment still, the datasets are relatively unbalanced. Therefore using accuracy metrics to evaluate the performance could be misleading. In order to further verified the validity of the proposed approach, Fig. 3, 4 and 5 shows the performance results based on F-score of the chosen classifiers when the proposed and existing filter methods were applied on the three datasets selected. Examining both the figures, we can see that the results obtained are in line with accuracy results obtained in Tables IV, V, and VI. The information depicted in Fig. 3 shows the approach recorded the highest F-score after TS with a relatively small difference. Moreover, in Fig. 4 and 5, the proposed approach attains the highest F- score with a minimal gap. Therefore, we conclude that the proposed method achieves the best classification performance in terms of the highest F-score on most of the cases. Although the proposed approach, does not always give the highest result on all the datasets such as with 20NewsGroups, but the F-scores results are still acceptable.

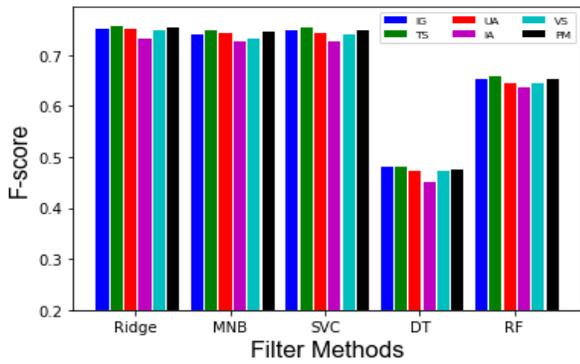


Fig. 3. Performance of the Existing and Proposed Filtering Method(s) in Terms of F-Score on 20 News Groups Dataset.

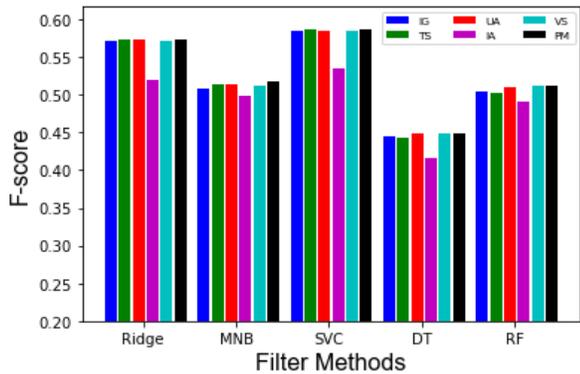


Fig. 4. Performance of the Existing and Proposed Filtering Method(s) in Terms of F-Score on News Category Dataset.

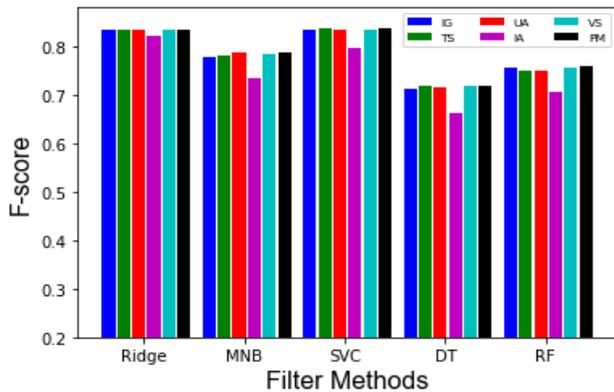


Fig. 5. Performance of the Existing and Proposed Filtering Method(s) in Terms of F-Score on Reuters-21578 Dataset.

V. CONCLUSION AND FUTURE DIRECTIONS

Filter-based Feature selection is one of the dimensional reduction techniques, and it is an important preprocessing step of any text classification problem. Selecting the most informative features is one of the main problems faced in building a robust classifier due to performance degradation and time complexity. Reducing the dimension of features by removing irrelevant and noise features as well as retaining the relevant features will significantly reduce the classifiers' computational complexity. There have been a quiet number of works done in the literature to address this problem. The

common filtering methods select features by considering a single theoretical approach. Recently, a hybrid approach that combines multiple filtering methods based on different theoretical approach receives more attention. These methods produce a discrepancy in the result that makes combining features subsets produced into a single subset and selecting significant features a difficult task. In this paper, we propose a novel Bi-strategy fileting approach that uses the combined scores of IG and t -test to produce refined features subsets by setting a new threshold. The method filters out common features with low V-scores from the considered subsets of features without sacrificing classifier's performance. First, two subsets with high ranked features based on IG and t -test are produced. This is done by defining the initial threshold K_1 . Then the method identified a feature with minimum IG and t -test scores that are present in both subsets produced and compute their V-scores. The minimum V-score is set as the new threshold K_2 , and it is used to further filter out insignificant features from the IG and t -test subsets.

In order to validate the performance of the proposed method, the study presents a comparison based on accuracy and F-score of the filtering approach with that of benchmark methods include IG, t -test (TS), and existing hybrid subsets merging approaches include Union (UA), Intersection (IA) and V-score (VS) using five classification algorithms. The experiment is conducted using three different text datasets, 20Newsgroups, NewsCategory, and Reuters-21578. Results in Fig. 2 show that our filter method produces a subset with features that is higher than that of IA in number but smaller than that of IG, TS, UA, and VS. It is the fact that our method ignored irrelevant and noisy features and at the same time retained much more informative features, unlike IA. Further experiment results showed that with the small size of features subset produced, our approach achieved a significant improvement in terms of accuracy and F-score of the classifiers used at the cost of a minimum running time. Lastly, a conclusion is reached that the proposed approach achieved a competitive performance even though it does not always give the highest result in most cases, but the results are still acceptable.

In future work, there is a need to investigate the following task: (1) To develop and in-cooperate a feature hashing method as the next step to our method that will consider the correlation between features. (2) To develop a method that has the capability to automatically determine optimal threshold parameter(s) between significant and non-significant features without any domain expert involvement.

ACKNOWLEDGMENT

The authors would like to thank the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia for the contribution and support toward this research.

REFERENCES

- [1] M. F. Karaca and S. Bayir, "Examining the impact of feature selection methods on text classification" *Int. J. of Adv. Comput. Sci. and Applications*, vol. 8, no. 12, 2017.
- [2] F. Salo, A. B. Nassif, and A. Essex, "Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection," *Comput.*

- Networks, vol. 148, pp. 164–175, 2019, doi: 10.1016/j.comnet.2018.11.010.
- [3] M. Shervin, K. Nal, C. Erik, N. Narjes, C. Meysam, and G. Jianfeng, “Deep learning text based classification: A comprehensive review,” *ACM compt. Surv.*, vol. 3, no. 62, pp. 1-40, 2021.
- [4] D. Wang, H. Zhang, R. Lui, and W. Lv, “Feature selection based on term frequency and t-test for text categorization,” *Pattern Recognit. Lett.*, vol. 45, no. 1, pp. 1–6, 2013, doi: 10.1016/j.patrec.2014.02.013.
- [5] K. Javed and H. A. Babri, “Feature selection based on a normalized difference measure for text classification,” *Inf. Process. Manag.*, vol. 53, no. 2, pp. 473–489, 2017, doi: 10.1016/j.ipm.2016.12.004.
- [6] N. Khan, M. S. Husain, and M. R. Beg, “Big data classification using evolutionary techniques: A survey,” in *Proceedings of IEEE International Conference on Engineering and Technology (ICETECH)*, 2015, no. March, pp. 243–247.
- [7] A. Faraz, “An elaboration of text categorization and automatic text classification through mathematical and graphical modelling,” *Comput. Sci. Eng. An Int. J.*, vol. 5, no. 2/3, pp. 1–11, 2015, doi: 10.5121/cseij.2015.5301.
- [8] J. Luengo, D. García-gil, S. Ramírez-gallego, S. Garcia, and F. Herrera, *Big Data Preprocessing*, 1st ed. Cham: Springer, 2020.
- [9] A. K. Nikhath, K. Subrahmanyam, and R. Vasavi, “Building a K-Nearest Neighbor classifier for text categorization,” *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 1, pp. 254–256, 2016.
- [10] A. M. Mostafa, “An evaluation of sentiment analysis and classification algorithms for arabic textual data,” *Int. J. Comput. Appl.*, vol. 158, no. 3, pp. 29–36, 2017.
- [11] F. Barigou, “Improving K-nearest Neighbor efficiency for text categorization,” *Neural Netw. World*, vol. 1, no. July, pp. 45–65, 2016, doi: 10.14311/NNW.2016.26.003.
- [12] Y. Chen, Q. Zhou, L. Wei, and D. Ji-Xiang, “Classification of chinese texts based on recognition of semantic topics,” *Cognit. Comput.*, vol. 8, no. 1, pp. 114–124, 2015, doi: 10.1007/s12559-015-9346-8.
- [13] L. N. H. Nam and H. B. Quoc, “A combined approach for filter feature selection in document classification,” in *In Proceedings of the International Conference on Tools with Artificial Intelligence, ICTAI*, 2016, vol. 2016-Janua, pp. 317–324, doi: 10.1109/ICTAI.2015.56.
- [14] R. Janani and S. Vijayarani, “Automated text classification using machine learning and optimisation algorithms,” *Soft Comput.*, vol. 25, no. 1, pp. 1129–1145, 2021.
- [15] D. S. Guru, M. Suhil, L. N. Raju, and N. V. Kumar, “An alternative framework for univariate filter based feature selection for text categorization,” *Pattern Recognit. Lett.*, vol. 103, no. January, pp. 23–31, 2018, doi: 10.1016/j.patrec.2017.12.025.
- [16] M. Micheal and L. Tony, “Evaluation of dimensionality reduction techniques: Principal feature analysis in case of text classification problems,” in *Proceeding of the 16th International Conference on Computing and Data Engineering*, 2020, pp. 75-79.
- [17] S. Ayesha, M. K. Hanif, and R. Talib, “Overview and comparative study of dimensionality reduction techniques for high dimensional data,” *Inf. Fusion*, vol. 59, no. January, pp. 44–58, 2020, doi: 10.1016/j.inffus.2020.01.005.
- [18] M. N. Asim, A. Rehman and U. Shoaib, “Accuracy based feature ranking metric for multi-label text classification” *Int. J. of Adv. Comput. Sci. and Appl. (IJACSA)*, vol. 8, no. 10, 2017.
- [19] L. N. Nam Hoai and H. B. Quoc, “A comprehensive filter feature selection for improving document classification,” in *Proceedings of 29th Pacific Asia Conference on Language, Information and Computation (PACLIC)*, 2015, pp. 169–177.
- [20] M. Habib, C. Sun, A. Abbas, and P. Prakash, “Big data reduction methods: A survey,” *Springer-Data Sci. Eng.*, vol. 1, pp. 265–284, 2016, doi: 10.1007/s41019-016-0022-0.
- [21] L. Mahdieh, M. Parham, A. Fardin, and J. Mahdi, “A novel multivariate filter method for feature selection in text classification problems,” *J. of Eng. Aapth. Of Artificial Intelli.*, vol. 70, pp. 25-37, 2017.
- [22] S. Suthaharan, *Machine Learning Models and Algorithms for Big Data Classification*, 2016th ed. Greensboro, USA: Springer, 2016.
- [23] N. Pilnenskiy and I. Smetannikov, “Feature selection algorithms as one of the python data analytical tools †,” *Futur. internet Artic.*, vol. 54, no. 12, pp. 1–14, 2020, doi: 10.3390/fi12030054.
- [24] D. Wang, H. Zhang, R. Liu, X. Liu, and J. Wang, “Unsupervised feature selection through Gram – Schmidt orthogonalization — A word co-occurrence perspective,” *Neurocomputing*, vol. 173, pp. 845–854, 2016, doi: 10.1016/j.neucom.2015.08.038.
- [25] J. Golay and M. Kanevski, “Unsupervised feature selection based on the morisita estimator of intrinsic dimension,” *Springer-Knowledge-Based Syst.*, vol. 135, no. Nov, pp. 125–134, 2017, doi: 10.1016/j.knsys.2017.08.009.
- [26] K. Ikeuchi, *Computer Vision: A Reference Guide*, 2014th ed., vol. 2. Boston, USA: Springer, 2014.
- [27] M. Li, H. Wang, L. Yang, Y. Liang, and Z. Shang, “Fast hybrid dimensionality reduction method for classification based on feature selection and grouped feature extraction,” *Expert Syst. Appl.*, vol. 150, no. July, pp. 1–10, 2020, doi: 10.1016/j.eswa.2020.113277.
- [28] A. Onan and K. Serar, “A feature selection model based on genetic rank aggregation for text sentiment classification,” *J. Inf. Sci.*, vol. 43, no. 1, pp. 25–38, 2015, doi: 10.1177/0165551515613226.
- [29] I. M. El-Hasnony, S. I. Barakat, M. Elhoseny, and R. R. Mostafa, “Improved feature selection model for big data analytics,” *IEEE Access*, vol. 8, pp. 66989–67004, 2020, doi: 10.1109/ACCESS.2020.2986232.
- [30] Z. M. Hira and D. F. Gillies, “A review of feature selection and feature extraction methods applied on microarray data,” *Adv. Bioinformatics*, vol. 2015, no. July, pp. 1–13, 2015, doi: 10.1155/2015.
- [31] A. Subasi, *Feature Extraction and Dimension Reduction.*, 1st ed., Elsevier Inc. 2019.
- [32] M. Rong, D. Gong, and X. Gao, “Feature selection and its use in big data: challenges, methods, and trends,” *IEEE Access*, vol. 7, pp. 19709–19725, 2019.
- [33] A. D. Rahajoe “Forecasting feature selection based on single exponential smoothing using wrapper method” *Int. J. of Adv. Comput. Sci. and Appl. (IJACSA)*, vol. 10, no. 6, pp. 139-145, 2019.
- [34] N. Gu, M. Fan, L. Du, and D. Ren, “Efficient sequential feature selection based on adaptive eigenspace model,” *Neurocomputing*, vol. 161, pp. 199–209, 2015, doi: 10.1016/j.neucom.2015.02.043.
- [35] M. Rong, D. Gong, and X. Gao, “Feature selection and its use in big data: challenges, methods, and trends,” *IEEE Access*, vol. 7, pp. 19709–19725, 2019, doi: 10.1109/ACCESS.2019.2894366.
- [36] B. Andrea, S. Xudong, B. Bernd, and L. Micheal, “Benchmark for filter methods feature selection in high-dimensional classification data,” *Compt. Stat. abd data Analy.*, vol. 143, no. 106839, pp. 1-19, 2020.
- [37] H. Zhou, S. Han, and Y. Liu, “A novel feature selection approach based on document frequency of segmented term frequency,” *IEEE Access*, vol. 6, pp. 53811–53821, 2018, doi: 10.1109/ACCESS.2018.2871109.
- [38] W. Xinzheng, G. Bing, S. Yan, Z. Chimin, and D. Xuliang, “Input feature selection method based on feature set equivalence and mutula information gain maximization,” *IEEE Access*, vol. 7, pp. 151525–151538, 2019.
- [39] B. Nurhayati, E. P. Armanda, and W. L. Kesuma, “Chi-square feature selection effect on Naïve Bayes classifier algorithm performance for sentiment analysis document,” in *Proceedings of the 7th International Conference on Cyber and IT Service Management*, 2019, pp. 1–7.
- [40] [40] C. Li and X. Jiucheng, “Feature selection with Fisher score followed by the Maximal Clique Centrality algorithm can accurately identify the hub genes of hepatocellular carcinoma,” *Sci. Rep.*, vol. 9, no. 1, pp. 17283, 2019.
- [41] M. Jafari, B. Ghavami, and V. Sattari, “A hybrid framework for reverse engineering of robust Gene Regulatory Networks,” *Artif. Intell. Med.*, vol. 79, pp. 15–27, 2017, doi: 10.1016/j.artmed.2017.05.004.
- [42] N. El Aboudi and L. Benhlima, “A review on wrapper feature selection approaches,” in *Proceedings of International Conferencr of Engineering and MIS (ICEMIS)*, 2016, pp. 1–5.
- [43] D. Ö. Şahin and E. Kılıç, “Two new feature selection metrics for text classification,” *J. Control. Meas. Electron. Comput. Commun.*, vol. 60, no. 2, pp. 162–171, 2019, doi: 10.1080/00051144.2019.1602293.

- [44] A. K. Das, S. Sengupta, and S. Bhattacharyya, "A Group Incremental Feature Selection for Classification using Rough Set Theory based Genetic Algorithm," *Appl. Soft Comput. J.*, vol. 64, no. April, pp. 400–411, 2018, doi: doi.org/10.1016/j.asoc.2018.01.040.
- [45] J. Kittler, *Feature selection and extraction*. Elsevier Inc., 2014.
- [46] G. Jesus and Q. G. John, "A new multi-objective wrapper method for feature selection-Accuracy and stability analysis for BCI," *Neurocomputing.*, vol. 333, pp. 407-418, 2017.
- [47] C. Hou, F. Nie, X. Li, D. Yi, and Y. Wu, "Joint embedding learning and sparse regression: A framework for unsupervised feature selection," *IEEE Trans. Cybern.*, vol. 44, no. 6, pp. 793–804, 2014, doi: 10.1109/TCYB.2013.2272642.
- [48] R. Chakraborty and N. R. Pal, "Feature selection using a neural framework with controlled redundancy," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 1, pp. 35–50, 2015, doi: 10.1109/TNNLS.2014.2308902.
- [49] T. Kari *et al.*, "Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm," *IET Gener. Transm. Distrib.*, vol. 12, no. 21, pp. 5672–5680, 2018, doi: 10.1049/iet-gtd.2018.5482.
- [50] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 67, no. 2, pp. 301–320, 2005, doi: 10.1111/j.1467-9868.2005.00527.x.
- [51] D. Loanni, "A review on variable selection in regression Analysis," *econometrics*, vol. 6, no. 25, pp. 1-27, 2018.
- [52] D. D. Lewis, "Feature selection and feature extraction for text categorization," in *In Proceedings of Speech and Natural Language: workshop held at Harriman*, 1992, pp. 212–217.
- [53] D. Mladenic and G. Marko, "Feature selection for unbalanced class distribution and naive bayes.," in *Proceedings of the Sixteenth International Conference on Machine Learning (ICML)*, 1999, no. January, pp. 258–267.
- [54] G. Forman, "BNS Feature Scaling : An Improved representation over TF · IDF for SVM text classification," in *ACM 17th Conference on Information and Knowledge Management*, 2008, pp. 1–8.
- [55] H. Uguz, "Knowledge-Based Systems A two-stage feature selection method for text categorization by using information gain , principal component analysis and genetic algorithm," vol. 24, pp. 1024–1032, 2011, doi: 10.1016/j.knosys.2011.04.014.
- [56] L. T. Vinh, L. Sungyoung, L. Y. Park, and B. J. D'Auriol, "A novel feature selection method based on normalized mutual information," *Appl. Intell.*, vol. 37, pp. 100–120, 2012, doi: 10.1007/s10489-011-0315-y.
- [57] B. Azhagusundari and A. S. Thanamani, "Feature selection based on information gain," in *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2013, no. 2, pp. 18–21.
- [58] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, "A mutual information-based feature selection method," *Expert Syst. Appl.*, vol. 2014, no. April, pp. 1–15, 2014, doi: 10.1016/j.eswa.2014.04.019.
- [59] C. Tsai and Y. Hsiao, "Combining multiple feature selection methods for stock prediction: Union , intersection , and multi-intersection approaches," *Decis. Support Syst.*, vol. 50, no. 1, pp. 258–269, 2010, doi: 10.1016/j.dss.2010.08.028.
- [60] K. K. Bharti and P. K. Singh, "Hybrid dimension reduction by integrating feature selection with feature extraction method for text clustering," *Expert Syst. Appl.*, vol. 42, no. 6, pp. 3105–3114, 2015, doi: 10.1016/j.eswa.2014.11.038.
- [61] K. D. Rajab, "New Hybrid features selection method : A case study on websites phishing," *Secur. Commun. Networks*, vol. 2017, no. March, pp. 1–10, 2017.
- [62] F. Kamalov and F. Thabtah, "A feature selection method based on ranked vector scores of features for classification," *Ann. Data Sci.*, pp. 1–20, 2017, doi: 10.1007/s40745-017-0116-1.
- [63] A. Al-thubaity, N. Abanumay, and Z. Mannaa, "The effect of combining different feature selection methods on arabic text classification," in *14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, 2013, pp. 219–224, doi: 10.1109/SNPDP.2013.89.
- [64] S. Vora and H. Yang, "A comprehensive study of eleven feature selection algorithms and their impact on text classification," in *proceeding of Computing Conference*, 2017, no. July, pp. 440–449.
- [65] D. Apriliani, T. Abidin, E. Sutanta, A. Hamzah, and O. Somantri "Sentiment analysis for assessment of hotel services review using feature selection approach based-on decision tree" *Int. J. of Adv. Comput. Sci. and Appl. (IJACSA)*, vol. 11, no. 4, pp. 240-245, 2020.
- [66] F. Thabtah and F. Kamalov, "Phishing detection : A case analysis on classifiers with rules using machine learning," *J. Inf. Knowl. Manag.*, vol. 16, no. 4, pp. 1–16, 2017, doi: 10.1142/S0219649217500344.
- [67] N. O. Essied, I. Othman, and A. H. Osman, "A novel feature selection based on one-way ANOVA F-Test for e-mail spam classification," *Res. J. Appl. Sci. Eng. Technol.*, vol. 7, no. 3, pp. 625–638, 2014, doi: 10.19026/rjaset.7.299.
- [68] Y. Liu, S. Ju, J. Wang, and C. Su, "A new feature selection method for text classification based on independent feature space search," *Math. Probl. Eng.*, vol. 2020, pp. 1–14, 2020, doi: 10.1155/2020/6076272.
- [69] A. O. Adi, and E. Celebi, "Classification of 20 news group with Naïve Bayes classifier," in *Proceedings of 22nd Signal Processing and Communications Applications Conference (SIU)*, 2015, pp. 2150–2153.
- [70] "Porter Stemming Algorithm (PSA): <http://tartarus.org/martin/PorterStemmer/> last access: October," 2019.
- [71] R. Ansari, M. Salwani, Z. Nor, and S. Faezeh, "Stemming text-based web page classification using machine learning algorithm: A comparison," *Int. J. of Adv. Comp. Sci. and Appl.*, vol. 11, no. 1, pp. 570–576, 2020.
- [72] K. M. M. Rajashekharaiyah, S. S. Chikkalli, P. K. Kumbar, and P. S. Babu, "Unified framework of dimensionality reduction and text categorisation," *Int. J. Eng. Technology*, vol. 7, no. November 2018, pp. 648–654, 2018, doi: 10.14419/ijet.v7i3.29.21397.
- [73] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: A survey," *Information.*, vol. 10, no. 4, pp. 1–68, 2019, doi: 10.3390/info10040150.
- [74] I. Zahidul, L. X. Jixue, L. Jiuyong, L.lin, and K. Wei, "A semantics aware random forest for text classification," in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management.*, 2019, pp. 1061-1070.
- [75] N. Jothi, W. Husain, A. N. Rashid Abdul, and S. M. Syed-mohamad, "Feature selection method using genetic algorithm for medical dataset," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, no. 6, pp. 1907–1912, 2019.
- [76] Z. Xiaoming, C. Wenhan, L.Zhoujun, L. Chunyang, and L.Rui, "Multi-modal kernel ridge regression for social image classification", *Applied Soft Cmmpt.*, vol. 67, pp. 117–125, 2018.

An Ontological Framework for Healthcare Web Applications Security

Mamdouh Alenezi

College of Computer and Information Sciences
Prince Sultan University, Riyadh
Kingdom of Saudi Arabia

Abstract—The current era of digitization and transformation causes various issues and advantages both at the same time in the healthcare sector. The beneficial advantages are exceptionally good but the issues that are gardening the business of attackers is a serious issue and requires effective prevention. Current statistics and attack vector analysis portray that technical breaches have the most common and highest priority in numbers. This type of information opens a need to prevent and develop an effective model that helps to exerts and healthcare practitioners in security management. In order to achieve this desired goal, we adopt and apply an ontology-based approach of security and development methodology and provide a model that effectively produces systematic secure pathways to design healthcare web applications. The conceptual framework discussed in the study has many effective and beneficial advantages namely, it gives a unified pathway to future developers; the model also attains focus on requirement identification during development and portrays its significance.

Keywords—Security; web application; healthcare; digitization; requirement

I. INTRODUCTION

The era of digitization led to the whole concept of businesses and services on a virtual platform. This type of transformation creates an exceptionally beneficial environment of working. It is unified that digitization has various beneficial and good effects for current businesses and services but there are also some adverse effects available that tell us about the possibility of exploitation and digital systems and the impact of harm after exploitations. It is very significant and necessary to investigate and facilitate models that reduce the possibility of exploitation in systems.

This type of exploitation possibility and attack situations raise the requirement of a solid and secure fundamental concept of security in web application security. These types of exploitations and attacks pose some serious harm and defects in the web application that cause serious harm for hosting organizations and businesses [1]. Any type of exploitation and vulnerability creates an adverse reputation situation for business because users do not want a type of interruption or loss of data during the use of web application and if it causes then the reputation of the application and its owner gets affected. Now, it is frequently shown that this type of exploitation possibilities and issues appeared only because of any type of flaw and vulnerability in the application.

Most of the time these vulnerabilities and flaws are overlooked by the development team because of the deadline and hurry issue [2]. Further, many of the organizations do not identify which factors need attention for security or which not. However, from some previous years, many researchers and experts are paying attention to factors that need considerable attention for security in software and web application and in this proposed article we are also paying attention to them. Web application security is something that associates the whole development functionality of it when we talk about its security. Security is something that can only be achieved by systematic security towards development in web applications [3]. Producing a secure development and development towards systematic security is something that demands an appropriate equipped team and laboratory.

To make this development and its system more clear and efficient SHIELDS discusses a secure development project that discusses various security models and models them towards web application security. This project opens a door for repositories that effectively help secure development by providing the best expert practices and their processes. This type of project is effective and systematic but unable to give foundation security functionality in web applications. The development of security is something that demands implementation from the basic level of development phases. To make it possible discussing and providing an artifact-based approach that gives an effective result is required.

The proposed paper tries to present a systematic step-wise framework based on the secure development of healthcare web applications from an ontology perspective. Ontology is nothing but a novel systematic idea or concept that has some exceptionally effective potential to portray solid and effective results. The proposed model in the paper associates security factors from basic development phases and tries to manage them systematically further. The proposed model in the paper is a conceptual ontology-based model that aims to provide effective and secure development for web applications.

II. MATERIALS AND METHODS

A. Relevant Literature Analysis

Semantic analysis and its tools provide support to ontology-based security modeling and help them in achieving their goals. The concept of ontology makes them very specific and conceptual [4]. Searching the ontology-related literature for security perspective is not so hard due to its wide adaptation.

However, it is surprisingly challenging for authors to get a relevant healthcare web application development-based ontology approach. After a long and exhausting literature search, we could not find any specific literature based on ontology that talks about artifacts and other functional security attributes of development phases. To summarize other studies that associate the ontology concept effectively and discuss it towards security in various other parts similar field, we discuss some literature below:

In order to secure the software by adopting the ontology concept, researchers proposed a model by associating the SecEval framework and managing it towards software security [1]. The researchers combine the developed model with software engineering approaches and produce some effective results as an outcome. This literature motivates authors of the proposed work to choose the ontology for web application security.

As the next relevant study, we find an article that extensively manages various factors of the possible risk that cause bad influence on software security and manage them through various models discussed in the paper [2]. This type of model manages the risk and provides long serviceability as well as security in software. The result discussed in the paper has potential and good efficiency.

Further, as a case study, a research study associates ontology-based security modeling for user perspective and models their possible threads and its perspective [3]. Moreover, to facilitate software development security a study summarizes the facts of security modeling from an ontology perspective and gives an extensive analysis of them [4]. The results discussed in this article are effective and sensitive.

Furthermore, to provide development knowledge security and management the relevant study talks about various contexts-based security artifacts and their features respectively [5]. This type of model provides a contextual ontology approach that produces effective results.

The above-discussed literature discusses various perspectives and previous use of ontology. We try to understand their work properly and then portray an appropriate model that has novel utilities and working processes. It is very significant to adopt ontology in secure healthcare web application development because there is no specific work available currently that talks about web application security in healthcare or normally about any web application context.

B. Needs for Ontology-Based Model

Healthcare is a sector that deals with sensitive and huge amounts of data. It is a challenging and crucial task for experts to manage and secure the whole healthcare data management and web application properly due to its complexities. The digitization era forces healthcare services to be digitized via various online platforms and web applications. In such situations, it is very important to apply security strategies and processes efficiently. Before discussing the proposed model and idea it is an important job to talk about the origin of the idea and its need in a relevant field.

Healthcare web applications are applications that get penetrated frequently by attackers in current situations. Due to the sensitivity and value of health information, it is very normal and frequent for attackers to target and exploit healthcare web applications for valuable information breaches. A survey about healthcare data breach victims tells that the amount of victims from 2005 to 2019 is 249.09 million. Further, 157.40 million are affected only in the last few years [6]. Now, if we look beyond the reported attack scenario of data breach attacks the total number of attacks implemented and registered with governing bodies are 2216 from all around the 65 countries of world, and 536 breaches are from the healthcare industry in all of them. This type of statistics portrays the severity of the healthcare sector and gives an understanding that healthcare is the most targeted and sensitive field for attacks [7].

Moreover, there are a lot of breach-related statistics for healthcare web applications are available like: a report published by IBM in 2019 shows that the average revenue loss caused by attacks in all the fields and world is around 3.92 million dollars and if we look at the healthcare field specifically then the cost of a breach was 6.45 million dollars [8]. These statistics are automatically sufficient to portray the sensitivity and thread on healthcare web application security. Managing these issues and problems of healthcare is the most prioritized and immense demand of the situation.

To understand the previous healthcare records and breach statistics, we provide Table I and Fig. 1 which contain data related to previous year records and the data breach counts. This table portrays that healthcare data attacks are increasing day by day. The increasing use of web applications for providing health services creates a sensitive condition for security measures because of the complex nature of healthcare it is very challenging to manage every possible security measure on web applications. Attackers attain these loopholes as their weapons and target the system effectively and exploit the system on a large scale.

TABLE I. PREVIOUS ATTACK STATISTICS

Breach Year	Count	Amount of Data (In Millions)
2010	199	5.530
2011	200	13.150
2012	217	2.800
2013	278	6.950
2014	314	17.450
2015	269	113.270
2016	327	16.400
2017	359	5.100
2018	365	33.200
2019	505	41.200
Total	3033	255.18

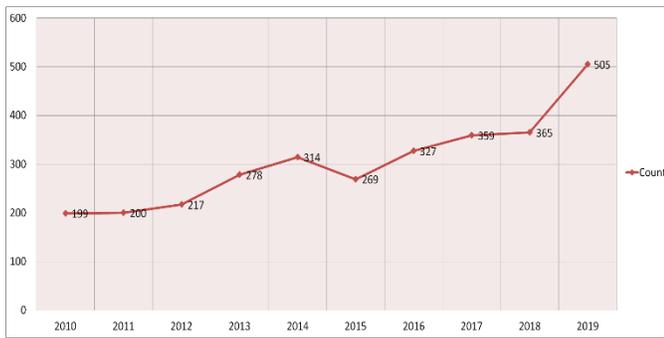


Fig. 1. Graphical Attack Illustration.

Table I and Fig. 1 portray data summarized and associated by HIPPA [9, 10]. The summarized information about the number of breach incidents and the number of breach records is totally surprising because these are officially registered and known cyber-attacks implemented on healthcare organizations. The amount of breach records is very high and significant because health is something that is totally and directly connected to the life of humans. A breach in this type of data can cause some serious life-threatening situations to the patients. Therefore, it is an immense need to manage web application security and attain security for healthcare.

Moreover, after understanding the attack statistics in healthcare it is a significant job to remediate its issues. To identify issues, it is necessary to find the loopholes first for healthcare breach incidents. To make it simple, we portray the various sources of breach registered and identified by experts in the following Table II. The following table represents the specific counts of breach incidents executed by any specific source that gives a clear point-wise understanding of healthcare loopholes and helps the experts to prepare preventive measures.

In Table II, it is clearly described and represented that technical issues are the most common and frequent exploit creators in healthcare. However, if we look at both Tables I and II comparatively then the data analysis tells that web application vulnerabilities cause maximum issues and exploits in healthcare in need of an instant solution. Thus, it is evident that technical faults cause and create issues in the current era and attackers grab these issues as an advantage and inject applications of healthcare very frequently on a large scale.

All in all, the whole statistics and attack data represented in this paper portray the current serious and sensitive situation of healthcare web application management as well as it also poses the need and requirement of functionality-based security measures that pillar the security of healthcare digitization from the deep down bottom. That is why the proposed article effectively aims to apply an ontology-based model that gives additional affectivity and conceptual understanding to the experts about how to build healthcare web applications.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar.

C. Adopted Ontology-Based Idea

As stated in the introduction ontology is something that portrays conceptual systematic steps that help in producing a secure and systematic manner of process and development methodology. In the development context of healthcare web application ontology adaptation is a novel that is why it is important to draw a proper connection between ontology concept and healthcare web application development. Developing a context-based model for effective security and efficiency in any type of process can only be achieved from ontology ideas [2].

Ontology ideology and concept are universally adopted and used in various fields like new mobile communication development and real human-machine interaction, etc. [11-15]. Ontology is nothing but simply a conceptual relationship of two ideas on one platform. This type of approach helps experts to relate the real live world issues into machine scenarios and portray or solve them through machines. Following figure 2 portrays the adopted ideology of ontology and tries to portray an overview of working.

TABLE II. SOURCE BASED ATTACK STATISTICS

Year	Technical Issues & Causes	Human Error & Disclosure	Theft	Miss Handling
2010	8	8	148	10
2011	17	27	136	7
2012	16	25	138	8
2013	25	64	150	13
2014	35	76	143	12
2015	57	101	105	6
2016	113	129	78	7
2017	147	128	73	11
2018	158	143	55	9
2019	274	142	51	7
Total	850	843	1077	90

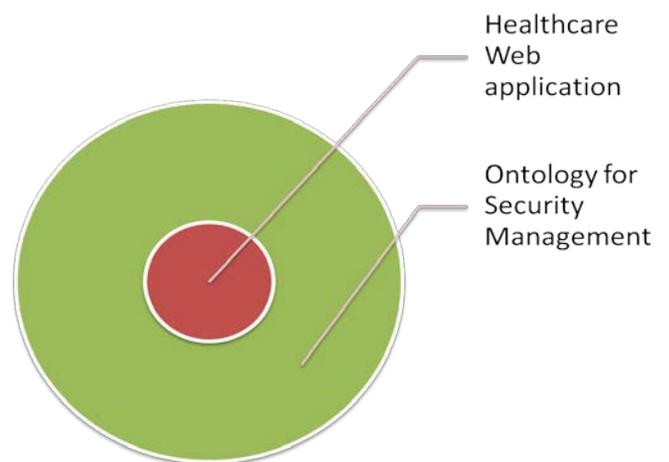


Fig. 2. Adopted Ontology Concept.

Fig. 2 illustrates the adopted concept for effective security measures in healthcare web application security. The core part of the figure represents the healthcare web application and its development and the outer layer of the figure portrays the ontology-based security management methodology that is going to be implemented on the healthcare web application development. Another issue that causes problems during development is the extensive and huge amount of information available on various online and offline resources that are adopted and used by developers during development steps to make their work easier. But did not identify that these adopting concepts are perfectly developed for their field or not. Development is a phase that demands only specific customized steps and utilities developed specifically for relevant processes; any other approach or utility can cause issues. Hence, the adopted concept for developing a more effective and secure model in healthcare web applications development is facilitated effectively by the proposed model in the article. The developed model based on this approach is discussed in the next section.

Moreover, the adopted methodology works on a conceptual ideology that tells about combining ontology-based security management with the classical widely adopted development phases for making it more systematic and secure without moving to a secure development process. There are various secure development ideologies available that are complex and very costly in implementation for organizations. In this type of situation, the proposed model by adopted ideology can produce some effective and good results and portray normal security managed overview of the healthcare web application.

III. PROPOSED MODEL

Modeling a systematic conceptual model for healthcare web application development by associating an ontology-based approach is a challenging task. There are various development guidelines and best practices are available but still it is hard for developers and designers to adopt them specifically for their development scenario [16]. This type of condition was created just because of the complexity and quick demand for applications in healthcare. The competitive market of web application development creates a challenging and risky culture of quick releasing in web application development. Business always tries to release their product or software quickly and this situation creates negligence in various fundamental aspects of designing.

Further, to tackle these situations in healthcare and provide them a systematic basic security model it is very important and necessary to portray the ontology ideology of concept development. Developing a concept before totally addressing it is the basic demand of any new approach. Concept-based models as proposed in this paper open a door for researchers to attain new ideas and develop their own by associating existing. In order to do this in the proposed article, we provide a classical development step-based model with ontology-based security management which is unique on its own.

Therefore, Fig. 3 portrays a five-step combined systematic model of healthcare web application development that associates classical development steps and ontology-based security management functionality. The proposed model

associates idea initiation to facilitate the phase of the web application with a blend of ontology-based security management. The conversation is something that can be referred to as comparative analysis in cybernetics [17] and managing the security always demands conversation in-between outcome and previous steps.

By focusing on this concept, a model is created that comparatively analyzes the requirement, design, and required the desired outcome of the developer to the same platform for producing effective results. The proposed model in this paper has five steps as Idea Understanding; Requirement Identification; Design & Code; Classification of Potential Possible Threats; Facilitate. All these steps are treated as classical none extra specific steps that demand extra from the developer or business.

As described and displayed in Fig. 3 the model has five fundamental steps for healthcare web application development. These steps work systematically one by one. We try to make it simple and non-complex because complexity is something that creates confusion as well as develops a lack of interest from the developer's side for the development process. Further, it is very important to understand the step-by-step working process of the proposed model that is described in the next heading of this paper.

A. Working Process of Proposed Model

A brief description of every step in the proposed model is described and discussed below.

a) *Idea understanding*: It is a step that initiates ideas about development. In this phase, the developer, coder, and owner of the web application discuss functionality and ideas about how the web application is going to be after development and what kind of functions the owner demands.

b) *Requirement identification*: After discussing and analyzing the idea of a web application it is necessary to identify which type of process, utilities, and qualities need attention during the development and design of the application. The requirement identification phase associates these demands and prepares a proper concept of relevant development for the systematic development of the application.

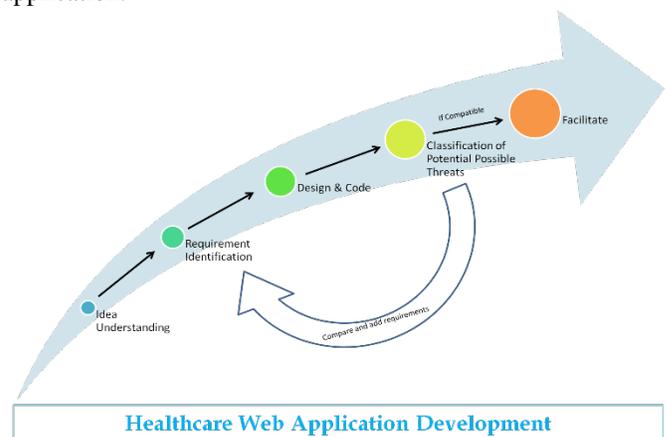


Fig. 3. Proposed Ontology-based Security Assured Model.

c) *Design and Code*: This is a step where developers design and code the application by associating identified requirements and prepare a systematic application according to them.

d) *Classification of Potential Possible Threats*: This is a step that is responsible for prevention from exploits in the system. It is a step that analyzes three perspectives and then compares the current development scenario. If the scenario was compatible according to the demand, then it's okay otherwise it again starts the loop from the requirement identification process. The whole phase description and internal working process are displayed in the following Fig. 4.

e) *Facilitate*: After comparing and classifying various threats and other issues of development when the whole cycle gets systematic without any interruption the proposed model allows developers to facilitate the developed healthcare web application in the industry.

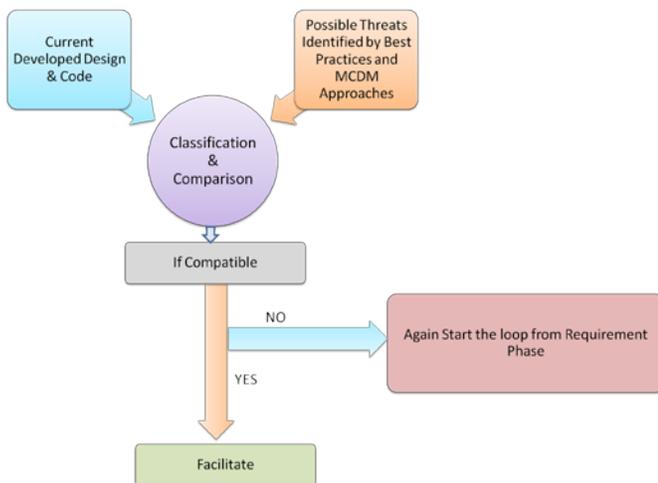


Fig. 4. Classification Phase of Proposed Model.

All in all, after discussing the concept, model, and working flow it is clear that the described model is the first initial step to facilitate healthcare web application development through an ontology-based approach. The proposed model portrays only the conceptual modeling of the idea and requires proper mathematical experimentation and validation in the future.

IV. DISCUSSION

The current healthcare digitization process demands various advanced requirements and utilities [18]. At the same time, increased attack vectors and processes created a situation where every type of post-developed prevention mechanism for breach attacks in healthcare gets destroyed by attackers. Now to tackle these situations experts and researchers believe that pre-security measures during the fundamental development phases can play a key role in healthcare security. Therefore, to facilitate this idea, and believe the authors of the proposed articles portray a conceptual model of healthcare systematic development by attaining a focus on security factor from the ontology-based approach.

V. CONCLUSION

In this rapid era of digitization healthcare services demand a new, novel, and fast development approach with a blend of security measures. To achieve this desired goal, we find that it is challenging and brainstorming work for them to facilitate the security measures after the development of healthcare web applications. Therefore, they think from different perspectives and manage the goal by applying ontology ideas and development phases at the same platform. This type of combination gives exclusive power to the authors about discussing and managing systematic development and security both at the same time for healthcare web applications. The proposed article portrays a model for ontology-based secure healthcare web application development. The proposed model has five simple but effective phases which aim to give a systematic secure development pathway to the healthcare web application developers. The proposed model has various beneficial advantages for healthcare web application security and portrays a pathway for future researchers to facilitate systematic development.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (references).
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, "Title of paper if known," unpublished.
- [5] R. Nicole, "Title of paper with only first word capitalized," *J. Name Stand. Abbrev.*, in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
- [7] M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.
- [8] IBM. Available online: <https://www.ibm.com/security/data-breach> (November, 2020).
- [9] HIPAA Journal. Available online: <https://www.hipaajournal.com/december-2019-healthcare-data-breach-report/> (November, 2020).
- [10] HIPAA Journal. Available online: <https://www.hipaajournal.com/largest-healthcare-data-breaches-of-2016-8631/> (November, 2020).
- [11] Alenezi, M., 2020. Ontology-Based Context-Sensitive Software Security Knowledge Management Modeling. *International Journal of Electrical and Computer Engineering (IJECE)*, 10(6), pp.6507-6520.
- [12] Kumar, R., Alenezi, M., Ansari, M.T.J., Gupta, B.K., Agrawal, A. and Khan, R.A., 2020. Evaluating the impact of malware analysis techniques for securing Web applications through a decision-making framework under fuzzy environment. *Int. J. Intell. Eng. Syst.*, 13(6), pp.94-109.
- [13] Zarour, M., Ansari, M.T.J., Alenezi, M., Sarkar, A.K., Faizan, M., Agrawal, A., Kumar, R. and Khan, R.A., 2020. Evaluating the impact of blockchain models for secure and trustworthy electronic healthcare records. *IEEE Access*, 8, pp.157959-157973. Bishop, M. "A Clinic for 'Secure' Programming." *IEEE Security & Privacy* 8.2 (2010).

- [14] Alenezi, M., Basit, H.A., Khan, F.I. and Beg, M.A., 2020. A Comparison Study of Available Software Security Ontologies. In Proceedings of the Evaluation and Assessment in Software Engineering (pp. 499-504). R. Kumar, A. Agrawal, and R. A. Khan, (2020), A wakeup Call to Data Integrity Invulnerability, Computer Fraud & Security, Volume 2020, Issue 4, pp. 14-19. Elsevier. Available at Thomson Reuters. DOI: [https://doi.org/10.1016/S1361-3723\(20\)30042-7](https://doi.org/10.1016/S1361-3723(20)30042-7).
- [15] Dubberly, H., & Pangaro, P. (2019). Cybernetics and design: Conversations for action. In Design Cybernetics (pp. 85-99). Springer, Cham.
- [16] Javed, Y., Arian, Q. A., & Alenezi, M. (2021). SecurityGuard: An Automated Secure Coding Framework. In Intelligent Technologies and Applications: Third International Conference, INTAP 2020, Grimstad, Norway, September 28–30, 2020, Revised Selected Papers 3 (pp. 303-310). Springer International Publishing.
- [17] Grinin, L., & Grinin, A. (2020). The cybernetic revolution and the future of technologies. In The 21st Century Singularity and Global Futures (pp. 377-396). Springer, Cham.
- [18] Zarour, M., Alenezi, M., Ansari, M. T. J., Pandey, A. K., Ahmad, M., Agrawal, A., ... & Khan, R. A. (2021). Ensuring data integrity of healthcare information in the era of digital health. Healthcare Technology Letters, 8(3), 66.

A Proposal to Improve the Bit Plane Steganography based on the Complexity Calculation Technique

Cho Do Xuan

Information Assurance Department
FPT University, Hanoi
Vietnam

Abstract—The video steganography technique is being studied and applied a lot today because of its benefits. In particular, the video steganography technique using Bit-Plane Complexity Segmentation (BPCS) has increasingly proven its effectiveness compared to other methods. In this paper, based on the theoretical basis of the BPCS method, we propose a new method to improve the efficiency of the steganography process. Accordingly, our improvement proposal in this paper is improving the complexity formula of the bit planes. Our new formula not only has improved the steganographic thresholds in the bit planes to find more planes hiding secret information, but also has ensured the amount of information hidden in the video and their safety. The experimental results in the paper have not only demonstrated the effectiveness of our proposed method but also provided a new mechanism for digital image analysis in general and video steganography techniques in particular.

Keywords—Steganography; video steganography technique; bit-plane complexity segmentation (BPCS); complexity formula

I. INTRODUCTION

A. Introduction to BPCS Method

In the study [1], Kawaguchi et al. presented an image steganography method based on the BPCS technique. The characteristic of this method is the use of images or video frames as the message vessel. Accordingly, the image is divided into bit planes based on the depth value of the image. With each bit-plane, we can divide it into noise-like blocks or informative blocks, then replacing noise-like by the blocks of secret information that have similar noisy property will not change the image quality. The process of embedding information in video using BPCS technique includes the following steps [1]:

1) *Determining the noise-like block.* Fig. 1 illustrates an example of the noise-like block. To determine noise-like block, the study [1] proposed Black-White border method. Accordingly, in Black-white border, the complexity is determined by the length of the border of the Black-White regions in blocks horizontally and vertically. For example, a black pixel surrounded by 4 white pixels has a border length of 4. The longer the block has the border length, the higher the complexity. Based on the above definition, we have the following formula to determine the complexity of a block with size $2n * 2n$ [1]:

$$\alpha = \frac{k}{2 * 2n * (2n - 1)} \quad (1)$$

Where:

- α is the complexity of the block.
- k is the number of borders in contact between 2 regions.
- $2 * 2n * (2n - 1)$ is the number of borders that are in contact maximum possible with the block. It is the number of borders in the chessboard.

2) *Identifying complexity threshold:* The process of identifying the complexity threshold to determine how much complexity is, the region is noise, and how much complexity is, the region is informative. To do that, the researchers tested on blocks of size $8 * 8$. The use of blocks $8 * 8$ is to match the complexity of the current method. For blocks $8 * 8$, the average value of α is 0.5, the informative blocks have α in between 0 and 0.5 and only accounted for $6.67 * 10^{-14} \%$. From these values, one normally choose a complexity threshold as $\alpha_0 = 0.3$ for block $8 * 8$ [1].

3) *The conjugation property.* This property ensures the safety of the information that needs to be hidden in the noise-like regions. Accordingly, when information is divided into appropriate blocks and is calculated complexity, if it is an informative block, it will be conjugated to be a block with higher noise for embedding information. The conjugation property is stated as [1]: With a block P with complexity α , there exists only a block conjugate of P denoted P* has complexity is $\alpha' = 1 - \alpha$. This block is an XOR result between P and the block W_c (alternating black and white blocks starting with white pixel).

4) *Canonical gray coding:* This is the process of processing images to embed information. Accordingly, suppose that b_i and g_i are respectively the i -th bit of the PBC and CGC codes with n -bits of data we have [2].

$$b_0 b_1 \dots b_{n-1} \text{ and } g_0 g_1 \dots g_{n-1}$$

The formula for switching between two coding systems is:

$$g_i = \begin{cases} b_0 & \text{Where } i = 0 \\ b_{i-1} \text{ xor } b_i & \text{Where } i > 0 \end{cases} \quad (2)$$

$$b_i = \begin{cases} g_0 & \text{Where } i = 0 \\ b_{i-1} \text{ xor } b_{i-1} = g_0 \text{ xor } \dots \text{ xor } g_i & \text{Where } i > 0 \end{cases}$$

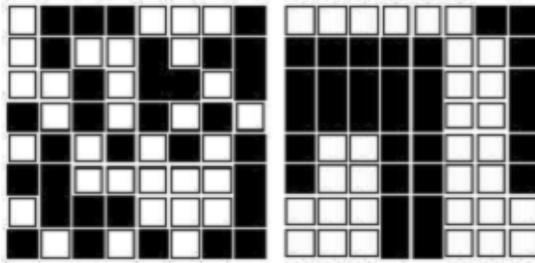


Fig. 1. Examples of Bit Planes for Determining the Noise-like Block.

B. The Problem of Improving the BPCS Method

Based on the process of processing and embedding information, it can be seen that improving the BPCS method is based on two main methods [2, 3, 4, 5]:

- Improvement in the pre-processing process. This is an improved method to be performed at the pre-processing step. The characteristic of this method is the modification of the input data (namely the image used for embedding) in order to make embedding more efficient. Some of the approaches are: Use of Gray coding; handling the vessel to increase the image sharpness; randomizing input data.
- Improving the embedded algorithm. The approaches usually focus on two main improvement methods: i) Improvement of complexity threshold. This approach mainly refers to 2 improvement methods: the complexity threshold for each bit plane and the Dynamic Threshold; ii) Improvement of the formula. There are 2 formulas currently used as Run-Length Irregularity and Border Noisiness.

C. Contributions of the Paper

Our research is presented as follows: the urgency of the research problem is presented in Section I. In Section II, we present the process of researching, surveying, and evaluating related works. In Section III, we present our surveys and reviews of problems of proposing improvements for BPCS. This review and survey help us have a basis to propose our new method in the paper. Finally, in Section IV, we present the experimental method to evaluate and compare the effectiveness of our proposed method with other methods. The practical significance and scientificity of our paper include:

- Proposing improvements for video steganography techniques based on the approach about the calculation formula. Accordingly, we propose a new calculation formula to evaluate the complexity of the frames. Details of the proposed algorithm and its feasibility assessment are described in section III.C of the paper.
- We conduct experiments, evaluate, and compare to see the effectiveness of our approach with other research directions. The experimental results in section IV.B.2 proved the correctness of our method. At the same time, the research results in the paper also provide a new approach to calculating the complexity of the image in order to serve for the analysis of digital images in general and hiding information by the BPCS method in particular.

II. RELATED WORKS

A. BPCS Technique and Some Improvements

The study [5] proposed the Modified BPCS technique by combining hybrid cryptography algorithms between RSA and DES to create noise for secret messages. Piyush and Paresh [6] proposed a combined method of cryptography, steganography, and multimedia data hiding. In this method, the authors used a reference database to provide higher security levels. First, the authors used the DES algorithm to encrypt the message, then using a modified bit encoding technique to save the cipher in the image. For each byte of data, one cover pixel will be edited. The study [7] combined the AES cryptography algorithm and the BPCS steganography algorithm to hide a large amount of data in image. In the study [8], the authors proposed applying the BPCS method and FPGA model to ensure information security for secret information. In the study [9], the authors proposed combining the BPCS method with the Huffman cryptography algorithm to improve the quality of hiding information.

In the study [10], the authors proposed a method of combining BPCS with fuzzy logic techniques. This study applied the principles of fuzzy sets to classify the bit-plane into three sets: informative, partly informative, and noise region. This is expected to classify the bit-plane in a more objective approach. Finally, the Mamdani fuzzy inference is used to make decisions on which bit-plane will be replaced with a message based on the classification of bit-plane and the size of the message that will be inserted. This helps improve the message capacity of the images. The research could improve BPCS steganography techniques to insert a message in a bit-plane with more precision. Thus, the container image quality would be better. Seeing that the PSNR value of the original image and the stego-image is only slightly different. In addition, the studies [11, 12, 13, 16] listed some video steganography methods, difficulties, challenges as well as the advantages and disadvantages of video steganography techniques. Besides, Spaulding et al. [14] presented the BPCS steganography method with the embedded zerotree wavelet (EZW) lossy compression. This method used the DWT coefficients to represent pixels of the original frame. Therefore, the BPCS steganography can be applied to DWT coefficient sub-bands containing different features. Similarly, Noda et al. [15] proposed a video steganography technique using the BPCS and wavelet compressed video.

On the other hand, Sharma et al. [17] presented a number of research and evaluation results of the effectiveness and safety of BPCS and LSB techniques. The research results showed that in the BPCS technique, all the data was embedded in a complex noisy blue plane. The embedding capacity of BPCS techniques and LSB techniques is high. On comparison, it is found that the LSB steganography and LSB using the secret key perform the best on the basis of PSNR. According to the entropy and correlation point of view, the best results are shown by the status bit and BPCS steganography techniques.

B. Some Studies of Combining Encryption with Steganography

In the study [18], Shifa et al. proposed a method of combining the AES encryption with LSB steganography

technology in order to distribute and exchange keys. The experimental results show that the LSB model combined with AES encryption has a good effect on the speed and impact degree on the message vessel. In the study [19], Dhall proposed the first method of using Quantum Cryptography to increase the security of the application. Saha [20] et al. published a steganography method based on Exploiting Modification Direction using the hashed-weightage Array. Kait [21] applied the BPCS technique as a technique to noise information in order to ensure data security using FPGA implementation. In addition, the studies [22-26] proposed to apply some common steganography techniques such as Discrete Cosine Transform (DCT), LSB, Least Significant bit Matched Revisited (LSBMR), DWT to protect data in communication applications.

III. EVALUATING AND RECOMMENDING IMPROVEMENTS OF BPCS

A. Improvement on Complexity Threshold

With basic BPCS, we usually choose only one complexity threshold for the whole bit-plane. The study [3] proposed a method to apply the complexity threshold for each bit plane. Accordingly, the authors demonstrated that the complexity threshold from the low bit plane will gradually decrease to zero. Usually, at the highest bit planes, no change will be made here because changing the bit here will strongly affect image quality. On the contrary, the lower bit planes have a higher threshold to increase the ability to embed information.

The study [3] proposed improvement direction for the BPCS algorithm based on Dynamic Threshold. This method makes some adjustments to the threshold setting. These threshold values have no default values. They can be adjusted manually according to the actual conditions. This algorithm can avoid the analyzer from detecting the existence of secret information using complex statistical analysis of the entire graph, thus further enhancing the security of steganography. The basis of the Dynamic Threshold improvement method is based on the Chaos theory. Accordingly, the Chaos theory is a type of behavior controlling nonlinear dynamical rule. In this research, the authors used the logistic mapping method to create chaos series according to the formula:

$$ak+1 = \mu * ak*(1 - ak), k= 0,1,2. \quad (3)$$

The moving value is in the range [0, 1] and μ is a control parameter or a split parameter. When $3.5699456... < \mu \leq 4$, logistic maps operate in a chaotic state [3]. The generated data stream is disordered and it is similar to random noise. Processing the obtained chaos series, mapped to {0, 1} and {-2, -1, 0, 1, 2}.

B. Improving the Formula Determining the Complexity

To assess whether the blocks are complex or not, the studies are based on the complexity of the black-and-white border region. However, it is not always the best approach. For example, blocks with periodic patterns like the chessboard will be recognized for complexity in this way as shown in Fig. 2 $\alpha_a = 1$ and $\alpha_b = 1/2$.

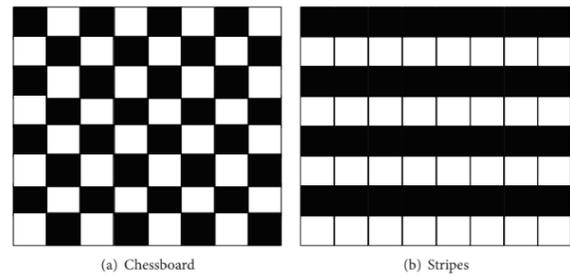


Fig. 2. Example of Blocks with Large α but those are not Complicated.

1) *Determining the noise-like block:* The Run-Length Irregularity is the histogram of the lengths of both black and white pixels in a row or a column [2]. Assume that:

- $h[i]$ is the frequency of runs of i pixels either in black or in white.
- n is the length of the sequence of pixels.
- h_c to measure the irregularity of a binary pixel sequence.

$$P_i = \frac{h[i]}{\sum_{j=1}^n h[j]} \quad (4)$$

The value $h_s \in [0; 1]$ and denoted by:

$$h_s = \sum_{i=1}^n h[i] \log_i p_i \quad (5)$$

If the size of the block is $n \times n$ and r_i and c_j are respectively the i -th row and j -th column of a block, then the run-length irregularity β of a block is defined with \bar{X} which is the mean of all the elements of X .

$$\beta = \min\{\widehat{H_s(r)}, \widehat{H_s(c)}\}$$

$$\widehat{H_s(r)} = \{\widehat{h_s}(r_0), \dots, \widehat{h_s}(r_{n-1})\}$$

$$\widehat{H_s(c)} = \{\widehat{h_s}(c_0), \dots, \widehat{h_s}(c_{n-1})\}$$

If according to the definition, the smaller row and column averages are taken as the value of the run-length irregularity β . As shown in Fig. 3, they are both periodic in row or column. The result is that every run-length irregularity β is 0, so they are simple and cannot be used for embedding information. The run-length irregularity β is only useful in rows or columns. If the block is frequently in other directions, there will be nothing to do with it.

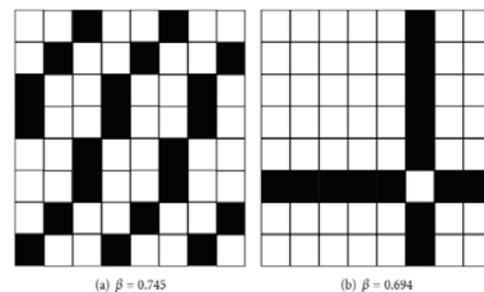


Fig. 3. Examples of Blocks with Large β but those are not Complicated.

2) *Border noisiness*: If the data is embedded on the boundary of the noise region and the informative region of the block in the image, then the noisy regions would grow. As a result, the image will be evidently changed. Border Noisiness is based on the difference between adjacent binary pixel sequences in a block. In a similar way, if the block size is $n \times n$ and r_i and c_j are respectively the i -th row and j -th column of a block, then the *Border Noisiness* γ of a block is defined as follows [2]:

$$\gamma = \frac{1}{n} \min\{E_f[P_x(r), E_f[P_x(c)]]\} \quad (6)$$

Where $\rho(x)$ is the number in the binary string X and

$$P_x(r) = \{\rho(r_0 \oplus r_1), \dots, \rho(r_{n-2} \oplus r_{n-1})\}$$

$$P_x(c) = \{\rho(c_0 \oplus c_1), \dots, \rho(c_{n-2} \oplus c_{n-1})\}$$

$$E_f(X) = \frac{1 - V(X)}{\max_X\{V(X)\}} \cdot \bar{X}$$

Where: $X = \{x_0, \dots, x_{m-1}\}$, $V(X)$ is the variance of X , and \bar{X} is the mean of X .

Border Noisiness is used to check if many black and white pixels are well distributed over a block along with both horizontally and vertically. Although blocks in Fig. 3(a) and 3(b) both have large run-length irregularity β ($a = 0.745$, $b = 0.694$), their Border Noisiness are 0.294 and 0.048 respectively. Therefore, they are not complicated according to Border Noisiness and are not suitable for embedding data.

C. Our Proposal

We noticed that there is no exact formula or judgment about the definition of complexity for an information block. The black-white border is just one of many formulas to determine this complexity. This formula is simple, useful, and gives the best results. But it still has some shortcomings. Therefore, we propose a novel method to calculate complexity. This method is quite similar to the Black-white border but it exploits another aspect that is the number of black and white regions in the block (the Black-white border uses the number of black and white borders). This formula is based on the assumption that the higher the number of black-and-white regions a block has, the higher the complexity of the block is and vice versa. From the above assumption, we propose the formula for calculating the complexity for a block of size $2n * 2n$ as follows:

$$\partial = \frac{k}{2n * 2n} \quad (7)$$

Where: k is the number of black and white regions in the block; $2n * 2n$ is the maximum number of black and white regions that a block of size $2n * 2n$ can receive. To verify this assumption and to show the better aspects of this formula than the previous formulas, we will conduct a comparison with some blocks (see Fig. 4 and Fig. 5, 6).

From Fig. 4(a) we calculate the following values: $\alpha = 0.42$; $\beta = 0.198$; $\partial = 0.04$.

Fig. 4(b) gives the following values: $\alpha = 0.42$; $\beta = 0.123$; $\partial = 0.156$.

Fig. 4(c) gives the following values: $\alpha = 0.58$; $\beta = 0.745$; $\partial = 0.14$.

Fig. 4(d) gives the following values: $\alpha = 0.285$; $\beta = 0.694$; $\partial = 0.14$.

With the naked eye, we can see which are informative blocks with clearly divided black and white regions. However, with the original formula, these images give high results, in contrast to the Border Noisiness formula and the formula that we proposed. Fig. 5 below is an example that shows a weak point of Border Noisiness. On the other hand, the proposed way still indicates that these are informative blocks.

From Fig. 5(a), we calculate the following values: $\alpha = 1$; $\beta = 0$; $\partial = 1$.

Fig. 5(b) gives the following values: $\alpha = 1$; $\beta = 0$; $\partial = 0.07$.

Can be seen that: Although giving good results with the previous blocks, but up to the chessboard block, the proposed method still gives high results. But for stripe block, it still gives low results in contrast to the Black-white border. Through the above examples, we can see that our proposed method gives better results than the original algorithm. Next, to evaluate the feasibility of the complexity formula, we will conduct the experiment of embedding information using the BPCS method and evaluate the embedded results based on measures.

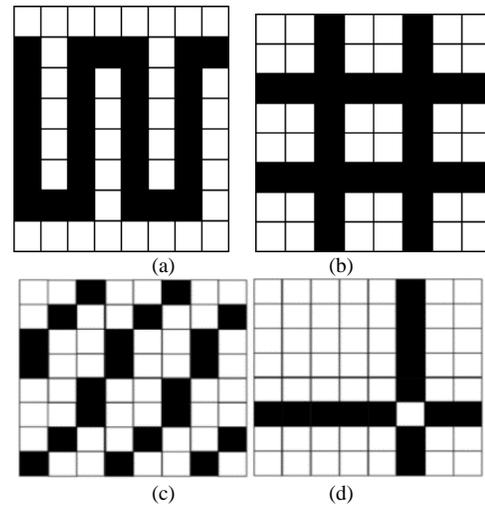


Fig. 4. Examples of Noise Blocks.

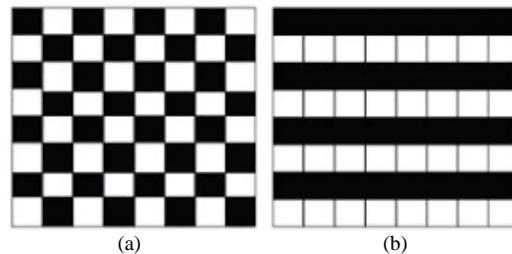


Fig. 5. Example of Informative Blocks showing the Weakness of Border Noisiness.

IV. EXPERIMENTS AND EVALUATION

A. Evaluation Criteria

To conduct experiments and evaluations, we use a number of criteria to evaluate as follows.

1) *Peak signal-to-noise ratio*: The Peak signal-to-noise ratio (PSNR) is a measure of the difference between the original image and the stego image. Comparing the results of two algorithms should be based on the same codecs and the same data content. The simplest way is defining through the mean squared error (MSE) used for 2-dimensional images with size $m \times n$ where I and K are the original image and the stego image [4].

$$MSE = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (C_{ij} - R_{ij})^2 \quad (8)$$

PSNR is defined by [4]:

$$PSNR = 10 \log_{10} \frac{(2^b - 1)^2}{MSE} \quad (9)$$

2) *Embedding capacity*: Embedding capacity is the maximum amount of data that an image can be successfully embedded by a steganography algorithm. This criterion depends on the input image data and the number of embedded bits [4].

$$bpp = \frac{\text{hidden bits}}{\text{total bits}} \quad (10)$$

B. Experiment And Evaluating Formula Improvement Methods

1) *Scenario and experimental data*: To evaluate the effectiveness of our proposed formula, we will do experiments on some images with a resolution of $512 * 512$, which is the resolution commonly used for hiding information. Fig. 6 below presents the images used to hide information.

2) *Experimental results*: Table I below shows the distribution of thresholds. From there, we select a complexity threshold that balances image quality and memory capacity.

Through the statistics in Table I and Fig. 7, we can see that with the proposed formula, blocks $8 * 8$ usually have a value in the range from 0 to 0.2. This proves that the majority of blocks have threshold values from 0 to 0.2. The distribution diagram is similar to that of the classic formula. The only difference is that the equilibrium threshold is 0.2 and the threshold of the classic formula is 0.5. With classic formula, the threshold will be selected from 0.3 to 0.5 so we can speculate that the threshold of the improved formula is from 0.12 to 0.2. With this range, we can also see the ability to hide is about 50%. With the selected complexity threshold of 0.14, the experimental process will proceed as follows. Specifically, to test the maximum ability to hide information, we will use 3 images in Fig. 6 as vessels, and then we check the maximum ability to hide information on each photo to give the results of the ability to embed. Besides, to test PSNR, we will embed the

Beagle.png image (Fig. 8) to produce the results and calculate PSNR between the original image and stego image.

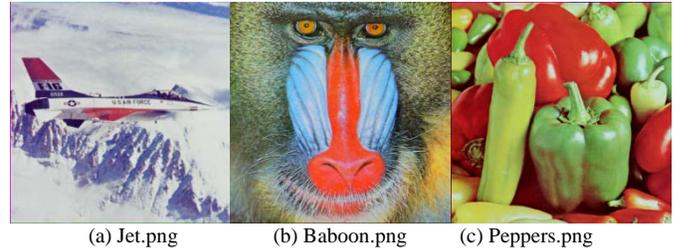


Fig. 6. Three Experimental Images.

TABLE I. THRESHOLD DISTRIBUTION OF THE IMAGES IN FIGURE 6

Capacity	Baboon.png	Peppers.png	Jet.png
0.05	92.85%	64.59%	60.56%
0.1	80.74%	52.48%	48.44%
0.15	56.52%	44.41%	40.37%
0.2	40.37%	24.22%	12.11%
0.25	20.18%	0.08%	0.08%
0.3	0.04%	0.00%	0.00%

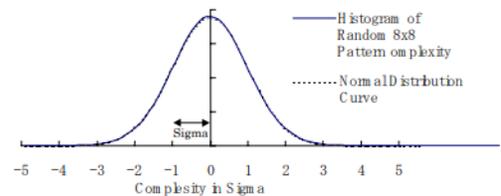


Fig. 7. Distribution Diagram of the Thresholds Calculated by the Classical Formula [1].



Fig. 8. Beagle.png

3) *Experimental results*: From the results in Tables II and III, it can be seen that the ability to hide information of improved BPCS, which we propose, is much better than classic BPCS with a complexity threshold of 0.4. The ability to hide information will increase by about 10% to 15% of the image capacity. Although this ability increases significantly, when compared to the PSNR index, it is not much different from the original algorithm. Although the PSNR index decreases by about 0.01, increasing the ability to hide information brings great benefits. This can be seen as a valuable improvement. Here we can increase the complexity threshold to 0.15 or 0.145 in exchange for a higher PSNR depending on the intended use.

TABLE II. COMPARING THE ABILITY TO HIDE INFORMATION OF EACH METHOD

Capacity	BPCS [1] $\alpha = 0.4$	BPCS [our proposal] $\delta = 0.14$
Baboon.png	3503297 bit = 55.68 %	3810240 bit = 60.56 %
Peppers.png	2612544 bit = 41.52 %	3048192 bit = 48.45 %
Jet.png	1821056 bit = 28.94 %	2540160 bit = 40.37 %

TABLE III. COMPARING IMAGE QUALITY AFTER HIDING INFORMATION

Capacity	BPCS [1] $\alpha = 0.4$	BPCS [our proposal] $\delta = 0.14$
Baboon.png	53.43	53.42
Peppers.png	53.44	53.43
Jet.png	53.44	53.43

V. CONCLUSION

Applying the BPCS steganography technique to hide secret information or protect vessels is one of the best current approaches. In this paper, based on the procedure and the mathematical basis of the BPCS steganography technique, we proposed an innovative method to improve the quality of hiding information. The experimental results, which we performed in Tables II and III, proven that the method of improving the complexity calculation algorithm in bit planes gave good results not only for the ability to hide information but also to ensure the image quality. The research results will allow having many criteria to select the bit planes based on calculating their complexity for hiding information. At the same time, our computation proposal will also provide a new way for problems related to analyzing digital images in the face of the diversity and rapid development of the multimedia field. In the future, we will apply many other improved methods, especially the applying of different thresholds for each plane bits, in order to increase the capacity to hide information as well as image quality when hiding.

REFERENCES

- [1] Eiji Kawaguchi, Richard O. Eason, "Principles and applications of BPCS steganography," *Multimedia Systems and Application*, vol. 35(2), pp. 464-473, 1999.
- [2] Peipei Shi, Zhaohui Li, "An improved BPCS steganography based on dynamic threshold," *Proceedings of International Conference on Multimedia Information Networking and Security*, Nanjing, Jiangsu, China, pp. 387-391, 2010.
- [3] Shuliang Sun, "A New Information Hiding Method Based on Improved BPCS Steganography," *Advances in Multimedia*, 2015.
- [4] Vipul J Patel, Ms. Neha Ripal Soni, "Uncompressed Image Steganography using BPCS: Survey and Analysis," *OSR Journal of Computer Engineering*, vol. 15, pp. 57-64, 2013.
- [5] Smita P. Bansod, Vanita M. Mane, R. Ragha, "Modified BPCS steganography using Hybrid cryptography for improving data embedding capacity," *Proceedings of International Conference on Communication, Information & Computing Technology (ICCICT)*, Mumbai, India, pp. 1-6, 2012.
- [6] P. Marwaha, "Visual cryptographic steganography in images," *Proceedings of Computing Communication and Networking Technologies (ICCCNT)*, Karur, India, pp. 1-6, 2010.
- [7] M. Goljan, J. Fridrich, R. Du, "Distortion-free data embedding," *Proceedings of 4th Information Hiding Workshop*, pp. 27-41, 2001.

- [8] Vikas S. Kait, Bina Chauhan, "BPCS steganography for data security using FPGA implementation," *Proceedings of 2015 International Conference on Communications and Signal Processing (ICCSPP)*, Melmaruvathur, India, pp. 1887-1891, 2015.
- [9] Tayal, N., Bansal, R., Dhal, S. et al, "A novel hybrid security mechanism for data communication networks," *Multimed Tools Appl*, vol. 76, pp. 24063-24090, 2017.
- [10] Rahmad Hidayat, "Klasifikasi Bit-Plane Noise untuk Penyisipan Pesan pada Teknik Steganography BPCS Menggunakan Fuzzy Inference Sistem Mamdani," *Jurnal Rekayasa Elektrika*, vol 11, pp. 101- 108, 2015.
- [11] Ramadhan J. Mstafa, Khaled M. Elleithy, Eman Abdelfattah, "Video steganography techniques: Taxonomy, challenges, and future directions," *Proceedings of 2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, Farmingdale, NY, USA, pp. 1-6, 2017.
- [12] Mstafa R.J., Elleithy K.M., "Compressed and raw video steganography techniques: a comprehensive survey and analysis," *Multimed Tools Appl*, vol. 76, pp. 21749-21786, 2017.
- [13] Gerrit Cornelis LANGELAAR, "Real-time Watermarking Techniques for Compressed Video Data," *Veenendaal*, ISBN 90-9013190-6, 2000.
- [14] Spaulding J, Noda H, Shirazi MN, Kawaguchi E, "BPCS steganography using EZW lossy compressed images," *Pattern Recogn Lett*, vol. 23, pp. 1579-1587, 2002.
- [15] Noda H, Furuta T, Niimi M Kawaguchi E, "Application of BPCS steganography to wavelet compressed video," *Proceedings of 2004 International Conference on Image Processing ICIP'04*, Singapore, pp. 2147-2150, 2004.
- [16] Y Liu, S Liu, Y Wang, H Zhao, S Liu, "Video steganography: A review," *Neurocomputing*, vol. 335(28), pp. 238-250, 2019.
- [17] Sharma, Rinku Ganotra, Reema Dhall, Sangeeta Gupta, Shailender, "Performance Comparison of Steganography Techniques," *International Journal of Computer Network and Information Security*, vol. 10, pp. 37-46, 2018.
- [18] S.A. Afgan et al. "Joint Crypto-Stego Scheme for Enhanced Image Protection with Nearest-Centroid Clustering," *IEEE Access*, vol. 6, pp. 16189-16206, 2018.
- [19] S. Dhall, R. Sharma, S. Gupta, "A multi-level steganography mechanism using quantum chaos encryption," *Multimed Tools Appl*, vol. 79, pp. 1987-2012, 2020.
- [20] Saha, S., et al. "Extended exploiting modification direction based steganography using hashed-weightage Array," *Multimed Tools Appl*, vol. 79, pp. 20973-20993, 2020.
- [21] V. S. Kait, B. Chauhan, "BPCS steganography for data security using FPGA implementation," *Proceedings of 2015 International Conference on Communications and Signal Processing (ICCSPP)*, Melmaruvathur, pp. 1887-1891, 2015.
- [22] G. L. Smitha, E. Baburaj, "A survey on image steganography based on Least Significant bit Matched Revisited (LSBMR) algorithm," *Proceedings of 2016 International Conference on Emerging Technological Trends (ICETT)*, Kollam, pp. 1-6, 2016.
- [23] A. Y. AlKhamese, W. R. Shabana, I. M. Hanafy, "Data Security in Cloud Computing Using Steganography: A Review," *Proceedings of 2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, Aswan, Egypt, pp. 549-558, 2019.
- [24] G. Maji, S. Mandal, N. C. Debnath, S. Sen., "Pixel Value Difference Based Image Steganography with One Time Pad Encryption," *Proceedings of 2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, Helsinki, Finland, pp. 1358-1363, 2019.
- [25] S. Timarchi, M. A. Alaei, H. Koushkbaghi, "Novel algorithm and architectures for high-speed low-power ConText-based steganography," *Proceedings of 19th International Symposium on Computer Architecture and Digital Systems (CADS)*, Kish Island, pp. 1-6, 2017.
- [26] G. Maji, S. Mandal, S. Sen, N. C. Debnath, "Dual image based LSB steganography," *Proceedings of 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom)*, Ho Chi Minh City, Viet Nam, pp. 61-66, 2018.

Multiplicative Iterative Nonlinear Constrained Coupled Non-negative Matrix Factorization (MINC-CNMF) for Hyperspectral and Multispectral Image Fusion

Priya K¹, Dr. Rajkumar K K²
Dept of Information Technology
Kannur University
Kerala

Abstract—Hyperspectral and Multispectral (HS-MS) image fusion is a most trending technology that enhance the quality of hyperspectral image. By this technology, retrieve the precise information from both HS and MS images combined together increase spatial and spectral quality of the image. In the past decades, many image fusion techniques have been introduced in literature. Most of them using Coupled Nonnegative matrix factorization (CNMF) technique which is based on Linear Mixing Model (LMM) which neglect the nonlinearity factors in the unmixing and fusion technique of the hyperspectral images. To overcome this limitation, we are going to propose an unmixing based fusion algorithm namely Multiplicative Iterative Nonlinear Constrained Coupled Nonnegative Matrix Factorization (MINC-CNMF) that enhance the spatial quality of the image by considering the nonlinearity factor associated with the unmixing process of in the image. This method not only consider the spatial quality but also enhance the spectral data by imposing constraints known as minimum volume (MV) which helps to estimate accurate endmembers. We also measure the strength and superiority of our method against baseline methods by using four public dataset and found that our method shows outstanding performance than all the baseline methods.

Keywords—Hyperspectral data; multispectral data; minimum volume; nonlinear mixing model; spectral variability; spectral image fusion

I. INTRODUCTION

Hyper-spectral (HS) images are enriched with high spectral resolution than conventional images. Therefore, the energy collected by hyper-spectral sensors is partitioned into several narrow wavelengths. Due to the partitioning of several narrow wavelengths band, the energy received by each band is limited. This makes the HS image to be easily influenced by many kinds of noise. This high signal-to-noise ratio reduces the spatial resolution of hyper-spectral image. So, it is mandatory to increase the spatial quality of the HS image. The fusion of high quality spatial data with the hyper-spectral image that possess high spectral resolution is one of the good approach for HS image enhancement [1].

Many existing traditional or multispectral sensors can capture images that have higher spatial resolution with lower spectral resolution compared with hyper-spectral sensors.

Therefore, data fusion with hyper-spectral (HS) and multispectral (MS) image is one good approach to increase the spatial quality of hyper-spectral image. This HS-MS fusion approach fuses the spectral data of low-spatial resolution hyper-spectral image (LR-HSI) with the spatial data of high-spatial resolution multispectral image (HR-MSI) of the same scene. Thus, generated a high-spatio-spectral resolution hyper-spectral image (HR-HSI) [2]. Some important HS-MS fusion problem is component substitution (CS), multiresolution analysis (MRA), Bayesian probability, Spectral Unmixing (SU). But recently, the Spectral Unmixing (SU) based HS-MS fusion become a trending attention in this area due to its straightforward description in the fusion process. In spectral unmixing a simple and effective method namely linear mixing models (LMM) are widely used in most of the literature [3].

In HS - MS data fusion approach, both LR-HSI and HR-MSI data are unmixed into spectral (endmember) and spatial (abundances) data [4]. Next, the high-spatial data of HR-MSI are fused with high-spectral data of LR- HSI using some constrained optimization techniques. Coupled Non-negative Matrix Factorization (CNMF) algorithm is a promising HS-MS data fusion approach based on unsupervised unmixing. Therefore, LMM based CNMF method yields the high-spatial-resolution HSI without any prior knowledge [5]. However, two main factors spectral variability and nonlinearity are still obstructing the performance of LMM method. The variation in spectral signature due to illumination, topography, atmospheric effects of the material are considered as spectral variability [6]. This variation of spectral signatures may cause significant errors during the spectral unmixing process. So, it is necessary to pay a considerable attention in spectral variability during the hyper-spectral unmixing process [7].

In this article, a multiplicative iterative nonlinear constrained coupled nonnegative matrix factorization (MINC-CNMF) algorithm is proposed. This MINC-CNMF algorithm aims to enhance the spatial as well as spectral quality of HSI along with considering the nonlinearity of the image. For this achievement, a multiplicative iterative algorithm is used alternately to update the endmember, abundance and outlier term that accounts for nonlinearity. This fusion algorithm enhances the spatial quality of hyperspectral image, but to

improve the spectral data we add minimum volume constraints to the simplex. This minimum volume (MV) of simplex controls the quality of spectral data. The simplex formed by connecting the endmembers selected for this process. The iterations continue until the algorithm reaches the convergence condition. Finally, the fusion image is created as a product of endmember and abundance with the outlier term. Thus, obtain high spatial and spectral quality fused image [8].

The MINC-CNMF algorithm is experimented on various synthetic data sets to evaluate the reconstruction quality in both spatial and spectral wise. This proposed method also compares with many other baseline algorithms available in literature to determine the performance of our fusion result. The main contribution the MINC-CNMF algorithm is that, 1) improve the accuracy for the extraction of spatial information from the MSI image, 2) improve the visual effect of the HSI image without any other distortion in the image, 3) propose an efficient image fusion algorithm that consider the nonlinearity effects in the image.

The succeeding sections of this paper are arranged as following order. In Section 2, a detailed study is done about the various spectral unmixing based fusion work that is related to our paper and identified the proper research gap in the literature. Then we proposed and formulated a model for MINC-CNMF algorithm in Section 3 followed by implementation of the same algorithm in Section 4. The details of experimental results and discussion are given in the Section 5, which also include the dataset and quality measures used for the performance evaluation. Finally, in the Section 6 conclusion and future enhancement of the work is explained.

II. RELATED WORK

Yokoya *et.al* in 2012 [16] proposed a coupled NMF (CNMF) algorithm which fuse HS and MS data based on unsupervised unmixing. Compared to other existing unmixing based fusion CNMF is straightforward and easy for mathematical formulation and its implementation. This method also minimizes the residual errors during unmixing process. Finally, this method results better local optimal solution and produce high-resolution hyperspectral image. Simoes *et.al* in 2015 [17] proposed a hyperspectral super resolution method, termed as HySure. This modal formulates data fusion as a convex optimization problem by adding an edge-preserving regularizer. This method uses vector total variation (VTV) regularizer to promote piecewise-smoothness to the image.

Lin *et.al* in 2018 [18], proposed an unmixing based fusion problem by incorporating two regularization terms such as sparsity and sum-of-squared-distances (SSD) regularizer. This method uses ℓ_1 -norm regularization to promote sparsity with well-known SSD regularizer to yield a fused image with high quality data. Therefore, this method upgrades the existing CNMF fusion performance by adding these regularization terms with two convex subproblems. Therefore, this algorithm with biconvex optimization is so called a convex optimization-based CNMF (CO-CNMF) algorithm. But as the noise level increases the CO-CNMF algorithm would degrade its performance rapidly. Therefore, it is necessary to add image denoising or smoothing constraints with the fusion method.

Yang *et.al* in 2019 [19], proposed an algorithm incorporating total variation and signature-based (TVSR) regularizations, into the CNMF method. Therefore, this algorithm is referred to as TVSR-CNMF. The total variation (TV) regularizer ensures the image smoothness to the abundance information. The signature-based regularizer provide high-fidelity signature reconstruction. Therefore, this method enhances both spatial and spectral quality of the fused data irrespective of high noise level in environment.

Borsoi *et.al* in 2019 [20], presented an unmixing based fusion algorithm that deals the spectral variability between images captured at different time instants. For this implementation, Generalized LMM (GLMM) that considers the variability of spectral signature by using individual scaling factor of each spectral band. This method divides the high-resolution images into subspace components and then represents the variability of the spectral signatures in each subspace separately. Then solve and combine each subproblem to obtain high fused data. Therefore, this algorithm is called HS-MS image Fusion with spectral Variability (FuVar). Due to the consideration of more complex spectral variability this method creates some difficulty to obtain an optimization solution.

Yang *et. al.*in 2019 [21], proposed an unmixing based fusion method by imposing sparsity and proximal minimum-volume regularizer. The minimum-volume regularizer control and minimize the distance between center of mass and the endmember at each iteration. Thus, it redefines the fusion method at each iteration until reaches the simplex volume to minimum. The algorithm is called SPR-CNMF and it reduced the computational complexity. This method also controls the loss of cubic structural information and thus improves the fusion performance by yielding high-fidelity reconstructed images and also shows good performance at high noise-level.

HS-MS image fusion is becoming an ever-increasing demand for the resolution enhancement of HS imagery. In this paper we analysed the literature based on several spectral unmixing based fusion methods. From the comprehensive and recent overviews of fusion models and associated spectral unmixing algorithms it is identified that unmixing-based fusion methods provide good and stable performance and estimated high-fidelity in the reconstructed image. The CNMF also shows a good classification capability in HS-MS fusion. Due to this reasons, CNMF is one of the very promising fusion methods based on spectral unmixing. The various algorithm based on CNMF models was introduced by imposing some constraints to the standard CNMF, to get better quality fused image. Even though all these methods had established well desired outcome, but most of the CNMF method are based on linear mixing model that means it does not consider the non-linearities of the image.

The LMM based method does not consider the nonlinear data in the image such as specific or localized areas of the image, the areas at the edges or boundary of heterogeneous regions. This property implies that, LMM assume only limited number of pixels. Due to this limited number of pixels assumptions, there occurs a lack of obtaining information during the unmixing process. Therefore algorithms based on

LMM are found difficulty for the accurate estimation of endmembers and their abundances fractions. So, in our work, we are planning to modify the CNMF algorithm by adding some regularizer for improving the image quality and also an additive term for considering the non-linearity of the image. Thus, improvement the performance and robustness of unmixing based image fusion algorithm.

III. PROBLEM FORMULATION

Let $Y_h \in \mathbb{R}^{L_h \times N_h}$ be an observed LR-HSI with L_h bands and N_h pixels, and $Y_m \in \mathbb{R}^{L_m \times N_m}$ be an observed HR-MSI with L_m bands and N_m pixels, with $L_m < L_h$ and $N_h < N_m$. Then data fusion of L_h band from LR-HSI, Y_h and N_m pixels from HR-MSI, Y_m to yield the desired high spectral and spatial resolution hyper-spectral image, $Z \in \mathbb{R}^{L_h \times N_m}$ [9].

$$Z = EA \quad (1)$$

The observed Y_h and Y_m can be represented as spectrally and spatially degraded version of fused image Z . This is represented as,

$$Y_m \approx DZ + R_m \quad (2)$$

$$Y_h \approx ZB + R_h \quad (3)$$

Where, $B \in \mathbb{R}^{N_m \times N_h}$ is a point spread function (PSF) which is used to blur the spatial quality of referenced hyper-spectral image to obtain LR-HSI, Y_h . $D \in \mathbb{R}^{L_m \times L_h}$ is a spectral response function (SRF) which is used to spectral downsampling of referenced hyper-spectral image to obtain HR-MSI, Y_m . The matrix R_m and R_h denote as residual noise are generally assumed as zero-mean Gaussian noises, but here this residual term R_m and R_h are consider as an outlier nonnegative matrix to accounts the nonlinearity effects [10].

The CNMF algorithm is a coupling of two NMF algorithm, which factorize a matrix into a product of two nonnegative matrices called endmembers and abundances. The CNMF algorithm starts from NMF by unmixing both hyper-spectral image Y_h and multispectral image Y_m [11]. That means CNMF simultaneously unmix both Y_h and Y_m by using NMF to estimate E and A , with the constraints. Then fuse the required data by using CNMF to obtain high quality image [12].

By the Eq. (1), (2) and (3) the cost functions of NMF unmixing for Y_h and Y_m , are defined as,

$$\|Y_h - EA_h\|_F^2 \quad \text{and} \quad \|Y_m - E_m A\|_F^2 \quad (4)$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm. The squared Frobenius norm are commonly used to minimize the cost function in hyper-spectral unmixing [10].

Then the CNMF method fuse the data from Y_h and Y_m to reconstruct the high spatial-spectral resolution hyperspectral image $Z = EA$ [13]. Therefore, the objective function CNMF can be defined as,

$$\text{CNMF}(E, A) = \|Y_h - EA_h\|_F^2 + \|Y_m - E_m A\|_F^2$$

subjected to, $E, A \geq 0$ (5)

The CNMF method has an ill-posed problem by nature, which means it may have more than one solution to the objective function. This ill-possedness problem of CNMF can be solved by adding some constraints or regularization term into spectral signature and/or fractional abundance.

A. Constrained CNMF Method

In this model we extend the standard LMM by considering the residual term R which accounts all possible nonlinear effects in the image. This additional term R gives the measurement of errors or novelty that deviates from the overall distribution of original sample data [14]. Therefore, the NMF unmixing for Y_h and Y_m , are defined as,

$$\|Y_h - (EA_h + R_h)\|_F^2 \quad \text{and} \quad \|Y_m - (E_m A + R_m)\|_F^2 \quad (6)$$

Then the objective function CNMF for Eq. (6) can be redefined as,

$$\text{CNMF}(E, A, R) = \|Y_h - (EA_h + R_h)\|_F^2 + \|Y_m - (E_m A + R_m)\|_F^2 \quad (7)$$

The symbol approximation (\approx) in Eqs. (2) and (1) indicates that it aims to obtain the minimum dissimilarity between both referenced and estimated image. The measure of dissimilarity between referenced image Y and estimated image $EA+R$ can be represented as, $D(Y|EA + R)$ [14]. Then this measure of dissimilarity on both hyperspectral Y_h and multispectral image Y_m in CNMF (E, A, R) method can be represented as:

$$\min_{E,A,R} \text{CNMF}(E, A, R) = D(Y_h|EA_h + R_h) + D(Y_m|E_m A + R_m) \quad \text{subjected to, } E, A, R \geq 0 \quad (8)$$

This additional term R itself is not enough to amounts the ill-posed problem of CNMF, so we further incorporate some priors to E and A based on its physical considerations. To reduce volume of the simplex a signature-based minimum volume (MV) constraint is imposed into our problem. To reconstructs the fused image, $Z = EA + R$, the objective function of our constrained CNMF method will be as following:

$$\min_{E,A,R} \text{CNMF}(E, A, R) + \alpha \phi_{MV}(E)$$

subjected to $E \geq 0, A \geq 0, R \geq 0$ (9)

where CNMF (E, A, R) is unconstrained CNMF method, $\alpha > 0$ are the parameters to control the MV constraints thus strengthen the spectral quality and this MV constraints are calculated as,

$$\phi_{MV}(E) = \sum_{i=1}^p \sum_{j=i+1}^p \|e_i - e_j\|_2^2 \quad (10)$$

$\phi_{MV}(E)$, is a well-known regularizer which minimize the volume of the simplex in hyper-spectral imagery, this minimum volume simplex is capable of estimating high-fidelity spectral signature [15].

IV. PROBLEM OPTIMIZATIONS

This proposed MINC-CNMF algorithm alternatively solved through the following convex subproblems until convergence:

$$R^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E^k, A^k, R) \quad (11)$$

$$A^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E^k, A, R^k) \quad (12)$$

$$E^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E, A^k, R^k) + \alpha \phi_{MV}(E) \quad (13)$$

where k indicates the number of iterations. This algorithm initializes E^0 with ATGP, A^0 with FCLS and outlier R^0 as a random matrix. Then updated each term (E, A, R) by using multiplicative updated rule [47]. Therefore, this algorithm starts with a given initial value (R^0, A^0, E^0) , then proceed in the order as $(R^k, A^k, E^k) \rightarrow (R^{k+1}, A^k, E^k) \rightarrow (R^{k+1}, A^{k+1}, E^k) \rightarrow (R^{k+1}, A^{k+1}, E^{k+1})$ where k is the current iteration stage. These steps repeated until it meets the convergence condition.

The method is implemented by extracting endmember E from LR-HIS, abundance matrix A from HR-MSI and the outlier R_h and R_m is estimated from observed HSI and MSI. After this, the outlier term R for final high resolution HR-MSI uses a low resolution Outlier Active Function (OAF) $S \in \mathbb{R}^{L_h \times N_m}$ that estimate the outlier matrix $R = R_h S R_m$. Consequently, the high spectral-spatial resolution fused image Z can be produced as follows,

$$Z = EA + R \quad (14)$$

This fused image Z cover all the information of ground truth image without any distortion in spectral and spatial information. The Algorithm 1 shown below summarized the proposed unmixing based fusion MINC-CNMF algorithm.

V. EXPERIMENTS AND PERFORMANCE ANALYSIS

To evaluate the performance of the unmixing based fusion method using our MINC-CNMF algorithm we conducted our experiment on four public dataset and then measured the quality of fusion method by using various quality measures. At last, the strength and superiority of our algorithm were evaluated by conducting experiments on four hyper-spectral datasets. We also compare the quality of our fusion methods with baseline fusion methods includes CNMF [16], HySure [17], CO-CNMF [18], TVSR-CNMF [19] and FuVar [20].

Algorithm 1: MINC-CNMF algorithm

Input:

LR-HSI $\rightarrow Y_h$

HR-MSI $\rightarrow Y_m$

Initialize:

$k=0$ and (R^0, A^0, E^0)

Step 1 : First unmix Y_h using nonlinear NMF unmixing algorithm by $\|Y_h - (EA_h + R_h)\|_F^2$

Optimize E, A_h and R_h as follows

$$R_h^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E^k, A_h^k, R_h) \quad A_h^{k+1} =$$

$$\arg \min_{E,A,R} \text{CNMF}(E^k, A_h, R_h^k)$$

$$E^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E, A_h^k, R_h^k) + \alpha \phi_{MV}(E)$$

Step 2 : Subsequently, unmix Y_m using same unmixing algorithm by $\|Y_m - (E_m A + R_m)\|_F^2$

Optimize E_m , A and R_m as follows

$$R_m^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E_m^k, A^k, R_m)$$

$$A^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E_m^k, A, R_m^k)$$

$$E_m^{k+1} = \arg \min_{E,A,R} \text{CNMF}(E_m, A^k, R_m^k) + \alpha \phi_{MV}(E_m)$$

Step 3: Repeat step 1-2 until convergence or predefined terminating condition is satisfied.

Step 4: Optimize the outlier R by using OAF $S : R = R_h S R_m$

Step 5: Reconstruct Z from E, A and R: $Z = EA + R$.

Output:

Fused HS image Z with high spatial-spectral dimension.

A. Dataset

Experimented and evaluated the proposed MINC-CNMF algorithm by using four real dataset such as Washington DC mall, Neon, Pavia University, Indian Pines. The first dataset is Washington DC Mall dataset is a well-known dataset captured by HYDICE sensor. This dataset contains image of size 1278×307 pixels. Due to large size of the image, we crop it to a 240×240-pixel size and that is selected for experiment which possesses 191 bands with 0.4 to 2.5 μm spectral range [16].

The second dataset is NEON Data, this dataset provides information on the National Observatory Networks San Joaquin Experimental Range field site. The image was collected over the San Joaquin field site located in California. The image selected for this experiment consists of 500×500 pixels with 107 bands with 0.4 to 0.85 μm spectral range [23].

The third HS dataset is Pavia University captured by the reflective optics spectrographic imaging system (ROSIS-3) over the University of Pavia, northern Italy in 2003. It consists of 610 × 340 pixels with 103 bands with 0.430 to 0.838 μm spectral range. The image select for this experiment is 560×320 pixel size [17].

The fourth HS image dataset AVIRIS Indian Pines is captured by AVIRIS sensor over the Indian Pines test site in northwestern Indiana, USA, in 1992. The image consists of 512×614 pixels. The selected image for experiment consists of 350×360 pixels size and 192 bands with wavelength range from 0.4 to 2.5 μm [16].

The high resolution ground truth images from these datasets were used as referenced image. The observed LR-HSI and HR-MSI are the input data for image fusion. These observed input data were generated by degrading spatial and spectral data from the referenced image according to Wald's protocol [22]. The observed LR-HSI Y_h is created by down sampling the ground truth image Z with a spatial blur factor $\omega = 6$ in both horizontal and vertical directions, respectively. The observed MSI Y_m , was produced with uniform spectral response functions corresponding to Landsat TM bands 1–5 and 7, which cover the 450–520, 520–600, 630–690, 760–900, 1550–1750, and 2080–2350 nm regions, respectively [10].

B. Quality Metrics

The strength of our fused images are measured by using four quality metrics: Spectral Angle Mapper (SAM), Signal-to-Reconstruction Error (SRE), Root-Mean-Square Error (RMSE), Peak Signal to Noise Ratio (PSNR), Universal Image Quality Index (UIQI). Using these metrics, the performance of our hyper-spectral image fusion algorithm is compared and also evaluated the quality of the estimated image by comparing it with ground truth HSI [8].

1) *SAM*: SAM identifies the spectral distortion between the estimated spectra E and a ground truth spectrum \hat{E} with n number of pixels. It measures the spectral similarity between the estimated and reference spectra by calculating the angle difference of the vectors between them as follows:

$$SAM(E, \hat{E}) = \frac{1}{n} \sum_{j=1}^n \arccos \left[\frac{E_j^T \cdot \hat{E}_j}{\|E_j\|_2 \cdot \|\hat{E}_j\|_2} \right] \quad (15)$$

The arccosine is defined as the inverse cosine function of the given value. If higher the spectral similarity between estimated spectra E and a ground truth spectrum \hat{E} , SAM values is closer to zero. That means, SAM value near to zero indicates high spectral quality [17].

2) *SRE*: The SRE measure the quality of reconstructed image based on the accuracy of estimated abundance data. Therefore, using this result, we can determine the quality and robustness of the proposed algorithm. The SRE is measured as follows,

$$SRE = 10 \log_{10} \left(\frac{\frac{1}{n} \sum_{i=1}^n \|\hat{A}_i\|_2^2}{\frac{1}{n} \sum_{i=1}^n \|\hat{A}_i - A_i\|_2^2} \right) \quad (16)$$

where the number of pixels is denoted as n , and \hat{A}_i and A_i are the abundance vectors of the estimated and original at the i^{th} pixel. Larger the SRE value, higher the spatial quality of the image [24].

3) *RMSE*: The result of RMSE gives the average difference between the original and estimated abundance map and so this result indicates the quality of the image as well as unmixing algorithm. That means, this matrix measures the spatial quality between the reference abundance \hat{A} and estimates abundance image A , is defined as:

$$RMSE(\hat{A}, A) = \frac{1}{\lambda_h n_m} \|A - \hat{A}\|_F^2 \quad (17)$$

Where, λ_h and n_m are the number of bands and the pixels in each of the band. The ideal value of RMSE is equal to zero and it can be achieved when $\hat{A} = A$ which signifies that there is no deviation. Smaller the RMSE value, better the quality of image [20].

4) *PSNR*: PSNR measure the reconstruction quality of spatial data in band wise. PSNR is the ratio between the signals to the residual errors. The PSNR of the l^{th} band is defined as

$$PSNR = \frac{1}{\lambda_h} \sum_{l=1}^{\lambda_h} PSNR_l \quad (18)$$

where $PSNR_l$ measures the spatial quality in the l^{th} spectral band, is defined as:

$$PSNR_l = 10 \cdot \log_{10} \left(\frac{\max(A^l)^2}{\|\hat{A}^l - A^l\|/P} \right) \quad (19)$$

where A^l is the pixel value of the l^{th} abundance band in the image. Higher the PSNR value, better the spatial quality of the estimated image [16].

5) *UIQI*: UIQI determines the similarity between the original and the estimated images by calculating the average correlation between the both images. The A_i denotes the image at the i^{th} band and \hat{A}_i denotes the original image at the corresponding band, then the correlation between $A^{(l)}$ and $\hat{A}^{(l)}$ are calculated as,

$$Q(A^l, \hat{A}^l) = \frac{\sigma_{A^l \hat{A}^l}}{\sigma_{A^l} \sigma_{\hat{A}^l}} = \frac{2\mu_{A^l} \mu_{\hat{A}^l}}{\mu_{A^l}^2 + \mu_{\hat{A}^l}^2} \frac{2\sigma_{A^l} \sigma_{\hat{A}^l}}{\sigma_{A^l}^2 + \sigma_{\hat{A}^l}^2} \quad (20)$$

Where, μ_{A^l} and $\mu_{\hat{A}^l}$ denote the mean vectors, σ_{A^l} and $\sigma_{\hat{A}^l}$ denote the variances and $\sigma_{A^l \hat{A}^l}$ is the covariance of both images respectively. The UIQI measure the average correlation Q over all the bands as follows,

$$UIQI(A^l, \hat{A}^l) = \frac{1}{\lambda_h} \sum_{l=1}^{\lambda_h} Q(A^l, \hat{A}^l) \quad (21)$$

The UIQI value range from $[-1, 1]$. When both images are similar, $A = \hat{A}$, then the value of $UIQI(A^l, \hat{A}^l) = 1$. For the final result, the overall UIQI of the estimated HSI can be computed by averaging the UIQI value of all bands [17].

C. Regularization Parameter

The regularization parameter controls the minimization of optimized problem, thus guarantees reconstruction of hyperspectral image with high spatial-spectral resolution. The original CNMF problem is reformulated into the regularized or constrained CNMF by incorporating minimum volume constraints to enhance the spectral quality during the enhancement of spatial quality by fusion. For regularizing this constraint a parameters α imposed into the problem. According to the comparative analysis of the existing literature study, the values of this parameter are set empirically in the range $\{1e-5, 1e-4, 1e-3, 1e-2, 0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4\}$ to obtain the best result [20].

The results obtained on our proposed method with this parameter α for set of above ranged values are observed in four selected datasets. It shows that when the α value exceeds 0.0015, the value of SAM and RMSE increases rapidly. Similarly, when the value of parameters α is below 0.0001, the RMSE value also starts rising. Therefore, the performance metrics provide better values in between 0.0015 and 0.0001. Other performance measures such as SRE, PSNR, and UIQI are also shows higher value in these ranges in all our four datasets. So, we set $\alpha = 0.00001$ (1e-4) to achieve better performance [20].

D. Performance Analysis of Fusion Algorithm

Many fusion algorithms are introduced in recent years to enhance the hyper-spectral image by fusion. But majority of these unmixing based fusion algorithms enhance the quality of either spectral or spatial information. Out of these existing algorithms our MINC-CNMF algorithm improves the quality of both spectral and spatial data in the hyper spectral image. In addition to this, enhancement is also considered the nonlinearity in the image during fusion by unmixing process. In this work our plan to enhance the spatial quality of LR-HSI. For this purpose, we extracted the high spatial quality of multispectral image and fused with high LR-HSI which having high spectral quality. The LR-HSI and HR-MSI are obtained by spatial and spectral down sampling of ground truth HSI.

In this paper, we proposed a MINC-CNMF algorithm for fusing the high quality spectral and spatial data of LR-HSI and HR-MSI by enhancing the both data without any distortion. The visual effect of our proposed algorithm on four datasets is presented in Fig. 1. Form Fig. 1 it is found that our proposed

algorithm gives better visual effect compared to other fusion method.

Then the performance capability of this proposed algorithm against various baseline algorithm are tested using the quality measure such as SAM, RMSE, SRE, PSNR and UIQI and the result are shown in the Table I. The result shown in Table I indicates that our proposed algorithm show superiority over all the baseline fusion methods. The less SAM value indicates that estimated fused image has less spectral distortion. The reduced value of RMSE shows the high reconstruction fidelity to the referenced image. The high PSNR value shows the good spatial quality of estimated image. Similarly, the higher value of SRE and UIQI shows much better performance of fusion algorithm. The performance of all these quality measures are shown in Fig. 2.

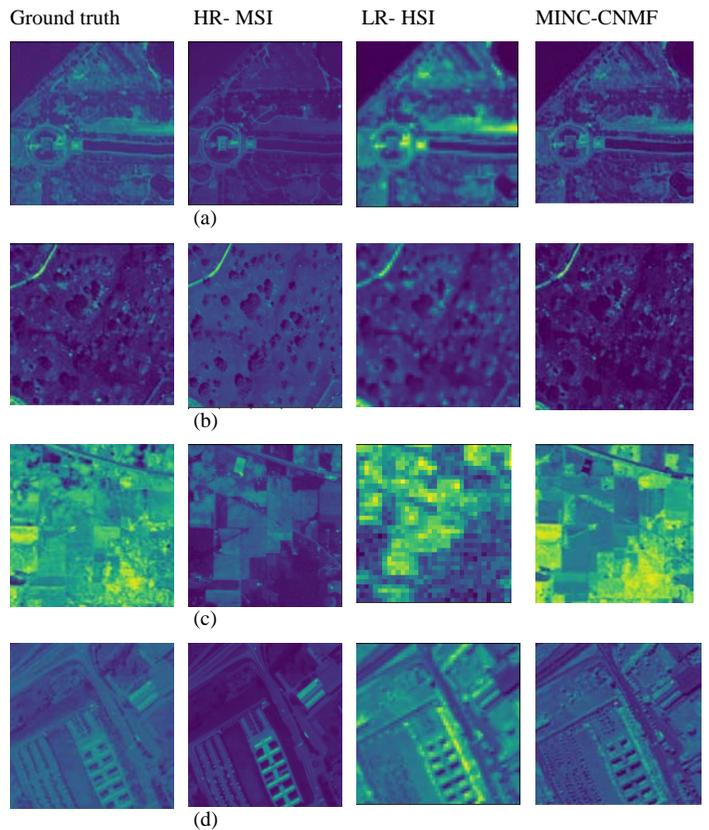


Fig. 1. The Ground Truth, HR-MSI, LR-HSI and our MINC-CNMF Algorithm Images of Four Dataset (a) Washington DC Mall, (b) NEON, (c) Indian Pines and (d) Pavia University.

TABLE I. PERFORMANCE VALUES OF DIFFERENT QUALITY MEASURES ON THE FOUR DATASETS (BEST VALUES ARE MARKED AS BOLD CHARACTER)

Dataset	Method	CNMF	HySure	CO-CNMF	TVSR-CNMF	FuVar	Proposed
Washington DC Mall	SAM	1.01	0.99	0.74	0.74	0.76	0.72
	RMSE	8.57	7.50	7.41	7.43	0.01	0.01
	SRE	15.26	15.46	15.64	15.66	16.17	16.58
	PSNR	76.23	76.02	75.35	76.01	75.33	76.31
	UIQI	0.01	0.09	0.02	0.01	0.11	0.30
NEON	SAM	0.59	0.69	0.52	0.86	0.51	0.49
	RMSE	5.56	5.58	5.57	5.58	5.58	4.45
	SRE	20.31	20.01	20.46	20.46	20.47	20.47
	PSNR	83.11	83.09	84.23	89.32	90.01	110.14
	UIQI	0.02	0.08	0.01	0.11	0.1	0.15
Pavia University	SAM	0.57	0.39	0.33	0.61	0.34	0.31
	RMSE	1.51	1.23	2.11	3.12	3.08	0.01
	SRE	16.13	17.03	17.33	16.03	17.01	17.83
	PSNR	70.67	71.90	71.43	72.30	71.40	72.35
	UIQI	0.03	0.01	0.13	0.09	0.25	0.30
Indian Pines.	SAM	0.53	0.38	0.02	0.48	0.01	0.01
	RMSE	2.01	2.24	2.72	2.77	2.91	0.01
	SRE	13.98	12.90	12.07	14.70	13.48	14.70
	PSNR	72.61	74.11	73.71	72.04	73.85	74.80
	UIQI	0.15	0.08	0.14	0.21	0.12	0.23

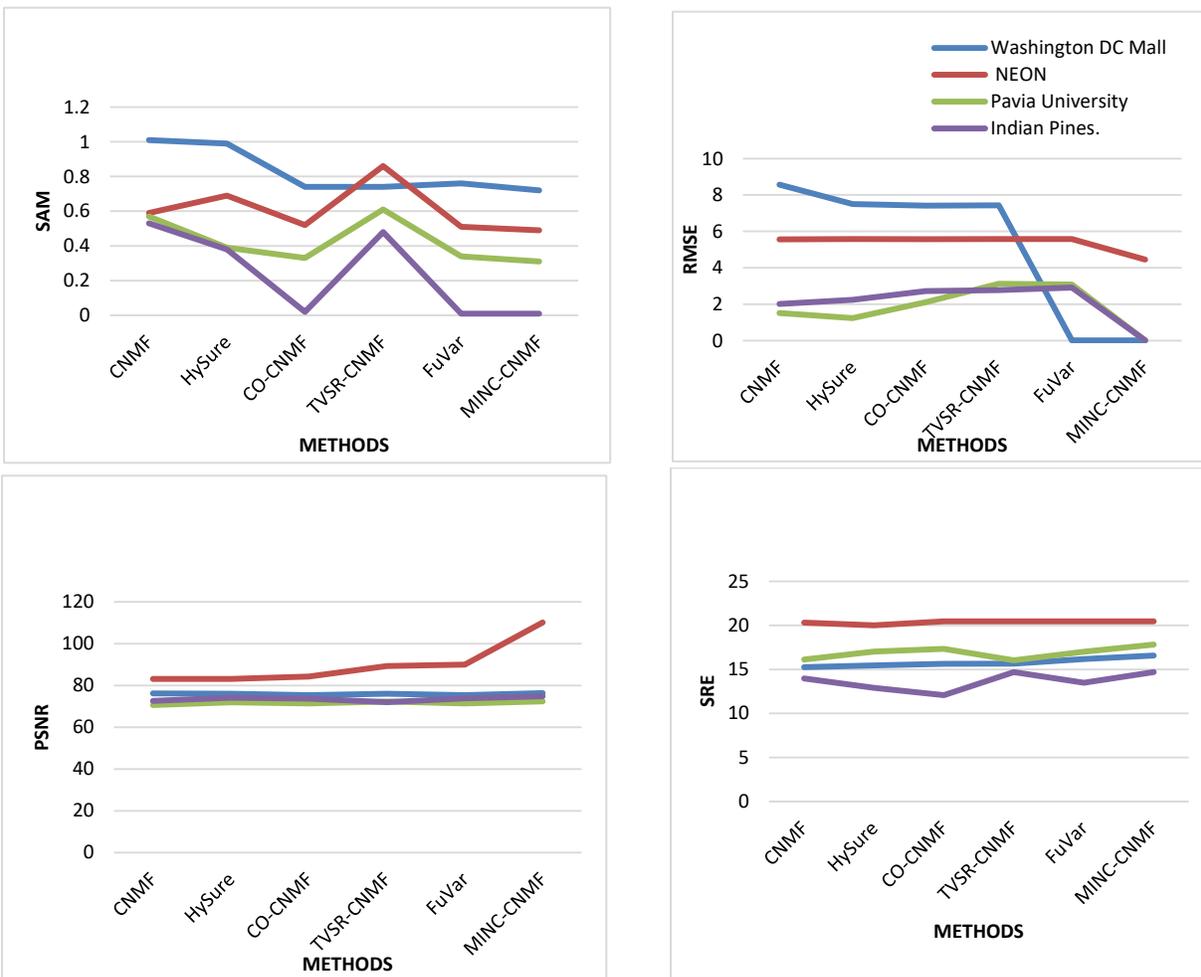


Fig. 2. The Performance of Four Quality Measures on our Proposed MINC-CNMF Algorithm.

VI. CONCLUSIONS

In this paper, we proposed an unmixing based fusion algorithm to enhance the LR-HSI using a multiplicative iterative nonlinear constrained coupled nonnegative matrix factorization (MINC-CNMF) algorithm. This algorithm enhances the quality of both spectral as well as spatial dimension of the image to the standard CNMF. In addition, this algorithm also considered the nonlinear factors in the image by considering the outlier data heterogeneous area, pixel illuminations, tiny spots in the image. This MINC-CNMF spectral unmixing based fusion algorithm updates each parameter by using a popular update method namely multiplicative update rule, which update the endmember signatures, abundances and the outlier matrix iteratively until it reaches the stopping criteria as explained in the algorithm implementation.

We experimented our fusion by unmixing method on four real-world dataset and analyze the performance of our methods on various quality measures such as SAM, SRE, RMSE, PSNR, UIQI. Then we compare the effectiveness of our MINC-CNMF algorithm with the existing methods. From the results produced by all the above method it is found that the proposed algorithm gives better fused image with high quality spatial and spectral dimension and consume less computational processing time than all other existing methods. This work mainly focused on exploring both spatial and spectral information by considering nonlinear effects in the hyper-spectral image. In future work, to furthermore improve the accuracy of unmixing performance by introducing some more constraints or prior information about endmembers and abundance to this modal.

ACKNOWLEDGMENT

The authors would like to express the gratitude to all the reviewers for their helpful comments and suggestions to improve the quality of paper.

REFERENCE

- [1] Ting Xu, Ting-Zhu Huang, Liang-Jian Deng, XiLe Zhao, Jie Huang. "Hyperspectral Image Superresolution Using Unidirectional Total Variation With Tucker Decomposition", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020.
- [2] Renwei Dian, Shutao Li, Leyuan Fang, Qi Wei. "Multispectral and hyperspectral image fusion with spatial-spectral sparse representation", Information Fusion, 2019.
- [3] Naoto Yokoya, Claas Grohnfeldt, Jocelyn Chanussot. "Hyperspectral and Multispectral Data Fusion: A comparative review of the recent literature", IEEE Geoscience and Remote Sensing Magazine, 2017.
- [4] Renwei Dian, Shutao Li, Leyuan Fang, Ting Lu, Jose M. Bioucas-Dias. "Nonlocal Sparse Tensor Factorization for Semiblind Hyperspectral and Multispectral Image Fusion", IEEE Transactions on Cybernetics, 2019.
- [5] Xuelong Li, Yue Yuan, Qi Wang. "Hyperspectral and Multispectral Image Fusion via Nonlocal Low-Rank Tensor Approximation and Sparse 11 1% 12 1% 13 1% 14 1% 15 1% Representation", IEEE Transactions on Geoscience and Remote Sensing, 2020.
- [6] Danfeng Hong, Naoto Yokoya, Jocelyn Chanussot, Xiao Xiang Zhu. "An Augmented Linear Mixing Model to Address Spectral Variability for Hyperspectral Unmixing", IEEE Transactions on Image Processing, 2019.
- [7] Danfeng Hong, Naoto Yokoya, Jocelyn Chanussot, Xiao Xiang Zhu. "Learning a lowcoherence dictionary to address spectral variability for hyperspectral unmixing", 2017 IEEE International Conference on Image Processing (ICIP), 2017.
- [8] Xinyu Zhou, Ye Zhang, Junping Zhang, Shaoqi Shi. "Alternating Direction Iterative Nonnegative Matrix Factorization Unmixing for Multispectral and Hyperspectral Data Fusion", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020.
- [9] Fei Ma, Feixia Yang, Ziliang Ping, Wenqin Wang. "Joint Spatial-Spectral Smoothing in a Minimum-Volume Simplex for Hyperspectral Image Super-Resolution", Applied Sciences, 2019.
- [10] D. Hong, N. Yokoya, J. Chanussot, X. Zhu, An Augmented Linear Mixing Model to Address Spectral Variability for Hyperspectral Unmixing, Geography, Computer Science, IEEE Transactions on Image Processing, 29 October 2018.
- [11] Naoto Yokoya, Takehisa Yairi, Akira Iwasaki. "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion", IEEE Transactions on Geoscience and Remote Sensing, 2012.
- [12] Li Sun ;Kang Zhao ,Congying Han,Ziwen Liu, " Enhancing Hyperspectral Unmixing With Two-Stage Multiplicative Update Nonnegative Matrix Factorization", Published in: IEEE Access (Volume: 7) Page(s): 171023 – 171031 Date of Publication: 26 November 2019 Electronic ISSN: 2169-3536 INSPECAccession Number: 19174338DOI: 10.1109/ACCESS.2019.2955982, 2019.
- [13] S A Gayathri, R J Renjith. "Spatial resolution enhancement of hyperspectral image by negative abundance oriented spectral unmixing", 2016 International Conference on Communication Systems and Networks, 2016.
- [14] Cédric Févotte and Nicolas Dobigeon, "Nonlinear Hyperspectral Unmixing With Robust Nonnegative Matrix Factorization", IEEE Transactions On Image Processing, Vol. 24, No. 12, December 2015.
- [15] Kewen Qu, Wenxing Bao. "Multiple-Priors Ensemble Constrained NonnegativeMatrix Factorization for Spectral Unmixing", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020.
- [16] Yokoya, N., Yairi, T., & Iwasaki, A. "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion". IEEE Transactions on Geoscience and Remote Sensing, Vol.50, No.2, page.no.528–537. doi:10.1109/tgrs.2011.2161320, 2012.
- [17] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspacebased regularization," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 6, pp. 3373–3380, Jun. 2015.
- [18] Lin, C.-H., Ma, F., Chi, C.-Y., & Hsieh, C.-H. "A Convex Optimization-Based Coupled Nonnegative Matrix Factorization Algorithm for Hyperspectral and Multispectral Data Fusion". IEEE Transactions on Geoscience and Remote Sensing, Vol.56, No.3, page.no.1652–1667. doi:10.1109/tgrs.2017.2746078, 2018.
- [19] Feixia Yang, Fei Ma, Ziliang Ping, and Guixian Xu, "Total Variation and Signature-Based Regularizations on Coupled Nonnegative Matrix Factorization for Data Fusion", Digital Object Identifier 10.1109/ACCESS.2018.2857943 IEEE Access, VOLUME 7, 2019.
- [20] Ricardo Augusto Borsoi, Tales Imbiriba, José Carlos Moreira Bermudez, Super-Resolution for Hyperspectral and Multispectral Image Fusion Accounting for Seasonal Spectral Variability, IEEE Transactions on Image Processing, DOI 10.1109/TIP.2019.2928895, 2019.
- [21] Feixia Yang, Ziliang Ping,Fei Ma, And Yanwei Wang, Fusion of Hyperspectral and Multispectral Images With Sparse and Proximal Regularization, Digital Object Identifier 10.1109/ACCESS.2019.2961240, IEEE Access, VOLUME 7, 2019.
- [22] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," Photogramm. Eng. Remote Sens., vol. 63, no. 6, pp. 691–699, 1997.
- [23] <https://www.neonscience.org/resources/learning-hub/tutorials/classification-endmember-python>.
- [24] Danfeng Hong, Naoto Yokoya, Jocelyn Chanussot, Xiao Xiang Zhu. "Learning a lowcoherence dictionary to address spectral variability for hyperspectral unmixing", 2017 IEEE International Conference on Image Processing (ICIP), 2017.

Optimal Operation of Smart Distribution Networks using Gravitational Search Algorithm

Surender Reddy Salkuti

Department of Railroad and Electrical Engineering, Woosong University
Daejeon, Republic of Korea.

Abstract—This paper proposes a methodology for an optimal operation of smart distribution network considering the network reconfiguration, distributed generation (DG) units allocation and optimally placing the shunt capacitors for reactive power compensation. In this work, the total power losses minimization objective is considered. By optimizing this objective, it can also result in the reduction of voltage deviation. The proposed problem is solved using evolutionary-based gravitational search algorithm (GSA). Simulation studies are performed on 33 bus radial distribution system (RDS). Simulation results reveal that there is a drastic reduction in the power losses by utilizing the network reconfiguration, DG allocation, and reactive power compensation.

Keywords—Distributed generation; renewable energy; meta-heuristic algorithms; network reconfiguration; smart grid; reactive power compensation

I. INTRODUCTION

Distribution system is the most important component of current smart power system, and it aims to provide the electricity to its customers in an efficient, reliable, economic, and environmentally friendly way while satisfying all the operating constraints of the system. It can be observed that during the last few years, there is a rapid expansion of power systems due to the drastic increase in the load demand which leads to high power losses and poor voltage regulation. To overcome this increased load demand and economic benefits, renewable energy has been introduced to the power system. In recent years, distributed generations (DGs) are emerging as an important alternative solution for the enhancement of smart power distribution systems [1]. Along with this, reactive power compensation and optimal network reconfiguration (ONR) are considered as the tools for the enhancement of smart power distribution systems.

Various objectives considered for the optimal reconfiguration of smart systems include investment and operational cost minimization, active power loss minimization, enhancement of reliability, etc. Several researchers proposed various approaches to handle the distribution networks include optimal network reconfiguration (ONR), optimal allocation of DG, and simultaneous optimization of both ONR and DG allocation. ONR is an approach to change the configuration of distribution system by changing the opening/closing status of the sectionalizing (normally closed) and the tie (normally open) switches of the network, so that the radiality and connectivity are well maintained [2]. The distribution networks need to be expanded to meet the increasing demand, however, it is a big

issue as it is associated with various economic and environmental factors. In this situation, ONR is a viable solution to this problem [3, 4]. ONR problem is a highly non-linear, complex, mixed-integer, large-scale, combinatorial, non-differential, and constrained optimization problem.

The author in [5] proposes an approach for the economic operation of distribution networks by considering the curtailment costs. A multi-objective-based ONR is proposed in [6] for the minimization of loss and enhancement of voltage profile. A day-ahead ONR model is proposed in [7] for smart distribution networks including the renewables-based DGs and storage systems by considering voltage deviations and operating cost minimization objectives. The author in [8] proposes an approach for calculating the daily profit and risk of RDSs by considering the uncertainties of power outputs from the DG units and electricity prices. An analytical optimization approach is proposed in [9] for optimal investment of DG units along with capacitors to minimize interconnection costs of renewable sources and to enhance the voltage profile of the system. An approach for simultaneous ONR and optimal DG allocations in a RDS by considering the voltage stability improvement and active power loss minimization objectives has been proposed in [10]. Simultaneous and optimal allocation of RESs and reconfiguration in RDSs with reliability enhancement and cost of power losses minimization objectives are solved by using the information gap decision theory has been proposed in [11]. The ONR and supply restoration approach based on the improved genetic algorithm (GA) has been proposed in [12]. Multi-criteria based ONR of RDS considering reliability, stability improvement, and loss minimization objectives have been proposed in [13].

The above literature review revealed that there is a need for simultaneous optimization of ONR, reactive power compensation, and optimal allocation of DG units. To address the complexity and computational burden involved with the loss minimization objective, it uses the gravitational search algorithm (GSA). The work presented in this paper is as follows: Distribution load flow (DLF) analysis of the radial distribution system (RDS) has been presented in Section 2. The proposed optimal network reconfiguration (ONR) approach has been presented in Section 3. In Section 4, gravitational search algorithm (GSA) is described. In Section 5, 33 bus RDS is used to demonstrate the proposed approach and analyze the results for the considered test system. Finally, conclusions are drawn from the proposed study are reported in Section 6.

II. DISTRIBUTION LOAD FLOW (DLF)

Load flow analysis is a tool for steady-state analysis of the distribution network in both operational and planning stages. In the literature, several researchers use conventional approaches like Fast Decoupled, Newton Raphson load flow approaches to address static and dynamic distribution network problems [14]. However, these load flow approaches are inefficient due to high resistance to reactance (R/X) ratio of the distribution line and its radial structure, which has resulted in the development of special load flows for the RDSs [15]. Most of the load flow calculations are based on forward-backward sweep methods. The load flow approach presented in this paper is based on an iterative approach which is based on the receiving end voltage of the RDS. Here, an effective power at each bus is calculated after the formulation of the adjacent node and adjacent branch matrices. Fig. 1 depicts the equivalent circuit of a line in a RDS [16].

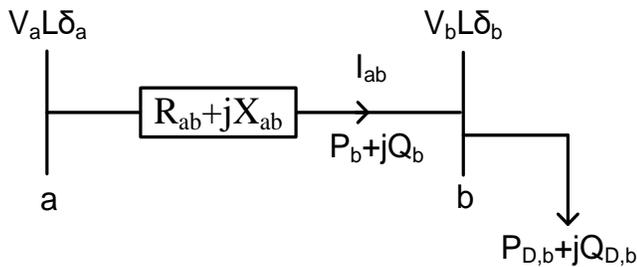


Fig. 1. Equivalent Circuit of a Branch in RDS.

The current in the branch (ab) connected between the sending end (a) and receiving end (b) can be expressed by [16],

$$I_{ab} = \frac{|V_a| \angle \delta_a - |V_b| \angle \delta_b}{R_{ab} + jX_{ab}} = \frac{P_b - jQ_b}{(|V_b| \angle \delta_b)^*} \quad (1)$$

Where,

$$P_b = \sum_{i=1}^{N_b} P_{D,i} + \sum_{l=1}^{B_b} P_{loss,l} \quad (2)$$

N_b represents all the nodes beyond node b and B_b represents all the branches beyond node b. $P_{D,i}$ is power demand at node i , and $P_{loss,l}$ is active power loss in the branch l . From equation (1), the active power (P_b) can be calculated by [17],

$$P_b = \frac{|V_a||V_b| \sin(\delta_a - \delta_b) + R_{ab}Q_b}{X_{ab}} \quad (3)$$

From equation (3), the voltage at the receiving end can be determined by,

$$|V_b| = - \left[|V_a| \left(\frac{R_{ab}}{X_{ab}} \sin \delta - \cos \delta \right) \right] + \left[|V_a| \left(\frac{R_{ab}}{X_{ab}} \sin \delta - \cos \delta \right) \right]^2 - 4Q_b \sqrt{\left(\frac{R_{ab}}{X_{ab}} + X_{ab} \right)} \quad (4)$$

Where $\delta = \delta_a - \delta_b$. The angle δ_b can be expressed by,

$$\delta_b = \delta_a - \tan^{-1} \left[\frac{P_b X_{ab} - Q_b R_{ab}}{|V_b|^2 + P_b R_{ab} + Q_b X_{ab}} \right] \quad (5)$$

The real and reactive power losses in a branch (ab) can be determined by [17],

$$P_{loss,ab} = \frac{R_{ab}(P_b^2 + Q_b^2)}{|V_b|^2} \quad (6)$$

$$Q_{loss,ab} = \frac{X_{ab}(P_b^2 + Q_b^2)}{|V_b|^2} \quad (7)$$

III. PROBLEM FORMULATION

Integration of renewable power generation along with ONR is performed to achieve optimum power losses in the system. As mentioned earlier, minimization of total power loss (P_T^{loss}) in the RDS is selected as a primary objective of ONR problem [18, 19]. However, the problem is also carried out with the objectives of voltage profile improvement, loadability and power quality enhancement, minimization of total network cost and emissions, economic and reliable operation. Real and reactive power losses in a branch connected between the nodes/buses a and b can be expressed by,

$$P_{a,b}^{loss} = \left(\frac{P_{a,b}^2 + Q_{a,b}^2}{|V_a|^2} \right) \times R_{a,b} \quad (8)$$

$$Q_{a,b}^{loss} = \left(\frac{P_{a,b}^2 + Q_{a,b}^2}{|V_a|^2} \right) \times X_{a,b} \quad (9)$$

Then total power losses (P_T^{loss}) in the entire RDS can be expressed as [20],

$$P_T^{loss} = \sum_{a=1}^{N_B} \left(\frac{P_{a,b}^2 + Q_{a,b}^2}{|V_a|^2} \right) \times R_{a,b} \quad (10)$$

Where N_B is number of buses in the RDS.

The power loss minimization objective can be expressed by [21],

$$f = \text{minimize } (P_T^{loss}) \quad (11)$$

In this work, optimization is performed, by considering the network reconfiguration, reactive power compensation by using the shunt capacitors, and by optimally allocating the DG units [22], then there is a minimization in the voltage deviation in the system. This voltage deviation can be expressed as,

$$\Delta V = \left(\frac{V_1 - V_a}{V_1} \right) \quad a = 1, 2, \dots, N_B \quad (12)$$

The above problem is solved by considering the following constraints.

A. Equality Constraints

These constraints are expressed as [23],

$$P_D = P_G^{Grid} + \sum_{i=1}^{N_{DG}} P_{DG,i} \quad (13)$$

$$Q_D = Q_G^{Grid} + \sum_{i=1}^{N_{DG}} Q_{DG,i} + \sum_{j=1}^{N_{RC}} Q_{RC,j} \quad (14)$$

B. Inequality Constraints

These constraints for the RDS are DG power limits, bus voltage, and branch current limits.

1) *DG power constraints*: The real and reactive power generations from DG units [24] are limited by,

$$P_{DG,i}^{min} \leq P_{DG,i} \leq P_{DG,i}^{max} \quad (15)$$

$$Q_{DG,i}^{min} \leq Q_{DG,i} \leq Q_{DG,i}^{max} \quad (16)$$

2) *Bus voltage constraint*: Lower and upper voltage limits of buses in the RDS are expressed as [25],

$$V_{Bus,i}^{min} \leq V_{Bus,i} \leq V_{Bus,i}^{max} \quad i = 1, 2, \dots, N_{Bus} \quad (17)$$

3) *Bus voltage constraint*: The current in each branch ($I_{b,k}$) is limited by,

$$I_{b,k} \leq I_{b,k}^{max} \quad k = 1, 2, \dots, N_{Br} \quad (18)$$

4) *Power flow constraint*: Power loss in each feeder (P_i) is limited by,

$$P_i \leq P_i^{max} \quad \text{for } i = 1, 2, \dots, N_F \quad (19)$$

Where N_F is number of feeders in the system.

5) *Capacitor constraint*: The capacity of shunt capacitor (C_k) is limited by,

$$C_k^{min} \leq C_k \leq C_k^{max} \quad \text{for } k = 1, 2, \dots, N_{Cap} \quad (20)$$

Where N_{Cap} is number of capacitor banks. C_k^{min} and C_k^{max} are minimum and maximum values of discrete controls of k^{th} capacitor banks.

The above problem is solved by retaining the structure of radial network. From an optimization perspective, these problems are considered as highly non-linear, highly constrained, mixed-integer, high dimension, and multi-modal optimization problems with a large number of local optimum solutions. So, determining the global solution is a complex optimization problem is challenging, which provides ample opportunity for further research. The distributed generators (DGs) can be wind energy generators, solar PV, biomass, small hydro, etc.

IV. GRAVITATIONAL SEARCH ALGORITHM (GSA)

GSA has been developed by E. Rashedi et al. in 2009 [26] by using Newton's law of gravity and motion. The gravitational force (F) can be expressed as,

$$F = G \frac{m_1 m_2}{r^2} \quad (21)$$

Where G is gravitational constant. m_1, m_2 are masses of the objects 1 and 2; r is distance between centers of masses. In this GSA, agents are represented as objects and their performance is measured by their masses. Let a system has N_m number of masses and the position of i^{th} mass (X_i) can be expressed as,

$$X_i = (x_i^1, x_i^2, \dots, x_i^d, \dots, x_i^n) \quad i = 1, 2, \dots, N_m \quad (22)$$

Where n is size of search space, x_i^d is position of i^{th} mass in d^{th} dimension. The flow chart of GSA is presented in Fig. 2. For detailed description of GSA, the reader may refer [26]-[29].

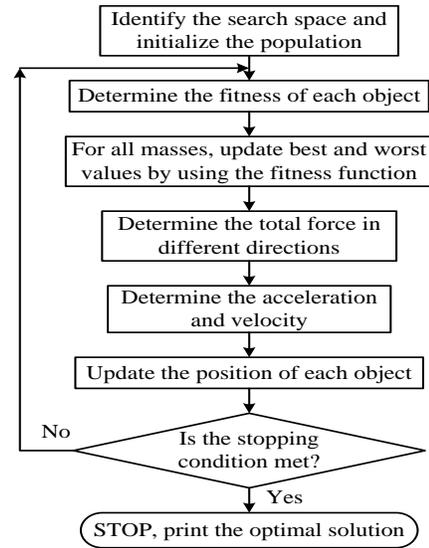


Fig. 2. Flow Chart of GSA.

V. RESULTS AND DISCUSSION

In this work 33 bus RDS is used for the analysis and explanation of network reconfiguration. The base voltage and base MVA for 33 bus is considered as 12.66 kV and 100 MVA [30]. In this test system, a maximum of three DG units are incorporated and the maximum size of DG is 2000 kW. In this paper, 3 case studies are considered, and they are:

- Case 1: Optimal operation with only reconfiguration.
- Case 2: Optimal operation with only reactive power compensation.
- Case 3: Optimal operation with only DG allocation.
- Case 4: Optimal operation with reconfiguration, reactive power compensation, and DG allocation.

This test system data has been taken from the [31, 32]. This test system has 33 buses, 32 lines, and bus 1 is assigned as the substation bus. In this test system, the active and reactive power demands are 3715 kW and 2300 kVAR. Single line diagram (SLD) of 33 bus system is depicted in Fig. 3. This system has 5 open tie switches which form the loops in the system and they are 33 to 37. This system has 32 sectionalizing switches marked as 1 to 32 which are normally closed and they are shown in Fig. 3.

A. Case 1

As mentioned earlier, a similar analysis can be made for a 33 bus RDS. Initial and final configurations of the system are depicted in Fig. 3 and 4. A comparison of node voltages before and after the reconfiguration is shown in Fig. 5. A comparison of various other parameters before and after reconfiguration is shown in Table I. Under the base case condition (i.e., before the network reconfiguration), the switches 33 to 37 are opened and the system power losses are 202.6592 kW. The minimum voltage has occurred at bus 18 and its value is 0.9038 p.u.

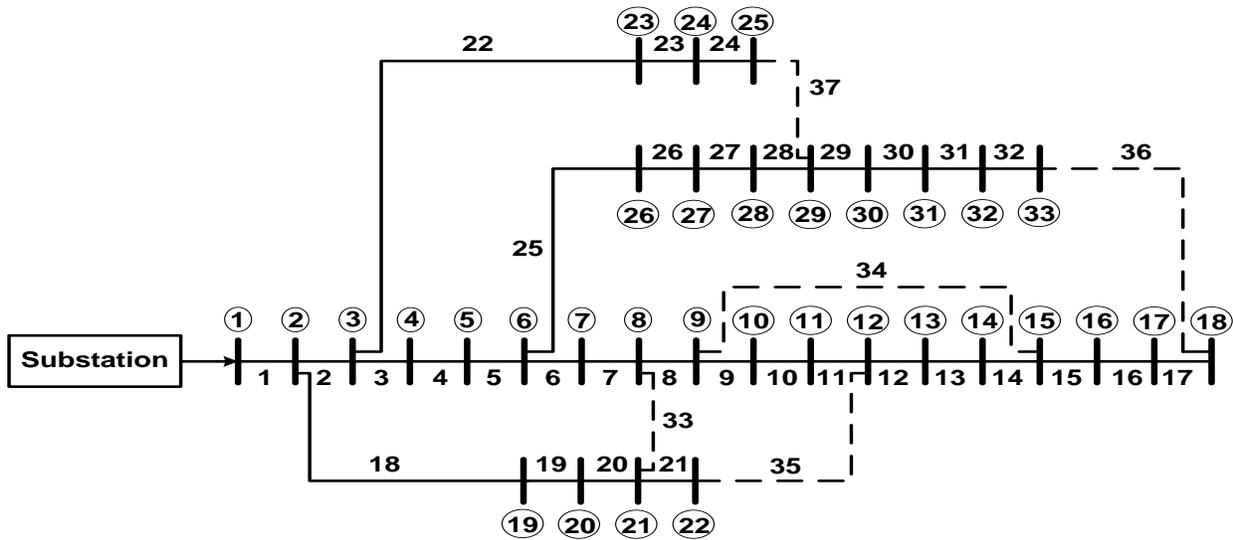


Fig. 3. Initial Configuration of a 33 Bus RDS.

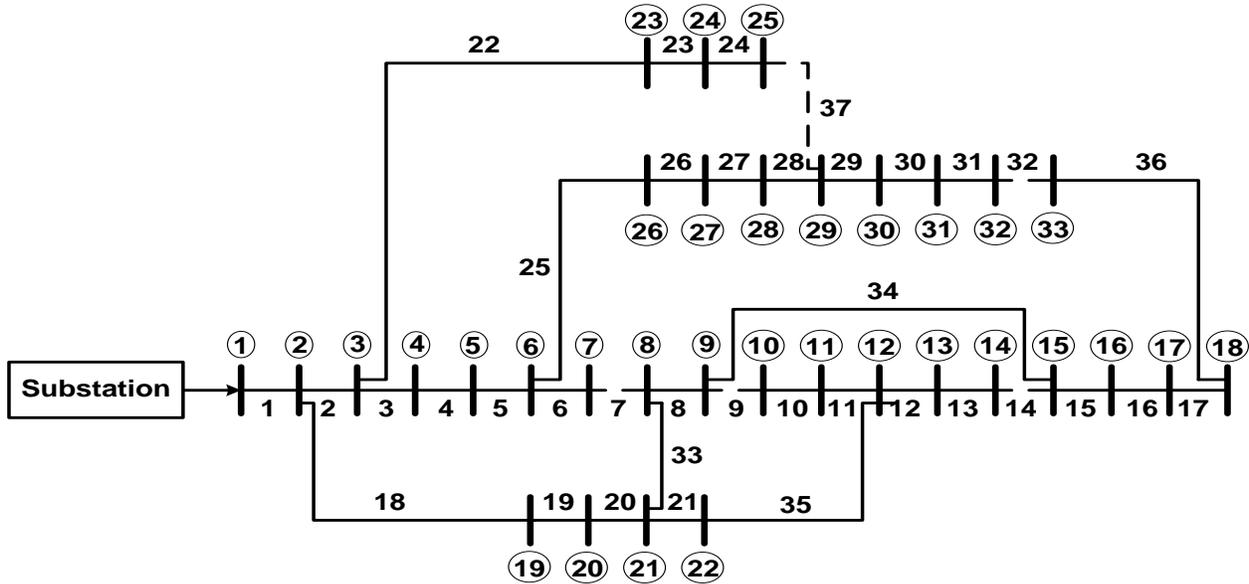


Fig. 4. Final Configuration of a 33 Bus RDS for Case 1.

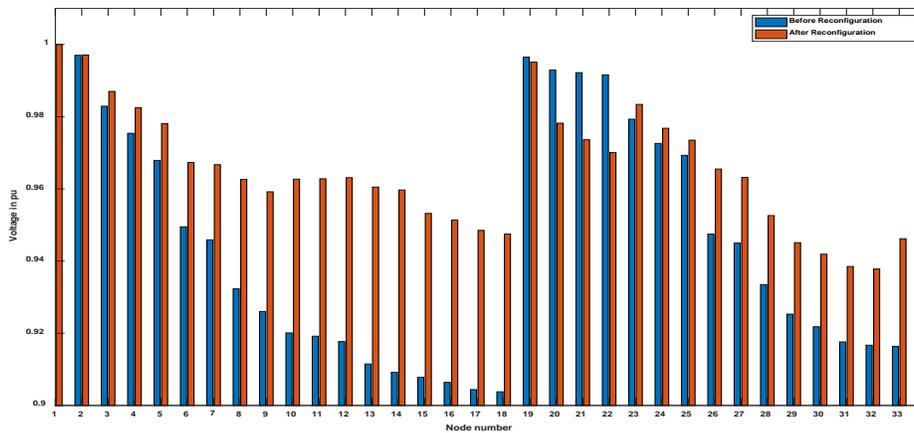


Fig. 5. Comparison of Node Voltages before and after Reconfiguration (Case 1) for a 33 Bus RDS.

TABLE I. SIMULATION RESULTS FOR 33 BUS RDS BEFORE AND AFTER THE NETWORK RECONFIGURATION

	Before network reconfiguration	After network reconfiguration
Tie switches opened for reconfiguration	33, 34, 35, 36, 37	7, 9, 14, 32, 37
Power loss (kW)	202.6592 kW	150.8846 kW
Reduction in power loss (%)	-	25.55 %
Minimum voltage (p.u.)	0.9038 pu at bus 18	0.9378 pu at bus 32

From Fig. 5, it can be seen that the voltages at node numbers 5 to 18 and 26 to 33 have drastically improved and at several other nodes, the voltage is improved to a reasonable extent. Though there is a decrease in voltages at some nodes, it is to a reasonable extent. From Table I it can be seen that there is a reduction in active power loss of 25.55%. The minimum voltage after reconfiguration is found to be 0.9378 p.u.

B. Case 2

In this case, reactive power compensation is used to reduce active power losses further. Table II presents the simulation results for 33 bus RDS for reactive power compensation using differential evolution (DE) and GSA.

TABLE II. RESULTS FOR 33 BUS RDS FOR REACTIVE POWER COMPENSATION USING DE AND GSA

	Reactive power compensation using DE	Reactive power compensation using GSA
Bus number and reactive power compensation (kVAr)	552.4 kVAr at bus 6	550.4 kVAr at bus 6
	560.5 kVAr at bs 28	562.8 kVAr at bs 28
	545.9 kVAr at bus 29	548.5 kVAr at bus 29
Total compensation	1658.8 kVAr	1661.7 kVAr
Power loss before the compensation	202.6592 kW	202.6592 kW
Power loss after the compensation	142.52 kW	140.83 kW
Power loss reduction	29.68%	30.51%
Minimum voltage	0.9794 pu at bus 33	0.9798 pu at buses 17 and 33

Active power losses before reactive power compensation is 202.6592 kW. In this paper, GSA is used for determining the optimum size of the capacitor at the potential candidate bus. In this work, it is considered that a maximum of 3 capacitors can be placed for the reactive power compensation. By using GSA, the total compensation required for this system is 1661.7 kVAr, and it is placed at buses 6, 28, and 29 with the compensation values of 550.4 kVAr, 562.8 kVAr, and 548.5 kVAr, respectively. The SLD of 33 bus RDS for Case 2 after incorporating the shunt capacitors is depicted in Fig. 6.

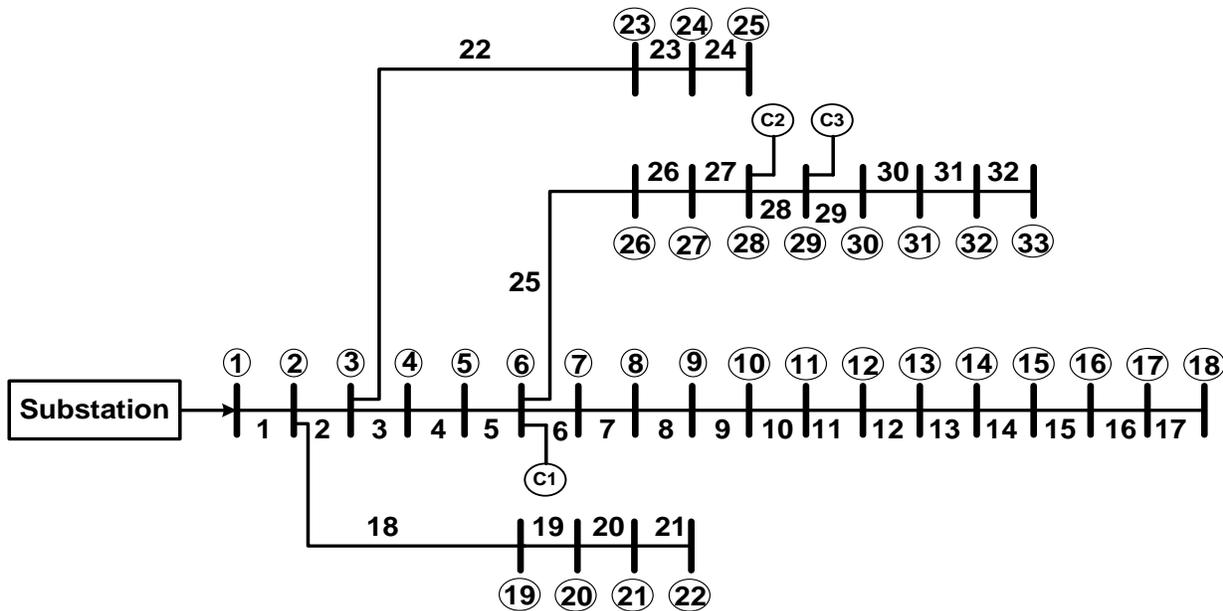


Fig. 6. SLD of 33 Bus RDS for Case 2 after Incorporating the Shunt Capacitors.

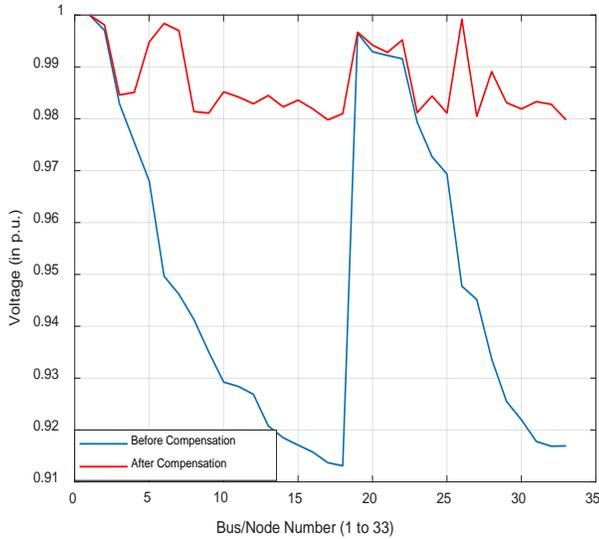


Fig. 7. Comparison of Node Voltages before and after Reconfiguration for a 33 Bus RDS.

The active power loss obtained in this case is 140.83 kW which resulted in the reduction of 30.51% compared to the case without any compensation. The results obtained with GSA are also compared with DEA, and the results are reported in Table II. Voltage profile of RDS has been improved after incorporating the shunt capacitors. Comparison of node voltages before and after the reconfiguration for 33 bus RDS is depicted in Fig. 7. Minimum voltage after installing the capacitors is found to be 0.9798 p.u.

C. Case 3

In this case, proposed optimization problem of RDS is solved by optimally allocating the DG units at various buses in the system. Table III presents the simulation results for Case 3. In this case, DG of 835.8 kW is placed at bus 8, 1105.6 kW is placed at bus 23, and 1084.9 kW is placed at bus 29. Hence, the total size of DG is 3026.3 kW. By optimally placing the DG units, the system power losses have been decreased to 84.62 kW from the base case loss of 202.66 kW. Therefore, in this case, there is a loss reduction of 58.25% compared to base case losses. Fig. 8 depicts the improved voltage profiles for case 3. The minimum voltage obtained in this case after DG allocation is 0.9742 p.u. which has been occurred at bus number 17, whereas in the base case, the minimum voltage occurred is 0.9038 p.u. at bus number 18.

TABLE III. SIMULATION RESULTS FOR 33 BUS RDS FOR CASES 3 AND 4

	Case 3	Case 4
Tie switches opened for reconfiguration	33, 34, 35, 36, 37	7, 9, 14, 32, 37
Bus number and DG size (kW)	835.8 kW at bus 8	822.4 kW at bus 13
	1105.6 kW at bus 23	1150.5 kW at bus 23
	1084.9 kW at bus 29	983.2 kW at bus 30
Total DG size (kW)	3026.3 kW	2956.1 kW
Bus number and reactive power compensation (kVAr)	---	560.5 kVAr at bus 6
	---	565.2 kVAr at bus 28
	---	540.8 kVAr at bus 29
Total compensation (kVAr)	---	1666.5 kVAr
Power loss (kW)	84.62 kW	55.98 kW
Power loss reduction	58.25 %	72.38 %
Minimum voltage	0.9742 pu at bus 17	0.9802 pu at bus 17

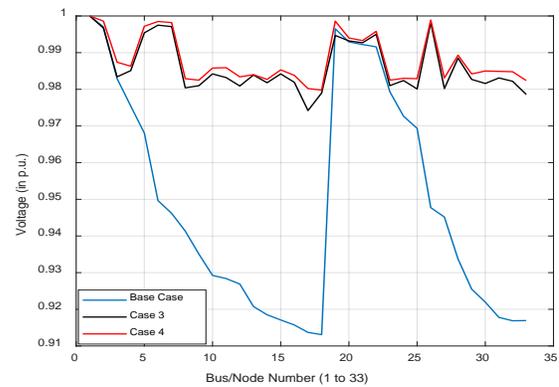


Fig. 8. Comparison of Node Voltages for the base Case, Case 3, and Case 4 for 33 Bus RDS.

D. Case 4

This case considers network reconfiguration, DG allocation, and reactive power compensation simultaneously for minimizing the system power losses. Table III presents the simulation results for Case 4. The opened tie-switches for ONR are 7, 9, 14, 32, and 37, and the system configuration for case 4 after the ONR has been depicted in Fig. 9. The total optimum DG size obtained is 2956.1 kW. The buses 13, 23, and 30 are placed with DG units with capacities of 822.4 kW, 1150.5 kW, and 983.2 kW, respectively. Here, the shunt capacitors are placed at buses 6, 28, and 29 with the reactive power compensation of 560.5 kVAr, 565.2 kVAr, and 540.8 kVAr, respectively.

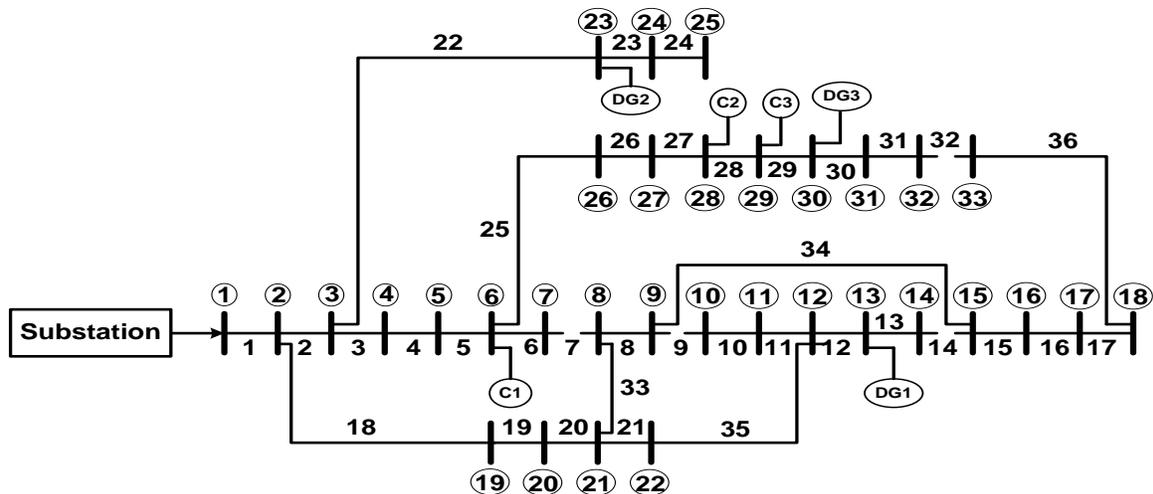


Fig. 9. Final Configuration of a 33 Bus RDS for Case 4 after Incorporating the Shunt Capacitors and DG Units.

The total optimum reactive power compensation required for Case 4 is 1666.5 kVar. The obtained optimum loss is 55.98 kW which is 72.38% less when compared to base case power loss (i.e., 202.66 kW). The voltage profile obtained in this case has been depicted in Fig. 8. From this figure, it can be observed that the voltage profile obtained in this case is better than all other cases studied in this work. Final configuration of a 33 bus RDS for Case 4 after incorporating the shunt capacitors and DG units has been depicted in Fig. 9. Minimum voltage occurred in this case after the ONR, DG allocation and reactive power compensation is 0.9802 p.u. at bus number 17, which has been improved from 0.9038 p.u. (i.e., base case).

VI. CONCLUSIONS

This paper proposed an approach for the optimal allocation of distributed generations (DGs), shunt capacitors along optimal network reconfiguration using the meta-heuristic-based gravitational search algorithm (GSA). Here, the minimization of total active power losses is considered as an objective function, which will also enhance voltage profile in the system and hence reduces voltage deviation. The proposed problem has been implemented on the standard 33 bus radial distribution system (RDS). The obtained results show the improved voltage profile and reduced power losses in the system. The proposed work can also be extended to unbalanced RDSs and the optimal allocation of electric vehicles (EVs) can also be studied. Optimal allocation of energy storage systems (ESSs) along with FACTS devices such as D-STATCOM is scope for future research.

ACKNOWLEDGMENT

This research work was funded by “Woosong University’s Academic Research Funding – 2021”.

REFERENCES

- [1] V.A. Evangelopoulos, P.S. Georgilakis, N.D. Hatziaargyriou, “Optimal operation of smart distribution networks: A review of models, methods and future research”, *Electric Power Systems Research*, vol. 140, pp. 95-106, 2016.
- [2] S. Haghifam, M. Dadashi, K. Zare, H. Seyedi, “Optimal operation of smart distribution networks in the presence of demand response aggregators and microgrid owners: A multi follower Bi-Level approach”, *Sustainable Cities and Society*, vol. 55, 2020.
- [3] L.H. Macedo, J.F. Franco, M.J. Rider, R. Romero, “Optimal Operation of Distribution Networks Considering Energy Storage Devices”, *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2825-2836, Nov. 2015.
- [4] S.R. Salkuti, “Feeder Reconfiguration in Unbalanced Distribution System with Wind and Solar Generation using Ant Lion Optimization”, *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 31-39, 2021.
- [5] H. Zhang, D. Zhao, C. Gu, F. Li, B. Wang, “Economic optimization of smart distribution networks considering real-time pricing”, *Journal of Modern Power Systems and Clean Energy*, vol. 2, pp. 350-356, 2014.
- [6] A.K. Fard, A. Khodaei, “Multi-objective optimal operation of smart reconfigurable distribution grids”, *AIMS Energy*, vol. 4, no. 2, pp. 206-221, 2016.
- [7] H.M.A. Ahmed, M.H. Ahmed, M.M.A. Salama, “Network Reconfiguration for the Optimal Operation of Smart Distribution Systems”, *IEEE Power & Energy Society General Meeting*, 2019, pp. 1-5.
- [8] R. Afshan, J. Salehi, “Optimal operation of distribution networks with presence of distributed generations and battery energy storage systems considering uncertainties and risk analysis”, *Journal of Renewable and Sustainable Energy*, vol. 9, no. 1, 2017.
- [9] S. Ouali, A. Cherkaoui, “Optimal Allocation of Combined Renewable Distributed Generation and Capacitor Units for Interconnection Cost Reduction”, *Journal of Electrical and Computer Engineering*, 2020.
- [10] A. Uniyal, S. Sarangi, “Optimal network reconfiguration and DG allocation using adaptive modified whale optimization algorithm considering probabilistic load flow”, *Electric Power Systems Research*, vol. 192, 2021.
- [11] R. Fathi, B. Tousi, S. Galvani, “A new approach for optimal allocation of photovoltaic and wind clean energy resources in distribution networks with reconfiguration considering uncertainty based on info-gap decision theory with risk aversion strategy”, *Journal of Cleaner Production*, vol. 295, 2021.
- [12] M.I. Pathan, M. Al-Muhaini, S.Z. Djokic, “Optimal reconfiguration and supply restoration of distribution networks with hybrid microgrids”, *Electric Power Systems Research*, vol. 187, 2020.
- [13] A.J. Nowdeh, M. Babanezhad, S.A. Nowdeh, A. Naderipour, H. Kamyab, Z.A. Malek, V.K. Ramachandaramurthy, “Meta-heuristic matrix moth-flame algorithm for optimal reconfiguration of distribution networks and placement of solar and wind renewable sources

- considering reliability”, *Environmental Technology & Innovation*, vol. 20, 2020.
- [14] S.R. Salkuti, “Multi-objective based Optimal Network Reconfiguration using Crow Search Algorithm,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 86-95, 2021.
- [15] A.O. Salau, Y.W. Gebru, D. Bitew, “Optimal network reconfiguration for power loss minimization and voltage profile enhancement in distribution systems”, *Heliyon*, vol. 6, no. 6, 2020.
- [16] S.R. Salkuti, “Optimal Allocation of DG and D-STATCOM in a Distribution System using Evolutionary based Bat Algorithm”, *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 4, pp. 360-365, 2021.
- [17] K. Nagaraju, S. Sivanagaraju, T. Ramana, P.V. Prasad, “A novel load flow method for radial distribution systems for realistic loads”, *Electric Power Components and Systems*, vol. 39, no. 2, pp. 128-141, 2011.
- [18] I.B. Hamida, S.B. Salah, F. Msahli, M.F. Mimouni, “Optimal network reconfiguration and renewable DG integration considering time sequence variation in load and DGs”, *Renewable Energy*, vol. 121, pp. 66-80, 2018.
- [19] S.R. Salkuti, “Optimal location and sizing of DG and D-STATCOM in distribution networks”, *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 16, no. 3, pp. 1107-1114, Dec. 2019.
- [20] K. Mahmoud, N. Yorino, A. Ahmed, “Optimal Distributed Generation Allocation in Distribution Systems for Loss Minimization”, *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 960-969, Mar. 2016.
- [21] R. Sanjay, T. Jayabarathi, T. Raghunathan, V. Ramesh, N. Mithulananthan, “Optimal Allocation of Distributed Generation Using Hybrid Grey Wolf Optimizer”, *IEEE Access*, vol. 5, pp. 14807-14818, 2017.
- [22] K. Mahmoud, M. Lehtonen, “Simultaneous Allocation of Multi-Type Distributed Generations and Capacitors Using Generic Analytical Expressions”, *IEEE Access*, vol. 7, pp. 182701-182710, 2019.
- [23] B.R. Pereira, G.R.M. da Costa, J. Contreras, J.R.S. Mantovani, “Optimal Distributed Generation and Reactive Power Allocation in Electrical Distribution Systems”, *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 975-984, July 2016.
- [24] S. Ganguly, D. Samajpati, “Distributed Generation Allocation on Radial Distribution Networks Under Uncertainties of Load and Generation Using Genetic Algorithm”, *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 688-697, July 2015.
- [25] S.R. Salkuti, Y.H. Lho, “Optimum Location of Voltage Regulators in the Radial Distribution Systems”, *International Journal of Emerging Electric Power Systems*, vol. 17, no. 3, pp. 351-361, Jun. 2016.
- [26] E. Rashedi, H. Nezamabadi-pour, S. Saryazdi, “GSA: A Gravitational Search Algorithm”, *Information Sciences*, vol. 179, no. 13, pp. 2232-2248, vol. 179, no. 13, 2009.
- [27] R.K. Swain, N.C. Sahu, P.K. Hota, “Gravitational Search Algorithm for Optimal Economic Dispatch”, *Procedia Technology*, vol. 6, pp. 411-419, 2012.
- [28] Z. Younes, I. Alhamrouni, S. Mekhilef, M. Rezasudin, “A memory-based gravitational search algorithm for solving economic dispatch problem in micro-grid”, *Ain Shams Engineering Journal*, 2021.
- [29] B. Shaw, V. Mukherjee, S.P. Ghoshal, “Solution of reactive power dispatch of power systems by an opposition-based gravitational search algorithm”, *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 29-40, 2014.
- [30] B. Venkatesh, R. Ranjan, “Optimal radial distribution system reconfiguration using fuzzy adaptation of evolutionary programming”, *International Journal of Electrical Power Energy Systems*, vol. 25, no. 10, pp. 775-780, Dec. 2013.
- [31] A. Bayat, A. Bagheri, R. Noroozian, “Electrical Power and Energy Systems Optimal siting and sizing of distributed generation accompanied by reconfiguration of distribution networks for maximum loss reduction by using a new UVDA-based heuristic method”, *International Journal of Electrical Power and Energy Systems*, vol. 77, pp. 360-371, 2016.
- [32] A.M. Imran, M. Kowsalya, “Electrical Power and Energy Systems A new power system reconfiguration scheme for power loss minimization and voltage profile enhancement using Fireworks Algorithm”, *International Journal of Electrical Power and Energy Systems*, vol. 62, pp. 312-322, 2014.

Analyzing the Performance of Stroke Prediction using ML Classification Algorithms

Gangavarapu Sailasya¹, Gorli L Aruna Kumari²

Department of Computer Science and Engineering
GITAM Institute of Technology, GITAM (Deemed to be University)
Visakhapatnam, Andhra Pradesh – 530045

Abstract—A Stroke is a health condition that causes damage by tearing the blood vessels in the brain. It can also occur when there is a halt in the blood flow and other nutrients to the brain. According to the World Health Organization (WHO), stroke is the leading cause of death and disability globally. Most of the work has been carried out on the prediction of heart stroke but very few works show the risk of a brain stroke. With this thought, various machine learning models are built to predict the possibility of stroke in the brain. This paper has taken various physiological factors and used machine learning algorithms like Logistic Regression, Decision Tree Classification, Random Forest Classification, K-Nearest Neighbors, Support Vector Machine and Naïve Bayes Classification to train five different models for accurate prediction. The algorithm that best performed this task is Naïve Bayes that gave an accuracy of approximately 82%.

Keywords—Stroke; machine learning; logistic regression; decision tree classification; random forest classification; k-nearest neighbors; support vector machine; Naïve Bayes classification

I. INTRODUCTION

According to the Centers for Disease Control and Prevention (CDC), stroke is the fifth-leading cause [1] of death in the United States. Stroke is a non-communicable infection that is liable for around 11% of total deaths. Consistently, over 795,000 individuals in the United States experience the ill effects of a stroke [2]. It is the fourth significant reason for death in India.

With the advancement of technology in the medical field, predicting the occurrence of a stroke can be made using Machine Learning. The algorithms present in Machine Learning are constructive in making an accurate prediction and give correct analysis. The works previously performed on stroke mostly include the ones on Heart stroke prediction. Very less works have been performed on Brain stroke. This paper is based on predicting the occurrence of a brain stroke using Machine Learning. The key components of the approaches used and results obtained are that among the five different classification algorithms used Naïve Bayes has best performed obtaining a higher accuracy metric. The limitation with this model is that it is being trained on textual data and not on real time brain images. The paper shows the implementation of six Machine Learning classification algorithms. This paper can be further extended to implementing all the current machine learning algorithms.

A dataset is chosen from Kaggle [3] with various physiological traits as its attributes to proceed with this task.

These traits are later analyzed and used for the final prediction. The dataset is initially cleaned and made ready for the machine learning model to understand. This step is called Data Preprocessing. For this, the dataset is checked for null values and fill them. Then Label encoding is performed to convert string values into integers followed by one-hot encoding, if necessary.

After Data Preprocessing, the dataset is split into train and test data. A model is then built using this new data using various Classification Algorithms. Accuracy is calculated for all these algorithms and compared to get the best-trained model for prediction.

After training the model and calculating the accuracy, an HTML page and a Flask application are developed. The web application is for the user to enter the values for prediction. The flask application is a framework that connects the trained model and the web application. After proper analysis, the paper concludes which algorithm is most appropriate for the prediction of stroke.

II. LITERATURE SURVEY

In [4], stroke prediction was made on Cardiovascular Health Study (CHS) dataset using five machine learning techniques. As an optimal solution, the authors used a combination of the Decision Tree with the C4.5 algorithm, Principal Component Analysis, Artificial Neural Networks, and Support Vector Machine. But the CHS Dataset taken for this work had a smaller number of input parameters.

In [5], stroke prediction has been carried out from the social media posts posted by people. In this particular work, the authors have used the DRFS method to find the various symptoms associated with stroke disease. The usage of Natural Language Processing to extract the text from the social media posts adds up to the overall execution time of the model which is not desirable.

In [6], the authors have performed the task of stroke prediction by using an improvised random forest algorithm. This was used to analyze the levels of risks obtained with the strokes. As suggested by the authors, this method is said to have performed better when compared to the existing algorithms. This particular research is limited to very few types of strokes and cannot be used for any new stroke type in the future.

Research paper [7] shows that the model was trained using Decision Tree, Random Forest, and Multi-layer perceptron for stroke prediction. The obtained accuracies for the three methods were quite close, with slight differences. The calculated accuracy for Decision Tree was 74.31%, Random Forest was 74.53%, and Multi-layer perceptron was 75.02%.

This paper suggests that Multi-layer perceptron is more accurate than the other two methods. Accuracy score was the only metric used for calculating the performance that might not always give favorable results.

Research carried out in [8] shows the implementation of machine learning model to predict heart stroke. They used various machine learning techniques like Decision tree, Naïve Bayes, SVM to build the model and later compared their performance. They obtained a maximum accuracy of 60% from the used algorithms which is pretty less.

In [9], the authors have used different data mining classification techniques to predict the possibility of a stroke. The dataset was taken from the Ministry of National Guards Health Affairs Hospitals, Kingdom of Saudi Arabia. The three classification algorithms used were C4.5, Jrip and Multi layers perceptron (MLP). With these algorithms, the model obtained an accuracy of around 95%. Even though the paper claims to obtain an accuracy of 95%, the time taken for training and predicting is higher as the authors have used a combination of complex algorithms.

Research carried out in [10], suggests the usage of three different algorithms to predict the possibility of stroke. These algorithms are Naïve Bayes, Decision Tree, and Neural Networks. This paper concluded that the Decision tree has the highest accuracy (about 75%) of the other two algorithms. But this model could not suit the real-world examples based on the values obtained from the confusion matrix.

In [11], the researchers have performed stroke prediction on Cardiovascular Health Study (CHS) dataset. They proposed a novel automatic feature selection algorithm that selects robust features based on their proposed conservative mean. They have combined this method with the Support Vector Machine algorithm for better efficiency. But this resulted in the generation of a number of vectors that tend to reduce the performance of the model.

Research in [12] proposes the prediction of thrombo-embolic stroke disease using Artificial Neural Networks. The method used for prediction was the Back-propagation algorithm. This model was able to get an accuracy of around 89%. But Neural Networks need more time to be trained and require higher processing time because of the complex structure with increasing number of neurons.

III. SYSTEM METHODOLOGY

To proceed with the implementation, different datasets were considered from Kaggle. Out of all the existing datasets, an appropriate dataset was collected for model building.

After collecting the dataset, next step lies in preparing the dataset to make the data more clear and easily understood by the machine. This step is called as Data preprocessing. This

step includes handling of missing values, handling imbalanced data and performing label encoding that are specific for this particular dataset.

Now that the data is preprocessed, it is ready for model building. For model building, preprocessed dataset along with machine learning algorithms are required. Logistic Regression, Decision Tree Classification algorithm, Random Forest Classification algorithm, K-Nearest Neighbor algorithm, Support Vector Classification and Naïve Bayes Classification algorithm are used. After building six different models, they are compared using five accuracy metrics namely Accuracy Score, Precision Score, Recall Score, F1 Score and Receiver Operating Characteristic (ROC) curve.

The comparison of the models gives the best model in terms of the accuracy metrics to proceed with the deployment phase. For deploying the model, an HTML page is developed to make it user-friendly for the user to enter the input parameters and get the result. The parameters entered by the user are sent to the model using a flask application which is basically a python framework that links the web application and the model together. The model takes the input parameters, predicts the output and returns the result to the flask application. Now this flask will display that result on the web page for the user to check the result.

The flow chart of the proposed system's methodology is in Fig. 1.

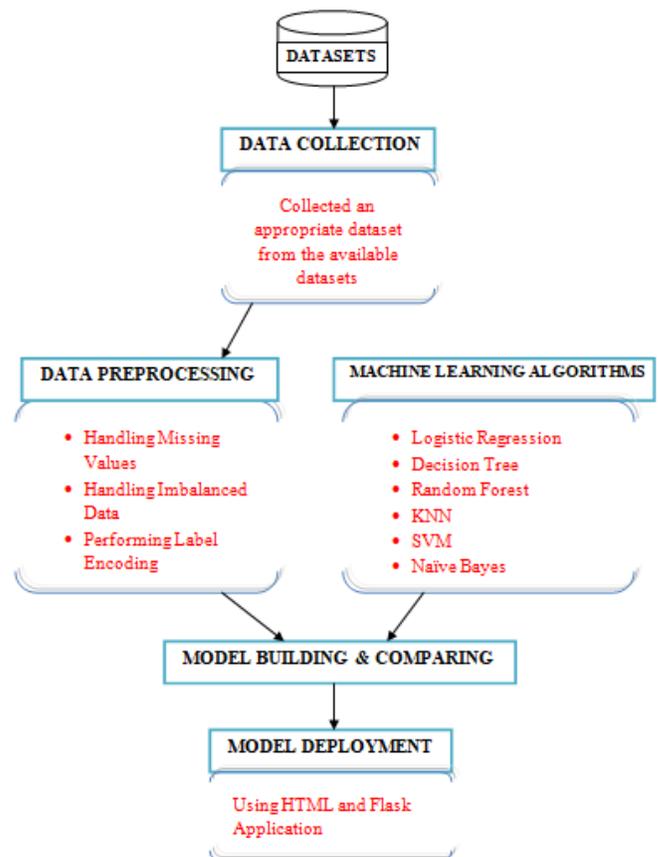


Fig. 1. Proposed System's Flow Chart.

IV. IMPLEMENTATION

The implementation of this project is as follows.

A. Dataset

The dataset for stroke prediction is from Kaggle [3]. This particular dataset has 5110 rows and 12 columns. The columns have 'id', 'gender', 'age', 'hypertension', 'heart_disease', 'ever_married', 'work_type', 'Residence_type', 'avg_glucose_level', 'bmi', 'smoking_status' and 'stroke' as the main attributes. The output column 'stroke' has the value as either '1' or '0'. The value '0' indicates no stroke risk detected, whereas the value '1' indicates a possible risk of stroke. This dataset is highly imbalanced as the possibility of '0' in the output column ('stroke') outweighs that of '1' in the same column. Only 249 rows have the value '1' whereas 4861 rows with the value '0' in the stroke column. For better accuracy, data pre-processing is performed to balance the data. The dataset discussed above is summarized in Table 1.

TABLE I. STROKE DATASET

Attribute Name	Type (Values)	Description
1. id	Integer	A unique integer value for patients
2. gender	String literal (Male, Female, Other)	Tells the gender of the patient
3. age	Integer	Age of the Patient
4. hypertension	Integer (1, 0)	Tells whether the patient has hypertension or not
5. heart_disease	Integer (1, 0)	Tells whether the patient has heart disease or not
6. ever_married	String literal (Yes, No)	It tells whether the patient is married or not
7. work_type	String literal (children, Govt_job, Never_worked, Private, Self-employed)	It gives different categories for work
8. Residence_type	String literal (Urban, Rural)	The patient's residence type is stored
9. avg_glucose_level	Floating point number	Gives the value of average glucose level in blood
10. bmi	Floating point number	Gives the value of the patient's Body Mass Index
11. smoking_status	String literal (formerly smoked, never smoked, smokes, unknown)	It gives the smoking status of the patient
12. stroke	Integer (1, 0)	Output column that gives the stroke status

B. Data Preprocessing

Data Preprocessing is required before model building to remove the unwanted noise and outliers from the dataset, resulting in a deviation from proper training. Anything that interrupts the model from performing with less efficiency is taken care of in this stage. After collecting the appropriate dataset, the next step lies in cleaning the data and making sure that it is ready for model building. The dataset taken has 12 attributes, as mentioned in Table I. Firstly, the column 'id' is dropped because its existence does not make much difference in model building. Then the dataset is checked for null values and filled if any found. In this case, the column 'bmi' has null values filled with the mean of the column data. After removing the null values from the dataset, the next task is Label Encoding.

C. Label Encoding

Label encoding encodes the string literals in the dataset into integer values for the machine to understand them. As the machine is usually trained in numbers, the strings have to be converted into integers. There are five columns in the collected dataset that have strings as their data type. On performing label encoding, all the strings get encoded, and the entire dataset becomes a combination of numerals.

D. Handling Imbalanced Data

The dataset chosen for the task of stroke prediction is highly imbalanced. The entire dataset has 5110 rows, of which 249 rows are suggesting the occurrence of a stroke and 4861 rows having the possibility of no stroke. The graphical representation of the imbalance is in Fig. 2. Training a machine-level model with such data might give accuracy, but other accuracy metrics like precision and recall are shallow. If such imbalanced data is not handled, the results are not accurate, and the prediction is inefficient. Therefore, to get an efficient model, this imbalanced data is to be first handled. For this purpose, the method of undersampling is used. Undersampling [13] balances the data wherein the majority class is undersampled to match the minority class. In this case, the class with a value as '0' is undersampled for the class with the value '1'. So after undersampling the resulting dataset will have 249 rows with value '0' and 249 rows with value '1'. The graphical representation of the output column in the resulting dataset is as shown in Fig. 3.

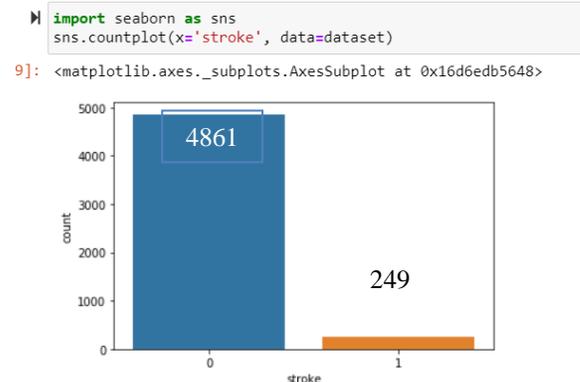


Fig. 2. Before Undersampling.

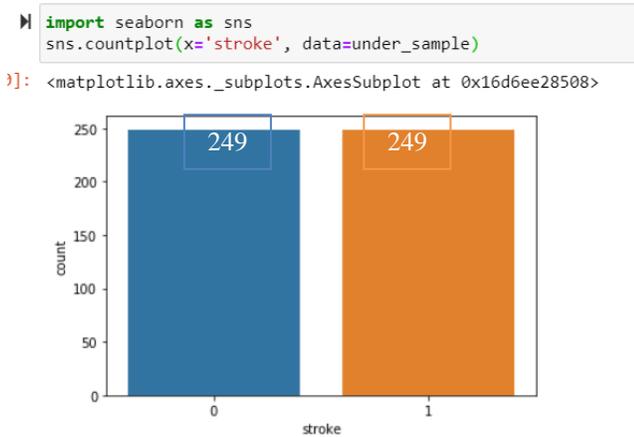


Fig. 3. After Undersampling.

V. MODEL BUILDING

A. Splitting the Data

After completing data preprocessing and handling the imbalanced dataset, the next step is building the model. The undersampled data is split into training and testing data for better accuracy and efficiency for this task keeping the ratio as 80% training data and 20% testing data. After splitting, various classification algorithms are used to train the model. The classification algorithms used for this purpose are Logistic Regression, Decision Tree Classification algorithm, Random Forest Classification, K-Nearest Neighbors Classification, Support Vector Machine and Naïve Bayes Classification.

B. Classification Algorithms

1) *Logistic regression*: Logistic Regression [14] is a supervised learning algorithm used for predicting the probability of the output variable. This algorithm is the best fit when the output variable has binary values (0 or 1). As the output attribute in the dataset has only two possible values, Logistic Regression is opted. After performing this algorithm on the dataset, the accuracy obtained is 78%. Efficiency of this algorithm can also be found by using various other accuracy metrics like precision score and recall score. Those two scores obtained in this case are equal, having a value of 77.6%. The F1 Score obtained with this algorithm is 77.6%. The Receiver operating characteristic (ROC) curve for Logistic Regression is 78% as shown in Fig. 4.

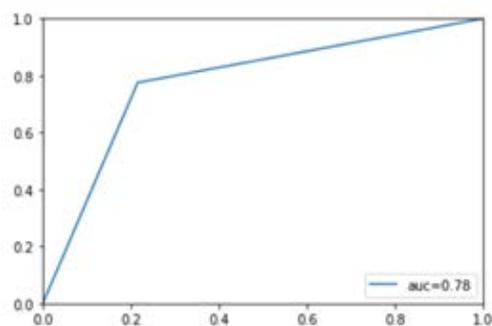


Fig. 4. ROC Curve for Logistic Regression.

2) *Decision tree classification*: Decision Tree classification [15] is to solve both Regression and classification problems. This algorithm is also a supervised learning method wherein the input variables already have their corresponding output variable. It is a tree-like structure. In this algorithm, the data continuously splits according to a particular parameter. A decision tree has two parts: Decision Node and Leaf node. The data splits at the former node, and the latter is the node that gives the outcome. In this case of stroke prediction, the decision tree classification algorithm obtained an accuracy of 66%, which is less than that obtained using logistic Regression. Similar to logistic Regression, the precision and recall scores are the same and equal to 77.6%. The F1 Score obtained with this algorithm is 77.6%. The Receiver operating characteristic (ROC) curve for Decision Tree Classification is 66% as shown in Fig. 5.

3) *Random forest classification*: The following classification algorithm chosen is Random Forest Classification [16]. Random Forests are composed of multiple independent decision trees trained independently on a random subset of data. These trees are generated at the time of training, and the outputs are obtained from each decision tree. For the final prediction from this algorithm, a method called "voting" takes place. This method means that each decision tree votes for an output class (in this case, the two classes are: 'stroke' and 'no stroke'). The random forest chooses the class with the maximum number of votes as the final prediction. The accuracy obtained by training the model using this particular algorithm is 73%. The precision and recall scores are 72% and 73.5%, respectively. The F1 Score obtained with this algorithm is 72.7%. The Receiver operating characteristic (ROC) curve for Random Forest Classification is 73% as shown in Fig. 6.

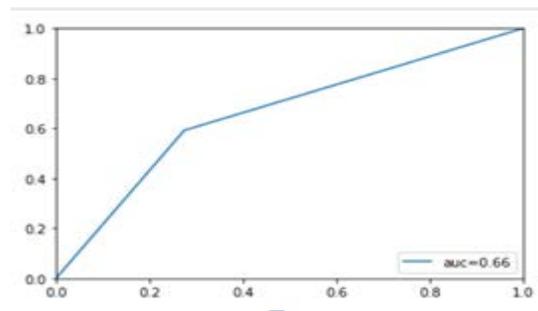


Fig. 5. ROC Curve for Decision Tree Classification.

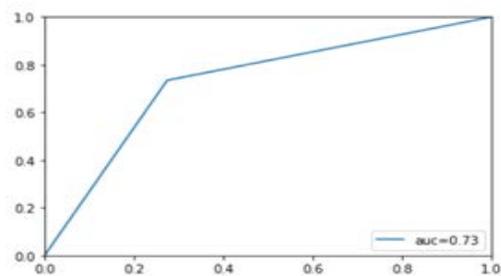


Fig. 6. ROC Curve for Random Forest Classification.

4) *K-nearest neighbors classification*: Another algorithm used for classification is K-Nearest Neighbors (KNN). It is also a supervised learning technique. KNN [17] is a lazy algorithm that would not train immediately on giving the dataset. Instead, it stores the dataset, and at the time of classification, it acts on the dataset. The working principle of KNN is to find similarities between the new case (or data) and available data and then map the new case into the category that is most similar to the available categories. The accuracy obtained is 80%. Precision and recall scores are 77.4% and 83.7%, respectively. The F1 Score obtained with this algorithm is 80.4%. The Receiver operating characteristic (ROC) curve for KNN is 80% as shown in Fig. 7.

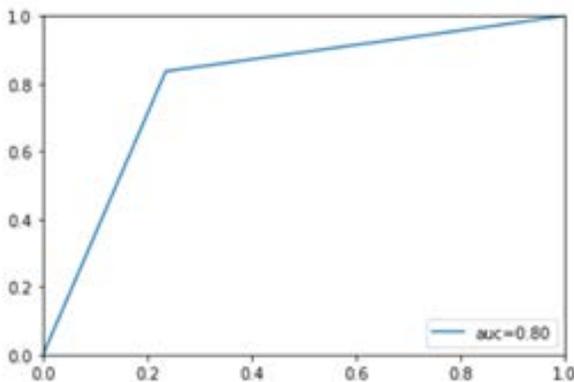


Fig. 7. ROC Curve for KNN.

5) *Support vector machine*: It is a supervised learning technique that can be associated with learning algorithms to analyze the data for both classification and regression. Support Vector Machine (SVM) [18] scales relatively well to high dimensional data. For this particular dataset, the algorithm obtained an accuracy of 80%, with precision and recall score being 78.6% and 83.8%. The F1 score is 81.1% for this algorithm. The Receiver operating characteristic curve (ROC) for Support Vector Classification is 80% as shown in Fig. 8.

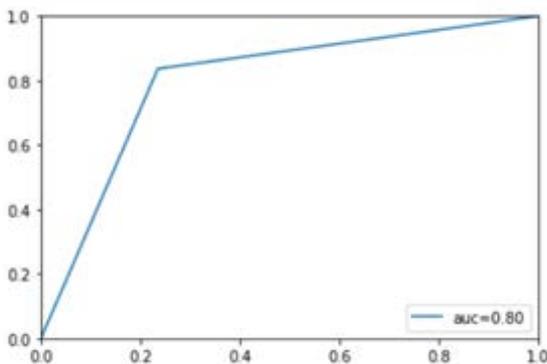


Fig. 8. ROC Curve for Support Vector Machine.

6) *Naïve bayes classification*: It is also a supervised learning technique. A Naïve Bayes classifier [19] assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. It is based on Bayes Theorem.

This algorithm follows the principle that 'every feature or attribute classified is independent of one another.' With this algorithm, the accuracy obtained was 82%, with the precision score being 79.2% and recall score being 85.7%. The F1 Score obtained with this algorithm is 82.3%. The Receiver operating characteristic (ROC) curve for Naïve Bayes Classification is 82% as shown in Fig. 9.

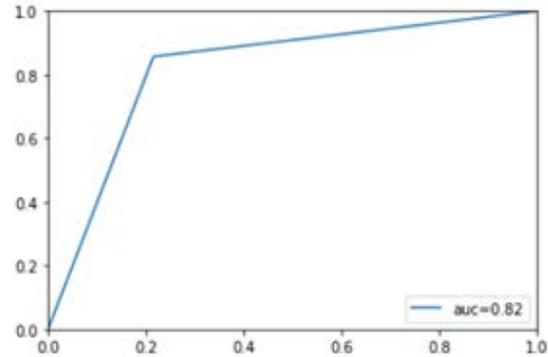
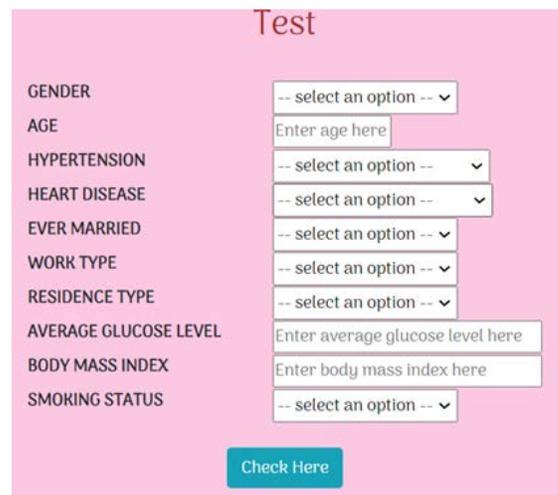


Fig. 9. ROC Curve for Naïve Bayes Classification.

After model building, it can be concluded that Naïve Bayes has performed better when compared to other algorithms. So, the model trained using Naïve Bayes classification is dumped using pickle. The next task is to develop a web application and a flask application to enter the input parameters.

The web page is built using simple HTML code. This application has an input form that will take the entered input values from the user to predict the occurrence of stroke. When the user clicks on the 'Check Here' button, the entered parameters are given to the flask application. A part of the HTML page is shown in Fig. 10.

The flask application which is basically a python code is the bridge between the web page and the trained machine learning model as shown in Fig. 11. The input values are sent to the flask application which in sends the values to the model for prediction.



The figure shows a web form titled 'Test' for inputting parameters to predict stroke occurrence. The form includes the following fields:

- GENDER: A dropdown menu with the option '-- select an option --'.
- AGE: A text input field with the placeholder 'Enter age here'.
- HYPERTENSION: A dropdown menu with the option '-- select an option --'.
- HEART DISEASE: A dropdown menu with the option '-- select an option --'.
- EVER MARRIED: A dropdown menu with the option '-- select an option --'.
- WORK TYPE: A dropdown menu with the option '-- select an option --'.
- RESIDENCE TYPE: A dropdown menu with the option '-- select an option --'.
- AVERAGE GLUCOSE LEVEL: A text input field with the placeholder 'Enter average glucose level here'.
- BODY MASS INDEX: A text input field with the placeholder 'Enter body mass index here'.
- SMOKING STATUS: A dropdown menu with the option '-- select an option --'.

A blue 'Check Here' button is located at the bottom of the form.

Fig. 10. Input Form in HTML Code.

```
from flask import Flask, request, render_template
import numpy as np
import pickle

model=pickle.load(open('stroke_model.pkl','rb'))

app = Flask(__name__)
@app.route('/')
def home():
    return render_template("home.html")
```

Fig. 11. Flask Application Linking the Model and Web Page.

Once the machine learning model gets the input parameters, it predicts the output. The obtained result then is reflected back on the web page through the flask application for the user to see the prediction.

VI. CONCLUSION

Stroke is a critical medical condition that should be treated before it worsens. Building a machine learning model can help in the early prediction of stroke and reduce the severe impact of the future. This paper shows the performance of various machine learning algorithms in successfully predicting stroke based on multiple physiological attributes. Out of all the algorithms chosen, Naïve Bayes Classification performs best with an accuracy of 82%. The comparison of accuracies obtained from various algorithms is as shown in Fig. 12. Among all the precision, recall and F1 scores obtained, Naïve Bayes has performed better. The comparison of Precision score, recall score and F1 score is as shown in Fig. 13, Fig. 14 and Fig. 15.

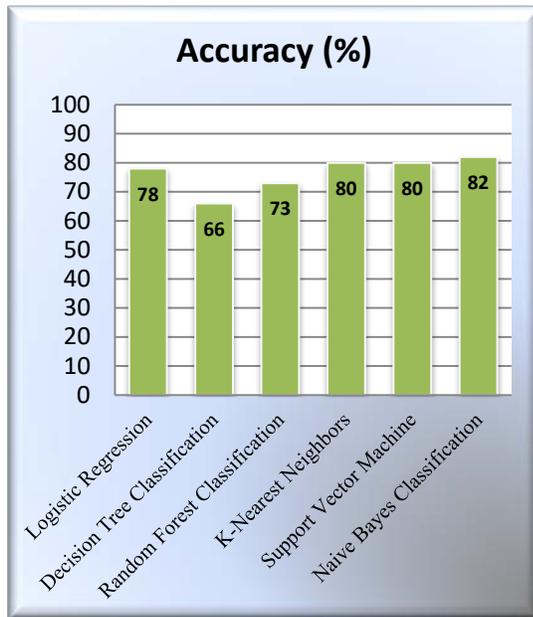


Fig. 12. Comparing the Accuracies of ML Algorithms.

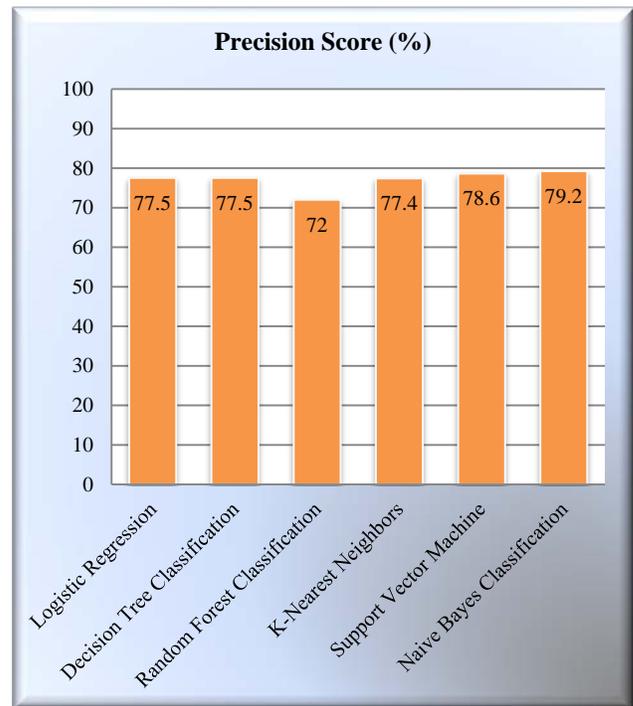


Fig. 13. Comparing the Precision Scores of ML Algorithms.

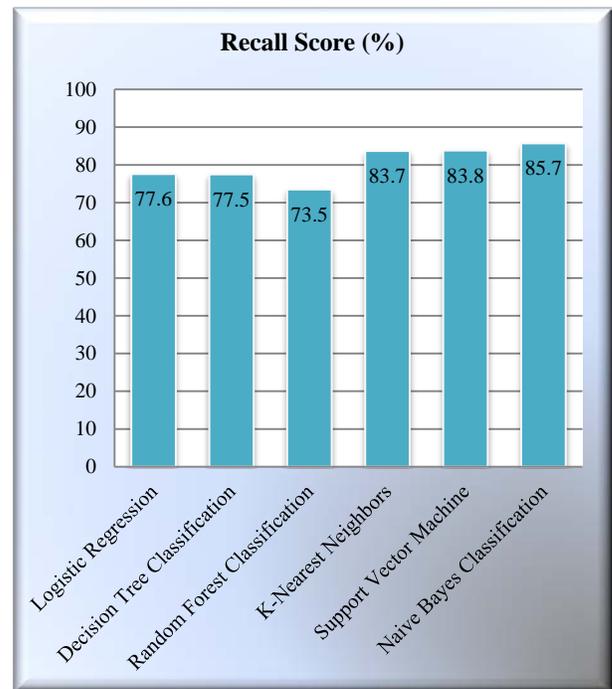


Fig. 14. Comparing the Recall Scores of ML Algorithms.

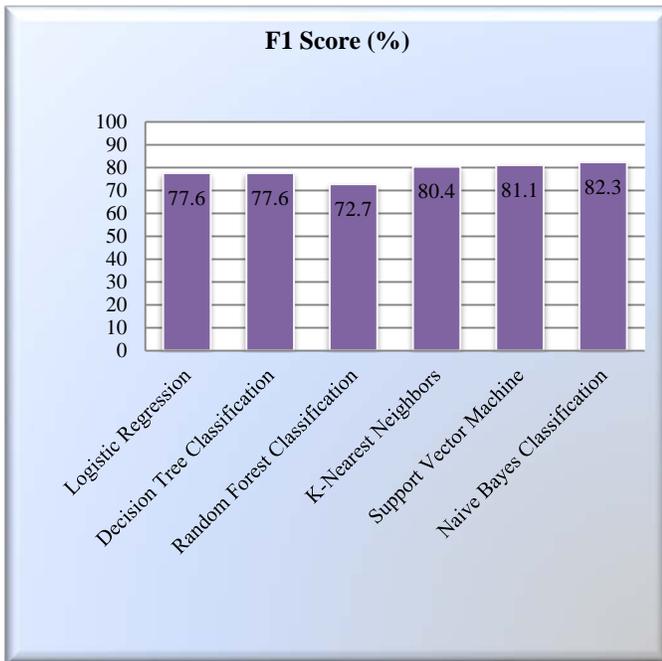


Fig. 15. Comparing the F1 Scores of ML Algorithms.

This paper suggests the implementation of various Machine learning algorithms on the dataset taken. This project can be further extended by training the model using Neural Networks. The comparison of the performance can be done by taking more accuracy metrics into consideration. This work is limited to textual data, which might not always be accurate for stroke prediction. Collecting a dataset consisting of images such as Brain CT scans to predict the possibility of stroke would be more efficient in the future.

REFERENCES

- [1] Concept of Stroke by Healthline.
- [2] Statistics of Stroke by Centers for Disease Control and Prevention..
- [3] Dataset named 'Stroke Prediction Dataset' from Kaggle: <https://www.kaggle.com/fedesoriano/stroke-prediction-dataset>.
- [4] Singh, M.S., Choudhary, P., Thongam, K.: A comparative analysis for various stroke prediction techniques. In: Springer, Singapore (2020).
- [5] Pradeepa, S., Manjula, K. R., Vimal, S., Khan, M. S., Chilamkurti, N., & Luhach, A. K.: DRFS: Detecting Risk Factor of Stroke Disease from Social Media Using Machine Learning Techniques. In Springer (2020).
- [6] Vamsi Bandi, Debnath Bhattacharyya, Divya Midhunchakkavarthy: Prediction of Brain Stroke Severity Using Machine Learning. In: International Information and Engineering Technology Association (2020).
- [7] Nwosu, C.S., Dev, S., Bhardwaj, P., Veeravalli, B., John, D.: Predicting stroke from electronic health records. In: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE (2019).
- [8] Fahd Saleh Alotaibi: Implementation of Machine Learning Model to Predict Heart Failure Disease. In: International Journal of Advanced Computer Science and Applications (IJACSA) (2019).
- [9] Ohoud Almadani, Riyad Alshammari: Prediction of Stroke using Data Mining Classification Techniques. In: International Journal of Advanced Computer Science and Applications (IJACSA) (2018).
- [10] Kansadub, T., Thammaboosadee, S., Kiattisin, S., Jalayondeja, C.: Stroke risk prediction model based on demographic data. In: 8th Biomedical Engineering International Conference (BMEiCON) IEEE (2015).
- [11] Aditya Khosla, Yu Cao, Cliff Chiung-Yu Lin, Hsu-Kuang Chiu, Junling Hu, Honglak Lee: An Integrated Machine Learning Approach to Stroke Prediction. In: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (2010).
- [12] Shanthy, D., Sahoo, G., Saravanan, N.: Designing an artificial neural network model for the prediction of thrombo-embolic stroke. Int. J. Biometric Bioinform. (IJBB) (2009).
- [13] 7 Techniques to Handle Imbalanced Data – Kdnuggets.
- [14] Documentation for Logistic Regression from Scikit-learn.org.
- [15] Documentation for Decision Tree Classification from Scikit-learn.org.
- [16] Documentation for Random Forest Classification from Scikit-learn.org.
- [17] Documentation for K-Nearest Neighbor from Scikit-learn.org.
- [18] Documentation for Support Vector Machine from Scikit-learn.org.
- [19] Documentation for Naive Bayes Classification Algorithm from Scikit-learn.org.

A Novel Threshold based Method for Vessel Intensity Detection and Extraction from Retinal Images

Farha Fatina Wahid¹, Sugandhi K²

Department of Information Technology, Kannur University, Kannur, Kerala, India

Raju G³

Department of Computer Science and Engineering
Faculty of Engineering, Christ (Deemed to be University)
Bengaluru, Karnataka

Biswaranjan Acharya⁵

School of Computer Engineering
KIIT Deemed to be University
Bhubaneswar, Odisha, India

Debabrata Swain⁴

Department of Computer Science and Engineering
School of Technology, Pandit Deendayal Energy University
Gandhinagar, Gujarat, India

Manas Ranjan Pradhan⁶

School of Information Technology
Skyline University College
Sharjah, UAE

Abstract—Retinal vessel segmentation is an active research area in medical image processing. Several research outcomes on retinal vessel segmentation have emerged in recent years. Each method has its own pros and cons, either in the vessel detection stage or in its extraction. Based on a detailed empirical investigation, a novel retinal vessel extraction architecture is proposed, which makes use of a couple of existing algorithms. In the proposed algorithm, vessel detection is carried out using a cumulative distribution function-based thresholding scheme. The resultant vessel intensities are extracted based on the hysteresis thresholding scheme. Experiments are carried out with retinal images from DRIVE and STARE databases. The results in terms of Sensitivity, Specificity, and Accuracy are compared with five standard methods. The proposed method outperforms all methods in terms of Sensitivity and Accuracy for the DRIVE data set, whereas for STARE, the performance is comparable with the best method.

Keywords—Retinal images; blood vessel detection; and segmentation; segmentation; hysteresis thresholding; cumulative distribution function introduction

I. INTRODUCTION

Retinal images are the fundamental diagnostic element to identify various ophthalmic diseases for medical experts. It can be captured using a specialized camera, fundus camera, with low power light capable of simultaneously illuminating and capturing the interior part of the eye [1]. The image captured by a fundus camera includes the retina along with the optic disc, macula, fovea, posterior pole, etc. [2].

A critical component of the retina is the blood vessels which are broadly categorized to retinal arteries and veins. The retinal artery carries oxygen and nutrients. It is through the optic nerve that the retinal artery enters the eye and then splits to superior and inferior branches [3]. The abnormalities in the structure of blood vessels lead to several ophthalmic diseases. In general, the abnormalities are due to the variations that occur in blood

vessel attributes such as its tortuosity, vessel width, diameter, bifurcation, etc. These abnormalities are identified by the ophthalmologist using a detailed analysis of vessel color, shape, and contrast [2], [4].

When Intra Ocular Pressure (IOP) is higher than normal, it leads to ocular hypertension. As a consequence of this situation, the retinal artery narrows considerably and may cause the bursting of blood vessels [5]. When a blockage occurs to the arteries carrying oxygen to the nerve cells, it leads to retinal artery occlusion, which may cause vision loss [6]. Changes in the structure of blood vessels may lead to bleeding or leakage of fluid and vision distortion. This situation, diabetic retinopathy, arises in the retinal vessels due to diabetes [7]. The obstruction of blood flow occurs when blockages are present in arteries or veins of the retina and lead to the narrowing of blood vessels or blood clots. This is normally known as eye stroke in ophthalmology [8].

In the early days, ophthalmologists faced difficulties in diagnosing these ophthalmic diseases as manual extraction may contain a human error. Also, the entire process of vessel extraction is time-consuming. Hence, blood vessel segmentation using medical image processing techniques has become significant for an ophthalmologist for accurate diagnostics.

Several blood-vessel segmentation algorithms for retinal images have been developed in recent years [2], [9]–[12]. The segmentation algorithms can be broadly categorized into algorithms which comes under pattern recognition techniques [10], [12], vessel tracking approaches [13], [14], model-based approaches [15], [16], hardware implementation [17], [18] and hybrid approaches [19], [20]. Further sub-classification is possible for pattern recognition (PR) techniques and model-based methods. The model-based methods can be categorized into region-based models and edge-based models [2]. On the other hand, PR techniques are divided into supervised and unsupervised algorithms based on the classification

methodology [2]. In supervised learning, prior model has to be trained using a set of reference images whereas in unsupervised learning, there is no need of prior training of samples. In the last decade, researchers proposed several deep learning models, specifically convolutional Neural Network models, for the segmentation of blood vessels. Compared to classical machine learning approaches, these methods are resource demanding.

A comprehensive review of the topic reveals the limitations of existing models in terms of recognition accuracy and algorithm complexity. This motivated the authors to explore the area. In this paper, a novel retinal vessel segmentation framework based on unsupervised learning is proposed to accurately extract blood vessels from fundus images. Vessel extraction is achieved by combining different segmentation steps from well-established unsupervised algorithms. The vessel detection step is constructed from an unsupervised coarse-to-fine vessel segmentation method developed by Câmara Neto et al. using an adaptive threshold [20]. The threshold is based on the cumulative distribution function of vessel intensity values [10]. The vessel extraction is carried out using the hysteresis thresholding technique [11]. The proposed vessel segmentation framework is compared with a set of state-of-the-art methods and is found to have the edge over them.

The paper is articulated as follows. Section 2 gives a comprehensive review of blood vessel segmentation algorithms developed in recent years. Section 3 describes the proposed methodology followed by experimental results discussed in section 4. Section 5 concludes the paper.

II. LITERATURE REVIEW

In this section, descriptions of prominent retinal vessel segmentation algorithms using conventional and Convolutional Neural Network (CNN) based approaches are discussed, highlighting their methodologies and merits.

Coye developed a novel vessel segmentation algorithm for retinal images in 2015 [21]. RGB retinal images are converted into LAB color space and enhanced using the CLAHE algorithm in the proposed algorithm. Background exclusion is carried out using an average filter followed by thresholding using the IsoData algorithm on the improved image. Further, isolated pixels are removed using morphological operations.

A retinal vessel segmentation algorithm from color fundus images using an ensemble classifier is developed by Zhu et al. in 2016 [22]. They have used a 36-dimensional feature vector for each pixel. These feature vectors are then given as input to the Classifier and Regression Tree (CART) classifier. Ultimately, an AdaBoost classifier is constructed by iterative training for vessel extraction. The authors have experimented with their algorithm on the Digital Retinal Images for Vessel Extraction (DRIVE) database, and average values of accuracy, sensitivity, and specificity of 0.9535, 0.8319, and 0.9607, respectively, are reported [22].

Orlando et al., in 2017, developed a segmentation algorithm for retinal blood vessels via a discriminatively trained fully connected conditional random field model. A structured output support vector machine is used to automatically learn the parameters. The authors carried out their experiments on images from DRIVE, Structured Analysis of Retina (STARE), Child

Heart and Health Study in England (CHASE_DB1), and High-Resolution Fundus (HRF) databases and claimed that the proposed method outperforms other existing methods based on sensitivity, F1-score, G-mean, and Matthews correlation coefficient [12].

A simple retinal vessel segmentation algorithm was developed by Dash et al. (2017) [23]. The extracted green channel from the retinal image is initially enhanced using the CLAHE algorithm in this algorithm. The enhanced image is then segmented using mean-c thresholding. Morphological cleaning is carried out on the segmented image to obtain the final segmented blood vessels. The authors carried out their experiments on DRIVE as well as CHASE_DB1 databases and claimed 0.9555 and 0.9540 accuracy, respectively.

Jiang et al. (2017) proposed a retinal vessel segmentation algorithm that mainly comprised of capillary detection and venules detection [9]. In this algorithm, the input retinal image is converted to gray scale by extracting its green channel. A morphological top-hat operation is performed on the extracted image to obtain characteristic features for the classification of vessels and non-vessels, followed by empirical intensity thresholding. This leads to venules detection. For capillary detection, centerline candidate pixels are highlighted using the first-order derivative of the Gaussian filter by rotating it at different angles with a variation of 180. In their work, instead of rotating the filter, the image itself is rotated in order to avoid the loss of image information. After applying the filter, a connectivity check is performed to remove low connective components. Finally, the extracted venules and capillaries are combined and denoised using morphological erosion. The authors obtained an accuracy of 95.88% for single database tests and 95.27% for cross-database tests on experiments carried out with images from DRIVE and STARE databases.

A coarse-to-fine algorithm for vessel segmentation was developed by Câmara Neto et al. (2017) [10]. In this method, a vessel segmentation algorithm is proposed in which coarse vessels are detected using intensity thresholding, and vessel refinement is carried out using the principal curvature method. In the coarse vessel segmentation stage, the preprocessed retinal image is divided into two categories- tissue intensities and vessel intensities using an intensity threshold which is obtained by maximizing the distance between the extreme intensity values of an image and its cumulative distribution function (CDF). Further, an adaptive threshold is used for the final segmentation of vessels which are refined using curvature analysis and vessel reconstruction. Balanced accuracies of 0.7819 and 0.8702 are reported by the authors for images from DRIVE and STARE databases.

In 2018, Oliveira et al. proposed a retinal vessel segmentation algorithm by combining multi-scale analysis provided by stationary wavelets and multi-scale fully convolutional neural network [24]. This method used rotation operation for data augmentation as well as prediction. The authors carried out their experiments on DRIVE, STARE, and CHASE_DB1 databases and reported accuracies of 0.9576, 0.9694, and 0.9653, respectively.

Dash and Bhoi (2018) [25] reported a study on thresholding, founded on the variational minimax optimization

algorithm, to segment retinal blood vessels. In this algorithm, gamma correction is applied on the extracted green channel of the fundus image, which is then enhanced using the CLAHE algorithm. On this enhanced image, thresholding is carried out and followed by morphological cleaning to obtain the final segmented blood vessels. The authors claimed that the experiments carried out on DRIVE and CHASE_DB1 databases give an average accuracy of 0.957 and 0.952, respectively.

In 2018, Wang et al. proposed a vessel segmentation algorithm that used Hessian-based linear feature filtering at preprocessing stage followed by fuzzy entropic thresholding. The authors reported an F1 score of 66.15% for their proposed method [26].

A vessel segmentation algorithm that used Gray-level Hit or Miss Transform (GHMT) with multi-scale, iteratively rotated multi-structuring elements to extract blood vessels from retinal images was proposed by Pal et al. in (2019) [11]. The vessels extracted using GHMT are further post-processed using hysteresis thresholding. The authors claimed that their method gives an average accuracy of 94.31% for images from the DRIVE database.

Preity and Jayanthi introduced Multi-threshold along with morphological operations for vessel segmentation [27]. Preprocessing was carried out using AHE, CLAHE, and average filter, followed by segmentation using a variant of the Otsu method followed by post-processing using morphological operations. Average accuracy of 95.3% was obtained using the DRIVE database.

An adaptive segmentation algorithm was used to extract retinal blood vessels by Kabir (2020) [28]. Anisotropic diffusion was applied on grayscale fundus images, followed by top-hat transformation to enhance the image. Local-property-based intensity transformation was introduced on sub-images, and finally, vessels were segmented using k-means clustering from all sub-images. 95.29% and 95.47% accuracies were obtained on DRIVE and STARE databases, respectively.

Jadoon et al. (2020) proposed a vessel extraction algorithm in which the CLAHE algorithm is applied on the green channel of fundus images, followed by the top-hat operation [29]. Prominent vessels are obtained using Frangi as well as high boost filters. The entire vessels were obtained using Isodata thresholding. 95.32% and 94.98% accuracies on DRIVE and STARE databases are reported.

Alhussein et al. (2020) used hessian along with intensity transformation information for blood vessel segmentation from fundus images. On CLAHE enhanced image, Wiener and morphological filter were applied to denoise it [30]. Thick and thin vessel enhanced images were obtained separately by using Eigenvalues of the hessian matrix at two different scales. Otsu and Isodata thresholding were applied to extract thick and thin vessels from the respective enhanced images. Accuracies of 95.59% and 95.01% on DRIVE and CHASE_DB1 databases were respectively reported.

In [31], a multi-scale convolutional neural network for retinal vessel segmentation was proposed. The network consisted of two consecutive convolution structures. After each convolution layer, ReLU was used as the activation function.

Once features are extracted from convolutional layers, feature maps were added and connected with a fully connected layer. The proposed network consisted of three fully connected layers. Softmax function was used for classification and the loss function used was cross-entropy. The network was implemented using Anaconda and TensorFlow. The training process took about 18h to complete. The proposed methods were evaluated on DRIVE and STARE databases with sensitivity, specificity, accuracy, and AUC values of 0.843, 0.980, 0.951, 0.979 on DRIVE and 0.823, 0.978, 0.956, and 0.974 on STARE databases, respectively.

In [32], a novel multi-label classification scheme for retinal vessel segmentation by a local de-regression model (LODESS) designed for multi-labeling was proposed. CNN classifier was designed for multi-label classification and learning of multi-label neighborhood relations. The proposed CNN model consisted of an encoder with max-pooling layers and a decoder with upsampling and convolutional layers. A focal loss function was used to solve the imbalanced classification problem. The algorithm was built on the Keras library with NVIDIA Titan X Pascal GPU. The network was trained for 20 epochs, and a standard stochastic gradient descent optimizer was used for training. Experiments were conducted on DRIVE and STARE databases. Accuracy, sensitivity, specificity and F1 scores of 0.952, 0.776, 0.979 and 0.813 on DRIVE and 0.970, 0.812, 0.990 and 0.855 on STARE databases were reported.

In [33], retinal vessel segmentation using a cross-connected convolutional neural network (CcNet) was proposed. Features from different layers were fused using a cross-connected structure. There were primary and secondary paths in the network, with each convolutional layer in the primary path connected with all layers in the secondary path. Each cross-connection from the primary to the secondary path was processed by a convolution, ReLU, and max-pooling module (CRM) in order to reduce the network parameters. The outputs of the primary path concatenated with each output of the secondary path. ReLU is the activation function used for the last layer. Green channel images were given as input to the network, and a pre-training step was used to accelerate the convergence of the network. The parameters of these pre-trained modes were employed as the initial values of the final trained model. CcNet was trained using the ADAM optimization method. GeForce GTX 1070 was used to accelerate the computation using the Caffe toolkit. Based on the experiments carried out on DRIVE and STARE databases, sensitivities of 0.7625 and 0.7709, specificities of 0.9809 and 0.9848, accuracies of 0.953 and 0.963, and AUC values of 0.968 and 0.970 were obtained, respectively.

Retinal vessel segmentation by making use of deep supervision with historically nested edge detector (HED) and smoothness regularization from CRFs was proposed in [34]. The network, known as Deeply Supervised and Smoothly Regularized Network (DSSRN), was an end-to-end network for a pixel to pixel segmentation that combined FCN and CRF strengths. Five staged HED network was used, and the segmented results from each layer and the last fuse layer were connected to mean-field CRF. The learning rate was initially set to 10^{-8} , momentum and weight decay parameters were set to 0.9 and 0.0002, respectively. The model was implemented on

NVIDIA GTX Titan X GPU using the Caffe framework. Training model on a single GPU took 12h for completion, whereas testing of one retinal image required only 0.3s. Accuracy and sensitivity values of 0.954, 0.763 on DRIVE, 0.960, 0.742 on STARE and 0.959, 0.782 on CHASE_DB1 databases were obtained respectively.

Based on the review, it is clear that even though various algorithms exist, the majority are computationally expensive. Moreover, the results obtained can be further enhanced. So, there is a requirement for a better method with maximum true positive values and minimal false negative values. This motivated the development of a new unsupervised framework for retinal vessel segmentation. The following section describes the proposal.

III. PROPOSED METHODOLOGY

A novel blood vessel segmentation framework that combines a coarse-to-fine vessel extraction algorithm with hysteresis thresholding is proposed in this section.

Retinal vessel extraction, in general, is carried out in multiple steps. For thresholding-based methods, the last steps include thresholding for classification of pixels and post-processing for fine-tuning/denoising. We have carried out an extensive experimental study of chosen retinal vessel segmentation algorithms [10], [11], [21], [23]. Based on the study, the proposed model is built.

The separation of vessel intensities from the tissue intensities as proposed in [10] is found to give a relatively good result. Also, the thresholding technique used by Pal et al. in their work [11] is simple, efficient, and convincing. The use of hysteresis thresholding ensures that final vessel segmentation is not based on a single threshold. Based on these observations, a novel retinal blood vessel segmentation framework is proposed by obtaining vessel intensities as described in [10], followed by the application of hysteresis thresholding [11].

The working of the proposed algorithm is as follows. Let, I denote the green channel of the input retinal image. In order to make the vessels in I brighter than their background non-vessels, I is inverted.

$$I_c = (L - 1) - I \quad (1)$$

where L is the number of possible intensity values.

I_c is then smoothed using a Gaussian filter to remove noise and further enhanced using morphological top-hat operation.

$$I_E = I_g - (I_g \circ S_e) \quad (2)$$

where I_g denotes the Gaussian smoothed image, the symbol \circ denotes morphological opening operation, and S_e , its structuring element, respectively.

The enhanced image, I_E , further undergoes contrast enhancement such that the top and bottom 1% of all the pixel values in I_E are saturated [10]. From this contrast-enhanced image, I_{CE} , vessels are extracted using a threshold defined based on the cumulative distribution function (CDF) of pixel intensities in I_{CE} . The idea behind the selection of CDF for threshold computation is that vessels constitute approximately

12% of total intensities in an image and the CDF becomes smoother when vessel pixels are reached [10].

$$I_{CDF}(x) = \sum_{n=0}^x P_n \quad (3)$$

where $x \in [0, (L-1)]$ and

$$P_n = \frac{\text{no. of pixels with intensity } n \text{ in } I_{CE}}{\text{total number of pixels in } I_{CE}} \quad (4)$$

where $n = 0, 1, 2, \dots, (L - 1)$.

From the CDF, the threshold is computed as follows.

$$\tau = \operatorname{argmax}_x \left\{ \frac{|I_{CDF}(x) + \Delta I_{CDF} * (1-x) - 1|}{\sqrt{1 + \Delta I_{CDF}^2}} \right\} \quad (5)$$

where

$$\Delta I_{CDF} = I_{CDF}(L - 1) - I_{CDF}(0) \quad (6)$$

The vessel intensity image is then extracted based on the threshold, τ , as follows:

$$I_{CE} \in V \text{ if } (I_{CE} > \tau) \quad (7)$$

From this vessel intensity image, vessels are segmented using hysteresis thresholding. Unlike most of the thresholding schemes, which uses a single global threshold or local threshold, hysteresis thresholding uses two threshold values, say T_1 and T_2 . All the pixels in V whose intensity values are greater than or equal to T_1 have been considered as vessel pixels, and all the pixels whose intensity values are less than T_2 , considered as non-vessel pixels. Also, for each pixel whose intensity values lies in between T_1 and T_2 , it is considered as a vessel pixel if at least one of the pixel in its 8-neighborhood has an intensity value greater than or equal to T_1 [11]. The working of hysteresis thresholding for vessel segmentation is mathematically expressed as follows.

$$S_v(x, y) \begin{cases} 1, V(x, y) \geq T_1 \\ 0, V(x, y) \leq T_2 \\ 1, (T_2 < v(x, y) < T_1) \text{ and} \\ \quad \exists i |\phi_i(x, y) \geq T_1 \end{cases} \quad (8)$$

where $\phi(x, y)$ denotes the 8-neighborhood of pixel with coordinate positions (x, y) . Fig. 1 depicts the working principle of the proposed methodology. The algorithm for the proposed framework is as follows.

Proposed Algorithm

- Read input fundus image.
 - Extract the green channel and invert it.
 - Smooth the inverted image using a Gaussian filter.
 - Enhance the smoothed image using morphological top-hat operation.
 - Apply contrast enhancement.
 - Detect retinal blood vessels from the contrast-enhanced image
 - Threshold-based on CDF of pixel intensities are used.
 - Final retinal vessels are extracted by applying Hysteresis thresholding.
-

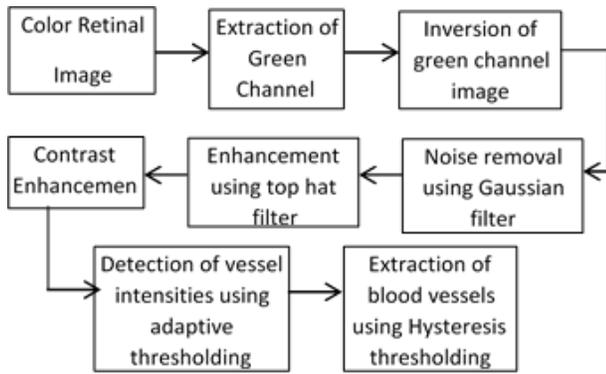


Fig. 1. Working Principle of the Proposed Methodology.

IV. EXPERIMENTAL SETUP

In this section, details of the data sets used, values of various parameters, and matrices for performance evaluation are described.

A. Data Sets

1) *Digital Retinal Image for Vessel Extraction (DRIVE)*: DRIVE is a well-known publicly available fundus image database used for comparative studies on retinal blood vessel segmentation algorithms [35]. The images in the database are obtained from a Diabetic Retinopathy (DR) screening program in the Netherlands with a population that comprised 400 subjects within the age group 25-90 years. From the photographs of these subjects, 40 images are randomly selected, among which 7 showed signs of mild early diabetic retinopathy, and the remaining had no signs of DR. Canon CR5 3CCD camera with a 45° field of view (FOV) is used for capturing the images. Each image in the database has a size of 584x568. The database is divided into two sets – training and testing set. Both the sets contain 20 images each, along with a corresponding mask that describes the FOV. Vasculature’s single manual segmentation is available for the training set, whereas the testing set contains two manual segmentations – one is used as gold standard/ ground truth and the other for comparison of an algorithm’s performance with that of a human observer.

2) *Structured Analysis of Retina (STARE)*: The STARE database is based on the STARE project initiated by Michael Goldbaum, M.D. in 1975, at the University of California, San Diego, funded by the U.S. National Institutes of Health [36]. The project is contributed by more than 30 people from the field of medicine, science, and engineering. The dataset consists of 400 raw images of size 700x605. The images are captured using the TopCon TRV-50 fundus camera at 35° FOV. For each image in the database, a list of diagnoses and diagnosis codes are also provided. Expert annotations of features visible in the database are also presented in a tabular format. For the blood vessel segmentation task, 20 images with manually segmented vasculature from two experts are provided. The database also contains artery/vein labeling of 10 images by two experts and 80 images with ground truth for optic nerve head detection.

B. Values for Various Parameters

The first parameter to be fixed is the sigma value for the Gaussian filter, used to remove noise from the inverted green channel image. The sigma value is fixed to 0.45, and a filter of size 3x3 is used as suggested in [10]. As far as the enhancement process is considered, the structuring element selected for the top hat filter is a disk with a radius of 6 [10]. For hysteresis thresholding, different combinations of threshold values are tried, and an optimum threshold value is fixed as $T_1 = 0.7$ and $T_2 = 0.2$ respectively. If the value of T_1 is reduced, it may lead to an increase in false-positive values. On the other hand, if the value of T_2 is increased, it may lead to decrease in true positive values.

C. Performance Evaluation Measures

“Sensitivity,” “specificity,” and “accuracy” are three performance evaluation measures considered for the study. These are powerful measures generally used to evaluate segmentation results. Based on the manually segmented vessels and the one obtained by applying a segmentation scheme, the numbers of pixels classified as true positive (TP), true negative (TN), false positive (FP), and false-negative (FN) are computed. From these values, the following formulas are derived [37].

$$Sensi = \frac{TPOS}{TPOS + FNEG} \quad (9)$$

$$Speci = \frac{TNEG}{TNEG + FPOS} \quad (10)$$

$$Accur = \frac{TPOS + TNEG}{TPOS + TNEG + FPOS + FNEG} \quad (11)$$

V. RESULTS AND DISCUSSION

Twenty images each from the DRIVE and STARE databases are selected for the experiments. The output of blood vessel segmentation is analyzed quantitatively using the performance evaluation measures – sensitivity, specificity, and accuracy.

In order to assess the merit of the proposed work, it is compared with the following algorithms: (i) Coye (2015): enhancement using CLAHE algorithm followed by segmentation using IsoData thresholding technique [21]; (ii) Jiang et al. (2017). : segmentation is carried out in two steps – venules detection and capillary detection [9]; (iii) segmentation technique which uses conditional random field model for discriminative feature extraction developed by Orlando et al. [12]; (iv) segmentation using mean-c thresholding [24] and (v) segmentation via variational minimax optimization thresholding [25].

Among the algorithms considered for implementation, each has its enhancement technique which precedes segmentation. These algorithms are applied on the green channel of color (RGB) retinal image and use the CLAHE algorithm for enhancement. The above algorithms are unsupervised methods except the one proposed by Orlando et al. [12], which is a supervised method.

Table 1 gives the list of methods used for the experimentations. The average, best and worst performance of the segmentation algorithms given in Table 1, based on

sensitivity, specificity, and accuracy on images from the DRIVE database, is given in Table 2.

From Table 2, it is clear that based on both sensitivity and accuracy, the proposed method outperforms the other methods for images from the DRIVE database. Based on specificity, though the proposed framework performs well, the supervised method based on the conditional random field [12] gives the highest value. Table 2 reveals that when the best performance of each algorithm is considered, the proposed method dominates the existing methods based on accuracy. For specificity, it performs closer to the method giving the best result. Also, it is evident that the worst performance of the proposed algorithm is more comparable to its average performance in terms of performance evaluation measures. The worst performance evaluation highlights the fact that for images where other methods fail to give a satisfactory result, the proposed method gives better results, which is a significant achievement.

In Table 3, the average, best and worst performances of the segmentation algorithms for images from the STARE database are given.

Methods that perform satisfactorily with the DRIVE dataset are found to give a poor performance with the STARE dataset. This is basically due to the difference in the quality of images.

From Table 3, it is clear that Method 1, which is a supervised method, shows good performance in identifying blood vessels with accuracy (high sensitivity and accuracy). The proposed method is better in terms of specificity, and overall performance is very close to Method 1. The performance of the other four methods is relatively poor.

Analysis of best and worst performance highlights the fact that variation of the performance of the proposed method with respect to images is low, which is a preferred quality.

From Table 3, it is clear that Method 1, which is a supervised method, shows good performance in identifying blood vessels with accuracy (high sensitivity and accuracy). The proposed method is better in terms of specificity, and overall performance is very close to Method 1. The performance of the other four methods is relatively poor. Analysis of best and worst performance highlights the fact that variation of the performance of the proposed method with respect to images is low, which is a preferred quality.

Table 4 gives a comparison of the proposed method with other state-of-the-art retinal vessel segmentation algorithms in terms of sensitivity, specificity, and accuracy on images from the DRIVE database. From Table 4, it is evident that the proposed method produces better results compared to other state-of-the-art methods.

Fig. 2 and Fig. 3 show the segmented vessels obtained using each algorithm for sample images from DRIVE and STARE databases, respectively.

The proposed framework fuses the merits of different works and is tested based on images from standard test databases. A possible challenge that may affect the algorithm is the quality of real-time input fundus images. We can correlate image statistics with input parameters by mapping image statistics with the parameters in such cases. It requires extensive experimentation and theoretical modeling and is not considered in the present study.

TABLE I. LIST OF SEGMENTATION ALGORITHMS USED FOR EXPERIMENTATIONS

Sl. No.	Method Name	Method Abbreviation
1	Conditional Random Field Model Segmentation [12]	M1
2	Unsupervised method using variational minimax optimization thresholding algorithm [25]	M2
3	Segmentation using mean-c thresholding [23]	M3
4	Fast, accurate retinal vessel segmentation [9]	M4
5	Taylor Coye algorithm [21]	M5
6	Proposed Framework	M6

TABLE II. AVERAGE, BEST AND WORST PERFORMANCE OF SEGMENTATION ALGORITHMS ON IMAGES FROM DRIVE DATABASE

Sl. No.	Method	Sensitivity			Specificity			Accuracy		
		Aver	Best	Worst	Aver	Best	Worst	Aver	Best	Worst
1	M1	0.7341	0.8245	0.6301	0.9822	0.9899	0.9758	0.9579	0.9648	0.9517
2	M2	0.6968	0.8539	0.3984	0.9752	0.9867	0.9615	0.9472	0.9607	0.9038
3	M3	0.7609	0.8933	0.3368	0.9608	0.9831	0.9115	0.9513	0.9715	0.9067
4	M4	0.7167	0.8419	0.5188	0.9499	0.9617	0.9406	0.9366	0.9471	0.9289
5	M5	0.6975	0.9484	0.1695	0.9641	0.9823	0.9487	0.906	0.9704	0.6834
6	M6	0.7992	0.8850	0.6567	0.9734	0.9850	0.9648	0.9595	0.9794	0.9527

TABLE III. AVERAGE, BEST AND WORST PERFORMANCE OF SEGMENTATION ALGORITHMS ON IMAGES FROM STARE DATABASE

Sl. No.	Method	Sensitivity			Specificity			Accuracy		
		Aver	Best	Worst	Aver	Best	Worst	Aver	Best	Worst
1	M1	0.8809	0.9541	0.7653	0.9462	0.977	0.8981	0.9427	0.9699	0.9018
2	M2	0.7409	0.9679	0.3343	0.9395	0.9761	0.896	0.9219	0.9657	0.8843
3	M3	0.6622	0.9041	0.2255	0.9220	0.9763	0.8546	0.9083	0.9726	0.8382
4	M4	0.7203	0.9125	0.3564	0.9348	0.9755	0.8785	0.9225	0.95	0.8761
5	M5	0.3821	0.7869	0.1446	0.9568	0.9805	0.9183	0.8079	0.9422	0.6713
6	M6	0.8002	0.9195	0.5983	0.9539	0.9875	0.8929	0.9392	0.9626	0.8912

TABLE IV. COMPARISON WITH EXISITING METHODS BASED ON PERFORMANCE ON IMAGES FROM DRIVE DATABASE

Sl. No.	Author, Year	Sensitivity	Specificity	Accuracy
1.	Coye, 2015	0.698	0.964	0.906
2.	Orlando et. al., 2017	0.734	0.982	0.958
3.	J. Dash and N. Bhoi,2017	0.761	0.961	0.951
4.	Jiang et. al., 2017	0.717	0.950	0.937
5.	J. Dash and N. Bhoi, 2018	0.697	0.975	0.947
6.	Fan et. al, 2019	0.736	0.981	0.960
7.	Alhussein et. al., 2020	0.785	0.972	0.956
8.	Kabir, 2020	0.784	0.982	0.953
9.	Preity and Jayanthi, 2020	0.695	0.985	0.953
10.	Jadoon et al., 2020	0.598	0.988	0.953
11.	Proposed Method	0.799	0.973	0.960

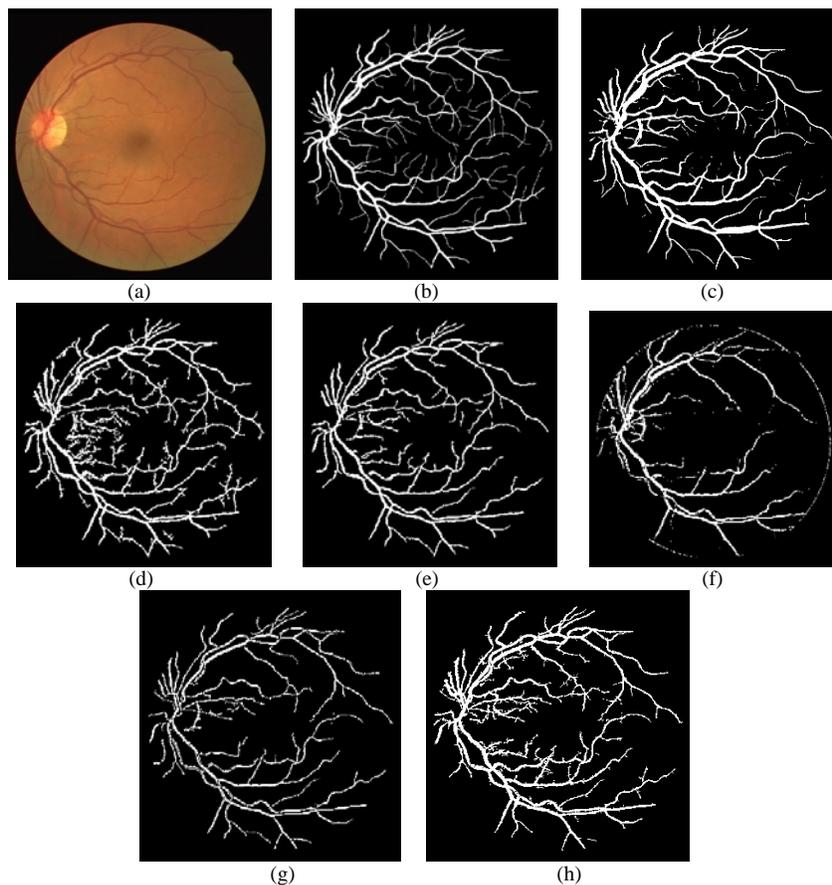


Fig. 2. Segmentation Results Obtained for Image 01_test from DRIVE database. (a) Original Fundus Image (b) Manual Segmented Vessels (c) M1[12] (d) M2[25] (e) M3[23] (f) M4 [9](g) M5 [21] (h) M6 (Proposed Method).

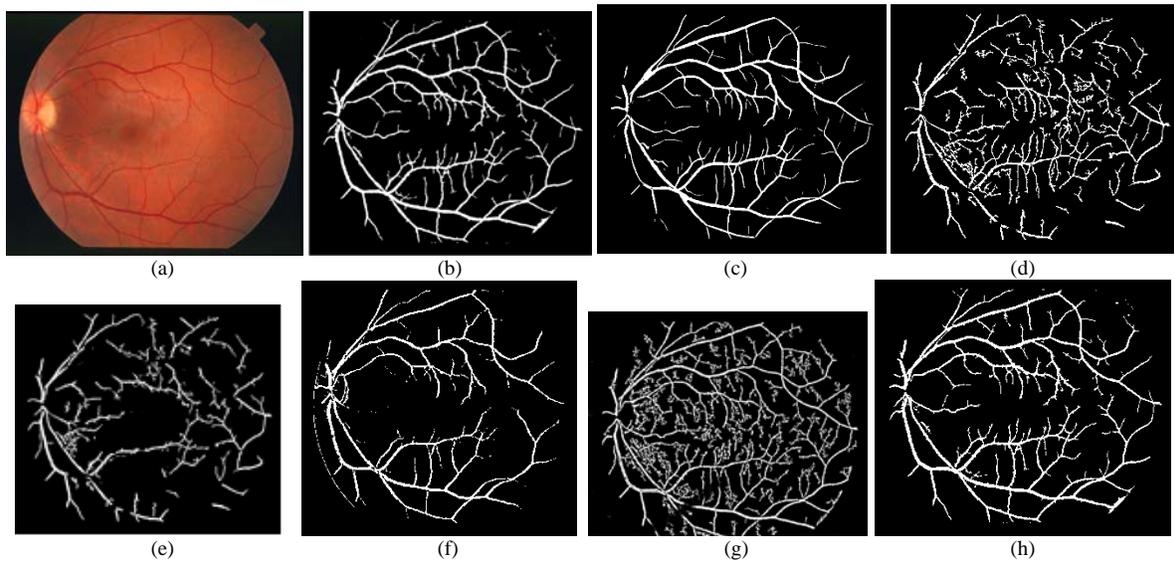


Fig. 3. Segmentation Results Obtained for Image im_0235 from STARE Database. (a) Original Fundus Image (b) Manual Segmented Vessels (c) M1[12] (d) M2[25] (e) M3[23] (f) M4 [9] (g) M5 [21] (h) M6 (Proposed Methodology).

VI. CONCLUSION

In this paper, a novel retinal vessel segmentation algorithm by performing vessel detection using an adaptive thresholding technique followed by hysteresis thresholding for final vessel segmentation is proposed. Vessel detection is carried out with an adaptive threshold that depends on the cumulative distribution function of image intensity values. This method is capable of detecting almost all vessel intensities within a retinal image. The detected vessel intensities are then extracted using hysteresis thresholding. The combination of the detection and extraction stage using the two approaches led to the development of an innovative idea for the extraction of retinal blood vessels.

Experiments carried out with the DRIVE and the state databases establish the merit of the proposed algorithm. The algorithm extends help to the ophthalmologist in an easy diagnosis of retinal diseases as the extracted vessels can identify changes in normal vessel structures and indicate the presence of abnormalities, if any. This opens scope for further research for a better segmentation algorithm. The present study can be further extended by incorporating the influence of image quality on the algorithm and making it suitable for all types of images.

REFERENCES

- [1] "Color Fundus Photography | Department of Ophthalmology." [Online]. Available: <https://ophthalmology.med.ubc.ca/patient-care/ophthalmic-photography/color-fundus-photography/>. [Accessed: 26-Nov-2018].
- [2] K. B. Khan et al., "A review of retinal blood vessels extraction techniques: challenges, taxonomy, and future trends," *Pattern Anal. Appl.*, pp. 1–36, 2018.
- [3] "About Retinal Blood Vessels | Retina Doctor Melbourne," 2019. [Online]. Available: <https://www.retina doctor.com.au/about-the-eye/the-retinal-blood-vessels/>. [Accessed: 16-Apr-2019].
- [4] M. M. Fraz et al., "Blood vessel segmentation methodologies in retinal images - A survey," *Comput. Methods Programs Biomed.*, 2012.
- [5] "Ocular Hypertension: Causes, Symptoms, Tests, and Treatment," 2019. [Online]. Available: <https://www.webmd.com/eye-health/ocular-hypertension>. [Accessed: 16-Apr-2019].
- [6] "Retinal Artery Occlusion - The American Society of Retina Specialists - The American Society of Retina Specialists," 2019. [Online]. Available: <https://www.asrs.org/patients/retinal-diseases/32/retinal-artery-occlusion>. [Accessed: 16-Apr-2019].
- [7] "Diabetic retinopathy - Symptoms and causes - Mayo Clinic," 2019. [Online]. Available: <https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/symptoms-causes/syc-20371611>. [Accessed: 16-Apr-2019].
- [8] "Eye stroke: Symptoms, risks, and treatment," 2019. [Online]. Available: <https://www.medicalnewstoday.com/articles/317877.php>. [Accessed: 16-Apr-2019].
- [9] Z. Jiang, J. Yepez, S. An, and S. Ko, "Fast, accurate and robust retinal vessel segmentation system," *Biocybern. Biomed. Eng.*, 2017.
- [10] L. Câmara Neto, G. L. B. Ramalho, J. F. S. Rocha Neto, R. M. S. Veras, and F. N. S. Medeiros, "An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images," *Expert Syst. Appl.*, 2017.
- [11] S. Pal, S. Chatterjee, D. Dey, and S. Munshi, "Morphological operations with iterative rotation of structuring elements for segmentation of retinal vessel structures," *Multidim Syst Sign Process*, vol. 30, no. 1, pp. 373–389, 2019.
- [12] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A Discriminatively Trained Fully Connected Conditional Random Field Model for Blood Vessel Segmentation in Fundus Images," *IEEE Trans. Biomed. Eng.*, 2017.
- [13] C. G. Owen et al., "Measuring retinal vessel tortuosity in 10-year-old children: Validation of the computer-assisted image analysis of the retina (caiar) program," *Investig. Ophthalmol. Vis. Sci.*, 2009.
- [14] C. M. Wilson et al., "Computerized analysis of retinal vessel width and tortuosity in premature infants," *Investig. Ophthalmol. Vis. Sci.*, 2008.
- [15] W. S. Oliveira, J. V. Teixeira, T. I. Ren, G. D. C. Cavalcanti, and J. Sijbers, "Unsupervised retinal vessel segmentation using combined filters," *PLoS One*, 2016.
- [16] A. E. Rad, M. S. Mohd Rahim, H. Kolivand, and I. Bin Mat Amin, "Morphological region-based initial contour algorithm for level set methods in image segmentation," *Multimed. Tools Appl.*, 2017.
- [17] F. Argüello, D. L. Vilariño, D. B. Heras, and A. Nieto, "GPU-based segmentation of retinal blood vessels," *J. Real-Time Image Process.*, 2014.
- [18] D. Koukounis, C. Ttofis, A. Papadopoulos, and T. Theodorides, "A high performance hardware architecture for portable, low-power retinal vessel segmentation," *Integr. VLSI J.*, 2014.

- [19] K. Jiang et al., "Isotropic Undecimated Wavelet Transform Fuzzy Algorithm for Retinal Blood Vessel Segmentation," *J. Med. Imaging Heal. Informatics*, 2015.
- [20] P. Dai et al., "A new approach to segment both main and peripheral retinal vessels based on gray-voting and Gaussian mixture model," *PLoS One*, 2015.
- [21] W. Zeng and C. Wang, "View-invariant gait recognition via deterministic learning," *Neurocomputing*, vol. 175, no. PartA, pp. 324–335, 2016.
- [22] C. Zhu, B. Zou, Y. Xiang, and J. Cui, "An Ensemble Retinal Vessel Segmentation Based on Supervised Learning in Fundus Images," *Chinese J. Electron.*, vol. 25, no. 3, pp. 503–511, 2016.
- [23] J. Dash and N. Bhoi, "A thresholding based technique to extract retinal blood vessels from fundus images," *Futur. Comput. Informatics J.*, vol. 2, no. 2, pp. 103–109, 2017.
- [24] A. Oliveira, S. Pereira, and C. A. Silva, "Retinal vessel segmentation based on Fully Convolutional Neural Networks," *Expert Syst. Appl.*, vol. 112, 2018.
- [25] J. Dash and N. Bhoi, "An Unsupervised Approach for Extraction of Blood Vessels from Fundus Images," *J. Digit. Imaging*, vol. 31, no. 6, pp. 857–868, 2018.
- [26] H. Wang, Y. Jiang, X. Jiang, J. Wu, and X. Yang, "Automatic vessel segmentation on fundus images using vessel filtering and fuzzy entropy," *Soft Comput.*, 2018.
- [27] Preity and N. Jayanthi, "A Segmentation Technique of Retinal Blood Vessels using Multi-Threshold and Morphological Operations," in *2020 International Conference on Computational Performance Evaluation, ComPE 2020*, 2020, pp. 447–452.
- [28] M. A. Kabir, "Retinal Blood Vessel Extraction Based on Adaptive Segmentation Algorithm," in *2020 IEEE Region 10 Symposium, TENSYPMP 2020*, 2020, pp. 1576–1579.
- [29] Z. Jadoon, S. Ahmad, M. A. Khan Jadoon, A. Imtiaz, N. Muhammad, and Z. Mahmood, "Retinal Blood Vessels Segmentation using ISODATA and High Boost Filter," in *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies: Idea to Innovation for Building the Knowledge Economy, iCoMET 2020*, 2020.
- [30] M. Alhussein, K. Aurangzeb, and S. I. Haider, "An Unsupervised Retinal Vessel Segmentation Using Hessian and Intensity Based Approach," *IEEE Access*, vol. 8, pp. 165056–165070, Sep. 2020.
- [31] M. Li, Q. Yin, and M. Lu, "Retinal Blood Vessel Segmentation Based on Multi-Scale Deep Learning," in *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2018, pp. 117–123.
- [32] Q. He, B. Zou, C. Zhu, X. Liu, H. Fu, and L. Wang, "Multi-Label Classification Scheme Based on Local Regression for Retinal Vessel Segmentation," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 2765–2769.
- [33] S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "CeNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features," *Neurocomputing*, 2019.
- [34] Y. Lin, H. Zhang, and G. Hu, "Automatic Retinal Vessel Segmentation via Deeply Supervised and Smoothly Regularized Network," *IEEE Access*, 2019.
- [35] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging*, 2004.
- [36] A. Hoover, "STRUCTURED ANALYSIS OF THE RETINA." [Online]. Available: <http://cecas.clemson.edu/~ahoover/stare/>.
- [37] "Sensitivity and Specificity." [Online]. Available: <http://www.med.emory.edu/EMAC/curriculum/diagnosis/sensand.htm>. [Accessed: 27-Nov-2018].

Integration of Identity Governance and Management Framework within Universities for Privileged Users

Shadma Parveen¹

School of Management and
Economics, University of Electronic
Science and Technology of China
Chengdu, 611731, China

Sultan Ahmad*²

Department of Computer Science,
College of Computer Engineering
and Sciences, Prince Sattam Bin
Abdulaziz University, Alkharj
11942, Saudi Arabia

Mohammad Ahmar Khan³

College of Commerce and Business
Administration, Dhofar University
Oman

Abstract—The development of high-tech progression around the world, with the attentions of governance body and private companies towards well-organized setup access control measures within the organization. The importance of these exceedingly essential perception, this article is proposing an integrated approach with the assimilation of IAM (identity access management) as an authentication tool and PAM (privileged access management) as a restricting accessing control measure in terms of an active directory. Originally the experimental setup organized within the Prince Sattam Bin Abdulaziz University, and it is analyzed using the real-time data which is available within the university database. We found that the proposed mechanism can be a vital method for protecting governance data or key business-oriented data from the unauthorized or adversarial attack. We also reviewed and compared other access control methods and find that the integrated method is relatively have an advantage to deal accessing task in any premier organization.

Keywords—Identity management; governance framework; privileged access management; security; privacy

I. INTRODUCTION

These days, propagation of the communication within the organization or outside world is very common. It is also very necessity to engage proper protocol to address effective and innovative things to manage secure access control over the system. The covid-19 disease outbreak has led to an inevitable rise with the use of digital technology nationwide subject to socioeconomic distancing requirements including territorial lockdowns, individuals and organizations now have to adjust a changing method of working as well as residing[1]–[5]. A rise in digitalization is leading industries as well as educational institutions to switch to work online. Governance and the management of identity will become relevant. They will require design and regulatory analysis. Online workers will likely increase in size, rising job distribution issues, teamwork, motivation, work overload, and present aspects. With rapid intensification in digital occurrence, tracking among the workplace and technostress problem will turn out to be prevalent[6]–[8]. Information technology (IT) improves companies' ability to compete in the 21st century's intensely competitive global marketplace has become increasingly evident. However, the actual practice of information technology count on profoundly on efficient and adequate IT governance. The flow of communication within the

organization and outside is inviting many security and privacy challenges among the user. The restriction of the adversarial attack is highly required and build such type of mechanism which can deal to provide a framework and use proper access among the authorized user[9]–[11]. To ensuring such type of mechanism our intension to encourage to develop an essential tool or any type of integration approach, concerning these technological enhancements, this article is going to optimized with the integration of the identity access management framework (IAM) and privileged access management (PAM) as an active directory which is demonstrated on the Fig 1. The architecture of the system incorporates as a managerial way such as life cycle manager, compliance manager, password manager and the layout of the system demonstrated such as identity intelligence, dashboards, reporting console, analysis console. This system is performed governance console including policy model, identity warehouse, role model, workflow engine, risk model, IT security, 3rd party provisioning, IT services management and mobile device management with the fronted communication in terms of external world.

Integrating as an active directory is necessary for almost all applications and contexts to organizations which can adapt within and across applications, systems, and boundaries. For instance, active directory can apply. One of the principal problems to be tackled is uncertainty. If an integrated framework is restricted or expensive, the development of trustworthy environments through adequate security and privacy policies and practices, a user-friendly interface, and commitment to user education and knowledge is another significant challenge for its successful implementation. Due to the organization's rapid growth and the application that they are using, the user will have to memorize all his credentials for each application[12]. Moreover, a crucial component of such an operation is digital identity management with privileged access management. Today public and private sector organizations vary dramatically in their approach to active directory designing their means of generating, checking, sorting, and using digital identities through their network and the internet. In our increasingly interconnected economies, the lack of shared policies and methods creates privacy, protection, and efficiency problems and hampers an organization's ability to deliver convenient service to customers. Besides, improving user conveniences is one of the

*Corresponding Author

significant advantages of active directory. Effective integration may reduce the inconveniences and inefficiencies caused by the need to keep track of various identities, passwords, and authentication requirements when used through organizations. Similarly, more functional user interfaces will boost accessibility and increase online services for registration and log-in processes. Finally, protection and privacy enhancement by minimizing the flow of data during transactions, only requesting, transferring, and storing what is needed, security and privacy are improved, and effective integrated approach can minimize the transactional details needed by multiple device users and reduce the security and privacy risks.

The significant contribution of the proposed study is provided a hybrid approach between the identity and access management and privileged access management to protect highly confidential access data within the organization or outside world. The incorporating this mechanism is validated that using these types of hybrid mechanism can be a better option to handle things and restrict some adversarial attacks

into the entire system. The adjustment between IAM and PAM, by describing a general project management methodology for IAM environments, we attempt at shutting down this research gap. It not only provides a generic high ranking research method, but also incorporates specific methods for the handling of attributes of quality. This fulfills the demand for the measurement of ABAC's quality evaluation and at the same time provides an interconnected process-oriented approach which is applicable to large-scale IAM scenarios.

The organization of the articles is as follows: Section 2 highlights literature background of the hybrid approach as a well-defined context about access control, Section 3 proposed an integrated methodology which provides a highly trusted architecture to deal with control access about the user within the organizations, Section 4 discussing optimization of proposed mechanism with advantages as well as disadvantages and finally Section 5 concludes with future work.

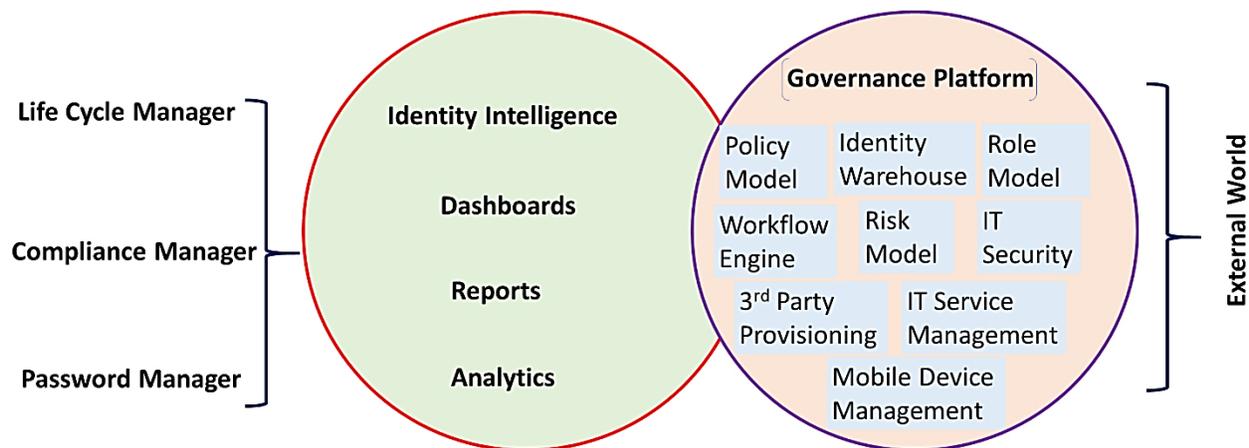


Fig. 1. Integrated Conceptual Design.

II. PRELIMINARIES

The intention of the governance and organizational setup is to improve control access methodologies within the institution from the external world. This requires due to day-by-day new threads coming to unauthorized access within the organization and stealing valuable secret data to destroy business as well as economic consequences around the world. Concerning these thread and new epidemical policies during covid-19, its hand to mouth to control adversarial data theft about the organization. So, it is required to provide an architecture to address these shortcomings, the adaptation of the identity and access management with integration of privileged access management is quite favorable option to adjust governance as well as business companies to fight back and control these valuable data and any miss happening things within the organization or institution. Identity management (IdM) is in a position to perform such functions as management, maintenance, discovery, management, enforcement of policies, exchange of information and substantiation. Identity and access management (IAM) formalizes the use and management of the same identity for all application areas and

guarantees simultaneous security[13]. The software is being utilised to validate devices, users or services as well as deny or allow access to data or other infrastructure components. In the event that the application is accessed, the system or service does not have to authenticate all its identity shops or authentication mechanism. Rather, the identity checking process can be customized such reliable identity provider, which actually decreases the volume of work among application. Management of identity and access simplifies management of transmitted distributed projects. IAM are utilised in whole business (B to B) or even amongst a private innovativeness with cloud provider within a company or outside the enterprise[14]. IAM has a broader organizational area to identify entities, cloud objects and manage admittance to resources oriented predetermined policies [15], [16]. Identity and access management are a number of functioning zones. Identity managing and provision, federated identity management, validation administration, compliance management and authorization management [17] are all functioning zones. These functioning zones ensure the secure and effective integration of authorized users at clouds. The service provisioning markup language (SPML) is an XML-

related framework which is used to manage identity. It enables communication on resources, users and services amongst officialdoms. One of SPML's failings, this employs multiple symmetric cryptography from innumerable providers which lead to a range of marginal applications (APIs). Since APIs not the identical provider, interacting with each other is difficult. Authentication management is the second industries application of IAM. This ensures secure management of credentials including passwords as well as digital documentations [18].

Federated identity management is the third operational field of IAM. The said identity management validates cloud facilities using the selected ID provider of the organization. Management of federated identities guarantees privacy, integrity as well as non-repudiation. It therefore ensures the confidence here between mobile application as well as the identity provider by transferring certified public keys from public key infrastructure (PKI). The fourth main activity is the management of authorizations. After authentication process, authorization management decides whether such an authenticated entity is permitted to perform an application. Compliance management is the last functioning zone of IAM. This ensures that the resources of an association are secured and accessible under the current policies and regulations [19]. In the field of cloud security, Identity management plays an important role. Privacy as well as compatibility are the major problems with current solutions to identity management, particularly in public cloud environments [10], [11], [18]. Currently, IAM systems are effective mechanisms for reducing cloud-based risks. Many institutions provide an IAM system to ensure information through the influence of each user's access authorization [20]–[22]. SailPoint, IBM, Oracle, RSA as well as Core security are the most popular IAM system providers. SailPoint's identity management approach is capable of handling passwords, compliance control, data access management, requests for access, automated delivery, as well as Single Sign-On [9]. Purpose of effective in web access requesting, providers to ensure, single sign-on company, multi-factor authentications, access control, privileged identity, and compliance of user-activity are available through the IBM identity and access management Suite. Oracle identity & access management offer four keys cloud security mechanisms. [23] explained an identity management system, and its functions, including security roles to prevent various scams and instances of identity theft. A distinction is presented between the conventional and new scope of identity management. As a prerequisite for the current IT period, an identity management framework is given. [24] reported that most emphasis on identity security has centered on securing customer information in the area of virtualization. However, as they use cloud services, users usually leave a trail, as well as the subsequent verification of accounts can potentially contribute to the leakage of confidential personal information. In the meantime, cyber criminals can do damage to cloud service providers by the use of fake identities. The author is introducing a credibility framework and developing a prestige identity management model for information computing to address these issues. Throughout the model, anonymous sources are created based on a credibility identity to ensure that aliases are untraceable,

as well as a framework is suggested to measure user credibility, that enables cloud providers recognize malicious users. Investigation verifies that perhaps the template will guarantee that data centers are accessed anonymously by consumers and also that cloud providers efficiently determine the reputation of users thereby infringing consumer privacy. [25] addressed and offers the efficient assessment model of IT administration that could be used in specific, by organizational management throughout the HEI. In addition, this research identifies the variables that lead to successful IT accomplish a specific objective upon that domain's published studies. The suggested model needs to be evaluated in future projects by someone else by empirical research.

Privileged access management offers an automatic credential managing as well as sessions managing solution for secure accessing control, audit, notification with logging for all privileged accounts. The methods are intended to control localized or domains shared admin privileges; a client's account and admin account, a customer's service, network device, operating system, application (A2DB) accounts and database. IT businesses can reduce privileged risks and meet assessment of progress by increasing the authority and ownership over authorized credentials. The benefit of approach, although, requires on all such situations of usage and even on the presence of resources on site, virtualized or cloud environments. Ecosystems also require to take account of advance reliability, breaking glass, disaster recoveries, as well as stint to recover when the application itself faults or underlying equipment part from connections to web access which could trigger a breakdown happens. Privileged information security used as a SaaS platform (software as a service) can only run in the clouds or allow special supervision nodes to drive as well as combine policies as well as events. These systems are fully maintained by PAM manufacturer as well as share data centers at multi-tenant installation with other PAM customers. Although nearby few PAM implementations at cloud utilizing SaaS at the moment, the trend implies that companies acquire assurance in the storage of PAM cloud credentials, administration tools and policies. This approach is determined by personal suppliers and system integrators, who offer economic services related on financial PAM capabilities as well as with no required skills by consumers.[26] explained in his article, amongst the most complex issues for device integration of 4G generation communication. Which is very complex in terms of safe multimedia delivery in current and future networks. This incorporation implies that in order to deliver their facilities to users, multiple service and network providers would have to collaborate. Such multi-domain setting poses a major threat to the consumer that has an agreement with the consumer and even just a restricted numeral of carriers and service providers confidence. [27] suggested about the model of digital identities. Confidence management should create and check the confidence to provide lead to increased results for online transactions such as reactants between the purchase and sale of products and services, improved retention and loyalty, expanded credibility, etc. Author discusses the need and value of confidence managing in the digital world, along with the different models and strategies required to mitigate confidence. For the existing options, a qualitative analysis was

presented. In view of the large-scale emergence of the digital world and electronic firms, the research seems to have a lot of significance. [28] Characterizes a case of developing a stable identity management system as well as its organizational structure in accordance with the VAHTI protection guidelines of the Finnish government. The construction project was about to be conducted in compliance with the directives for government protection, while adopting the management structure of the supplier itself. [29] referred the goal of his article is to examine how or when integrated knowledge management of a seaport terminal operation can be efficiently shared between the most advantageous and cost-effective great skills, to potentially boost their access control, a road haulier as well as a rail operator. The automated exchanging of characteristics depends on the standardization of the participating actors among information technology. In this analysis, compatibility is accomplished through an existing simple object access protocol. His research paper adds to previous studies by creating a cost comparison that identifies the characteristics affecting four principles (from low advantage/low cost to high advantage/high cost) data cultivation, committed sharing of data, opportunistic exchange of information, as well as preventing sharing of information.

III. MATERIAL AND PROPOSED METHOD

The idea behind to integration or hybrid approach of the IAM and PAM are to facilitate secure ecosystem about the end user within the organization or outside the organization. Nowadays, there are so many malicious, adversarial threads available in the world to make highly risk and easily destroy organization with stealing credential, informative data. The key objective of this project is to clarify that identity management and governance are central to good cybersecurity, and role-based access control is among the essential characteristics of identity and access management (RBAC). Role-based access control enables device users to delegate positions[30]. Moreover, permissions are required to execute specific functions across these roles. This implies that users are not explicitly granted authorizations but rather obtain them through their allocated profession function or responsibilities, meaning whether someone enters the business, switches offices, goes on leave, or leaves the company, their access rights are easy to handle and stay in charge. Instead of addressing user access rights at a granulated level, operator access rights are combined into several positions across different systems[31]. This means that you mechanically have one set (combination) of defined access rights if your effort in the Finance team, which is different from if you work in the marketing team. Organizations minimize both the difficulty of granting user access rights and the related costs by role-based access management. It offers the ability to evaluate access rights in order to ensure compliance with different legislation, as well as to refine processes such that new workers can be up as well as successively work from the beginning, as it is already defined that the new member of staff would access systems, all based on his or her position in the company. The market advantages are various. This also increases productivity, in addition to the apparent improvement in security across the company, resulting in smoother onboarding and off-boarding processes

and enforcement, as an organization has a higher degree of control and understanding of who has access to what and why, as well as decreasing directorial work, IT care and creating price savings.

A. Integration Methods

In numerous systems of government, digital identity management initiatives and processes have been developed in order to deal with identity-related risks, compliance, and operational gaps. In order to stay one step ahead of the competition, companies must regularly evaluate their identity solutions' abilities. An increase in the number of password-related breaches has made the issue of IAM access rights even more critical. To safeguard data from internal and external threats, the several organizations still have yet to implement mature capabilities that enable them to effectively manage privileged access, even though the frequency of the compromise of privileged accounts has increased. Compounding the risk of compromise has been added to the equation. This type of organization has invested in a product, but it hasn't implemented the processes as well as governance to make it work. Other organizations may have well-established processes, but they are missing the necessary supporting technologies that would be required for addressing privileged access threats at an enterprise scale. Identity governance and PAM solutions have been implemented by some organizations, but few have combined the two. For many companies, this could lead to the inconsistent application of processes and policies across silos of tools, which would result in incorrect reporting and failed audits.

In order to better manage both privileged as well as nonprivileged user access requests, authorizations, accreditations, provisioning, and restoration, organizations should implement Identity Governance and PAM. The importance of the proposed approach as shown in the Fig. 2 using cloud identity platform's component is that it makes it possible to integrate with and govern a wide variety of enterprise applications as well as directories, whether they are in the cloud, on-premise, or a combination of both. Creating a service account is typically required, which must be set up for each application to gain access to identity information.

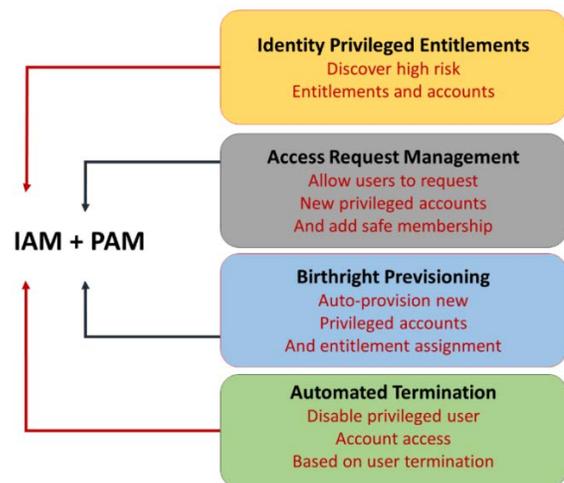


Fig. 2. Integrated Approach.

B. Session Flow within the Optimised System

The proposed incorporation of the hybrid approach is significantly organized as shown in the Fig. 3. Where it is described the conceptual session flow within the organization. The implemented mechanism is targeting to achieve better connection-oriented way within the organization to tackle users accessing policies and its managing session to accomplishing the task. In the session flow operation, some sort of the protocol handler is incorporated to setup its goal and provide accessing things according to user’s authorization skills. In which the used protocol handlers are listed as follows:

1) *Secure connection:* Communication lines prior among each node are optimised securely. This can be validated in the Fig. 3 where it is demonstrated very well how the transformation happen within the entire system. Our optimised system has taken care secure communication and tracing each activity with the monitoring capacity easily.

2) *Launch application locally:* Initially the optimised setup launched with locally installed server to administering each activity. The incorporated adjustment was quite well to reach desired outcomes.

3) *Session recording:* The additional process is inbuilt in the optimised setup which is known as session recording concept. This feature is enabling to monitoring task within the university setup and this can provide real proof about the adversarial attacks or miss use within the organization very well. Utilising this mechanism gives freedom to know who want to access our setup then it can be possible to take immediate action against the happened activities.

4) *Terminating session:* The administration of the entire setup can easily terminate suspected activity within the university accessing lines. The incorporation the whole process it can be easily seen that each process is providing distinguished outcomes to handle security a privacy within the system to integrate IAM and PAM.

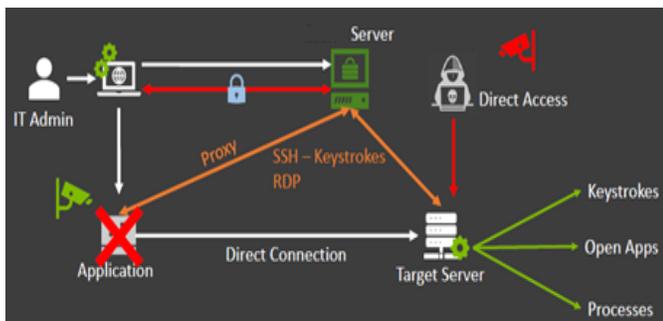


Fig. 3. Session Flow Diagram within the Applied Mechanism.

C. Implementation with Hardware and Software

The implementation of the integrated approach is needed appropriate hardware and supporting software utilities to successfully run proposed thought. The proposed integrated method is needed highly equipped hardware resources which are illustrated in Table 1 very well. The required supporting software is also defined very well in the Table 2. The

collectively incorporation among the hardware and software utilities produced an integrated mechanism which can be easily used among the organizations as well as any universities for the security and safety from the external world or adversarial attacks.

TABLE I. HARDWARE SPECIFICATION FOR THE IMPLEMENT

VM Name	CPUs >=2 GHz	RAM (GB) >=4 GB	Storage Space (GB) >=100 GB	OS Advance Version	Comments
Server Node 1	8	16	500	Windows Server 12 or Higher	
Server Node 2	8	16	500	Windows Server 12 or Higher	
Database Cluster	8	16	1000		It entails if customer does not have present DB
Site Connector	2	4	100	Windows Server 12 or Higher	

TABLE II. COMMUNICATING PORT FOR THE IMPLEMENTATION USING SOFTWARE

Application/Process	Traffic Types	Port	Sources	Destination
Discovery	Microsoft DS SSH RPC Port Range Epmap	445 22 49152- 65535 135	Secret Server IP	Work Stations
Web Server	HTTP HTTPs	80 443	IT Admin IP	Secret Server
Active Directory Syn	NTLM LDAP LDAPS Kerberos	445 389 636 88	Secret Server IP	Active Directory
Remote Changer password	Oracle Telnet NTLM LDAPS MS SQL SSH LDAP Sybase Kerberos	1521 23 445 636 1436 22 389 5000 88	Secret Server IP	Work Stations
Database	TCP/UDP SQL Connections	1433	Secret Server IP	Database Cluster
Load Balancer	HTTP HTTPs	80 443	Load Balancer IP	Secret Server IP
RADIUS Server	RADIUS	1812	Secret Server IP	RADIUS Server
Email	SMTP	25	Secret Server IP	Email Server
Rabbit MQ	MQTT	5672	Secret Server IP	Rabbit MQTT

IV. DISCUSSION TOWARDS OPTIMIZATION

The integrated mechanism for the optimization towards security and safety from external world within the organization and universities is very necessary. Discussing an integrated method which is the highly acceptable due to achievement of the various factors including major access controls mechanism incorporated with hybrid approach of IAM and PAM. The incorporation of this method reflected as a better mechanism which gives three-way access control mechanism as shown on Fig. 4. The first one is endpoint privilege managing. This managing administration gives very least privileged as well as credential theft protection within the organization to remain safe and protected towards external things. The second one is core privileged access security managing. This managing administration is provided remotely vendor access, risk-based credential security, protection of session management attacks and protection of least privilege server with domain controller. The third one is application access managing. This managing administration is illustrated tools, secret managing applications, containers and other develops. These all protection is coming from the optimization of IAM and PAM in a certain degree of integration methods as mentioned in the proposed section. During incorporating and implementation it seems some advantages to explore which is quite justifiable to integrate these applications.

A. Benefit of the Integrated Methods

One integrated identity and access management (IAM) and privilege access management (PAM) implementation can resolve this issue and make it possible for businesses to reliably respond to incidents and help facilitate regulatory compliance. When used, it can be used to automate use cases that involve the management of privileged accounts in the real world, proposing a unified, policy-driven approach to IAM across all users.

- Discover privileged accounts in addition passwords installed by the IAM solution in the PAM program.
- Implement a single, set of policies IAM solution for all users.
- Supplying fresh privileged accounts automatically by role-based access provisioning or IAM application authorization policies.
- Utilize user profile characteristics including title, consulting firm including profile to allow sufficient access to privileged accounts.
- Automation of periodic accounts access checks.
- Automation and implementation of duties segregation (SOD) strategies on privileged and non-privileged accounts.
- Modularizing privileged accounts abortions based on player separation or termination incidents as per the active directory solution.
- Excludes the doubt as to who is permitted to receive privileged or restricted information.
- We've developed better security on both the outside and inside of the organization.
- The efficiencies created by autonomous systems reduce costs, freeing companies to focus on building and protecting their networks.
- To stop the breach from happening, implement a process to avoid hackers from breaking in will save both time and money.
- Enforcing new and existing security policies with the system for the easiness.

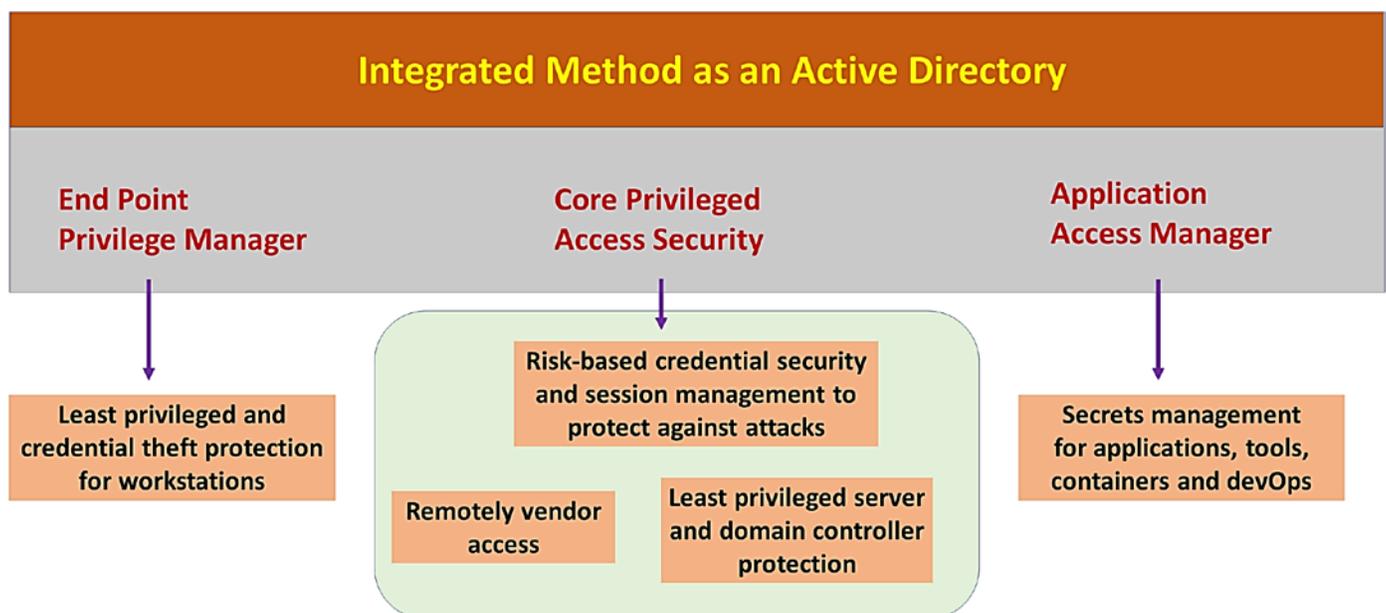


Fig. 4. Outcomes of Applied Integrated Mechanism.

B. Limitation of the Integrated Methods

Protecting identity and restricting access control is the key reason for using an integrated method. The system is controlled by only passwords. This is to say that the employee access code is only has some value when they have the ability to affect it. Regardless of the size of the enterprise, installing and configuring an integrated method can be expensive and time-consuming. For security purposes, it needs to be incorporated with the current security systems. To the extent possible, many enterprises are depending on IT security experts to develop and enforce the relatively better framework such as two factor or three factor authentications to eliminate interruptions for staff and company operations.

V. CONCLUSION

The emergence of the technological advancement around the world a well-organized governmental as well as business oriented smart control measure focuses within the organization. The emphasis of these highly required concept, this article is proposed an integrated approach with the incorporation of IAM as an authentication tool and PAM as a restricting accessing control measure. This method is enhanced the secure accessing policies among the governance as well as business organization to know each activity and record the adversarial attack from the proposed system. The whole experimental setup implemented within the Prince Sattam Bin Abdulaziz University, Saudi Arabia and it is analyzed using the real-time data which is available within the university database. We found that the proposed mechanism can be a vital method for protecting governance data or key business-oriented data from the unauthorized or adversarial attack which is always a challenging task in any premier organization. Using this integrated method, we believe that there will be so many controls within the governance body or business companies which can help a lot to restrict any miss happening being a large company. This article is just an initiative towards the better digitized system in terms of cyber security to employ at public as well as private corporate system to enrich trusted ecosystem for the people of kingdom of Saudi Arabia.

For the future work, we are still thinking to incorporate some cryptographic based strong authentication setup and integrate with effective privilege access control measure which can give highly trusted and relatively better outcomes for the governance system or business-oriented institutions.

ACKNOWLEDGMENT

We thank the Deanship of Scientific Research, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia for help and support.

REFERENCES

- [1] Y. Zhang, R. H. Deng, G. Han, and D. Zheng, "Secure smart health with privacy-aware aggregate authentication and access control in Internet of Things," *J. Netw. Comput. Appl.*, vol. 123, pp. 89–100, Dec. 2018.
- [2] L. Malina, P. Dzurenda, J. Hajny, and Z. Martinasek, "Secure and efficient two-factor zero-knowledge authentication solution for access control systems," *Comput. Secur.*, vol. 77, pp. 500–513, Aug. 2018.
- [3] C. Lin, D. He, X. Huang, K.-K. R. Choo, and A. V. Vasilakos, "BSEIn: A blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0," *J. Netw. Comput. Appl.*, vol. 116, pp. 42–52, Aug. 2018.
- [4] A. Sharma, S. B. Borah, and A. C. Moses, "Responses to COVID-19: The role of governance, healthcare infrastructure, and learning from past pandemics," *J. Bus. Res.*, vol. 122, pp. 597–607, Jan. 2021.
- [5] U. Iqbal and A. H. Mir, "Secure and scalable access control protocol for IoT environment," *Internet of Things*, vol. 12, p. 100291, Dec. 2020.
- [6] Y. Inoue, "Indirect innovation management by platform ecosystem governance and positioning: Toward collective ambidexterity in the ecosystems," *Technol. Forecast. Soc. Change*, vol. 166, p. 120652, May 2021.
- [7] M. Janssen, P. Brous, E. Estevez, L. S. Barbosa, and T. Janowski, "Data governance: Organizing data for trustworthy Artificial Intelligence," *Gov. Inf. Q.*, vol. 37, no. 3, p. 101493, Jul. 2020.
- [8] P. Ashley, M. Vandenwauver, and F. Siebenlist, "Applying authorization to intranets: architectures, issues and APIs," *Comput. Commun.*, vol. 23, no. 17, pp. 1613–1620, Nov. 2000.
- [9] J. Khan, H. Abbas, and J. Al-Muhtadi, "Survey on Mobile User's Data Privacy Threats and Defense Mechanisms," *Procedia Comput. Sci.*, vol. 56, pp. 376–383, 2015.
- [10] J. Khan et al., "Efficient secure surveillance on smart healthcare IoT system through cosine-transform encryption," *J. Intell. Fuzzy Syst.*, vol. 40, no. 1, pp. 1417–1442, Jan. 2021.
- [11] J. Khan et al., "SMISH: Secure Surveillance Mechanism on Smart Healthcare IoT System With Probabilistic Image Encryption," *IEEE Access*, vol. 8, pp. 15747–15767, 2020.
- [12] P. White, "Identity Management Architecture: A new direction," in *2008 8th IEEE International Conference on Computer and Information Technology*, 2008, pp. 408–413.
- [13] I. Indu, P. M. R. Anand, and V. Bhaskar, "Identity and access management in cloud environment: Mechanisms and challenges," *Eng. Sci. Technol. an Int. J.*, vol. 21, no. 4, pp. 574–588, 2018.
- [14] I. Indu and P. M. Rubesh Anand, "Identity and access management for cloud web services," in *2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, 2015, pp. 406–410.
- [15] P. M. Rubesh Anand and V. Bhaskar, "A unified trust management strategy for content sharing in Peer-to-Peer networks," *Appl. Math. Model.*, vol. 37, no. 4, pp. 1992–2007, Feb. 2013.
- [16] S. Parveen, S. Yunfei, J. P. Li, J. Khan, A. U. Haq, and S. Ruinan, "E-waste Generation and Awareness on Managing Disposal Practices at Delhi National Capital Region in India," in *2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing*, 2019, pp. 109–113.
- [17] Z. Wang, D. Huang, Y. Zhu, B. Li, and C.-J. Chung, "Efficient Attribute-Based Comparable Data Access Control," *IEEE Trans. Comput.*, vol. 64, no. 12, pp. 3430–3443, Dec. 2015.
- [18] Z. Yan, M. Wang, Y. Li, and A. V. Vasilakos, "Encrypted Data Management with Deduplication in Cloud Computing," *IEEE Cloud Comput.*, vol. 3, no. 2, pp. 28–35, Mar. 2016.
- [19] J. Li, X. Huang, J. Li, X. Chen, and Y. Xiang, "Securely Outsourcing Attribute-Based Encryption with Checkability," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 8, pp. 2201–2210, Aug. 2014.
- [20] J. Khan et al., "An Authentication Technique Based on Oauth 2.0 Protocol for Internet of Things (IoT) Network," in *2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2018, pp. 160–165.
- [21] J. Khan et al., "Medical Image Encryption Into Smart Healthcare IOT System," in *2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing*, 2019, pp. 378–382.
- [22] Q. Liu, G. Wang, and J. Wu, "Secure and privacy preserving keyword searching for cloud storage services," *J. Netw. Comput. Appl.*, vol. 35, no. 3, pp. 927–933, May 2012.
- [23] A. Bhardwaj and V. Kumar, "Identity Management Services in the Present IT Era," *IPASJ Int. J. Inf. Technol.*, vol. 6, no. 4, pp. 6–11, 2018.

- [24] L. Wu, S. Zhou, Z. Zhou, Z. Hong, and K. Huang, "A Reputation-Based Identity Management Model for Cloud Computing," *Math. Probl. Eng.*, vol. 2015, pp. 1–15, 2015.
- [25] M. Q. Huda, M. C. Utami, N. A. Hidayah, and Q. Aini, "Effective IT Governance in Higher Education Institutions: The Conceptual Model," vol. 149, no. Icosat 2017, pp. 148–151, 2018.
- [26] G. Karopoulos, G. Kambourakis, S. Gritzalis, and E. Konstantinou, "A framework for identity privacy in SIP," *J. Netw. Comput. Appl.*, vol. 33, no. 1, pp. 16–28, Jan. 2010.
- [27] P. Pradhan and V. Kumar, "Trust Management Models for Digital Identities," *Int. J. Virtual Communities Soc. Netw.*, vol. 8, no. 4, pp. 1–24, Oct. 2016.
- [28] K. Rindell, S. Hyrynsalmi, and V. Leppanen, "Case Study of Security Development in an Agile Environment: Building Identity Management for a Government Agency," in 2016 11th International Conference on Availability, Reliability and Security (ARES), 2016, pp. 556–563.
- [29] S. Jacobsson, P. O. Arnäs, and G. Stefansson, "Automatic information exchange between interoperable information systems: Potential improvement of access management in a seaport terminal," *Res. Transp. Bus. Manag.*, vol. 35, p. 100429, Jun. 2020.
- [30] M. umar Aftab, Z. Qin, Zakria, S. Ali, Pirah, and J. Khan, "The Evaluation and Comparative Analysis of Role Based Access Control and Attribute Based Access Control Model," in 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2018, pp. 35–39.
- [31] M. U. Aftab et al., "Negative Authorization by Implementing Negative Attributes in Attribute-Based Access Control Model for Internet of Medical Things," in 2019 15th International Conference on Semantics, Knowledge and Grids (SKG), 2019, pp. 167–174.

Dorsal Hand Vein Extraction in Uncontrolled Environment

Nisha Charaya¹, Anil Kumar², Priti Singh³

Department of ECE
Amity University, Gurgaon
Haryana, India

Abstract—Biometrics is an inseparable part of our day to day life. A major development in this area has been observed in past few decades. Over the recent years, dorsal hand veins have emerged as a promising biometric attribute due to its universality, stability and anti-forgery characteristics. However, detecting the veins of different thickness under different illumination is a challenging task. The traditional vein extraction approaches based on thresholding does not find their applicability in these situations. This paper presents a hybrid approach for vein segmentation for these hand images. The proposed approach is a combination of two techniques, i.e. repeated line tracking and maximum curvature points. The technique has been tested over Bosphorus hand vein dataset which contains 1575 images of different age groups captured under different illumination conditions. From the results, it is evident that this technique is suitable to extract vein pattern from all types of images. Further, these images have yielded an accuracy of more than 98% when subjected to feature extraction and classification steps.

Keywords—Biometrics; security; forgery; hand vein; vein segmentation; vein extraction; thick veins; unclear veins

I. INTRODUCTION

With the advances in technology and implementation of automation everywhere, security has been put at stake and needs to be ensured before bringing a new technology into practice. Starting from minor to major, all the activities in our lives has become automated with the help of biometrics. Attendance marking systems in offices, online transactions in banks and door opening at home; everything has become fast and automatic. But this automation requires a perfect, accurate and precise method to identify a person [1][2]. The technique to identify or verify a person with the help of physiological or behavioral attributes is known as biometrics. Iris, veins, fingerprints, and face are examples of physiological biometrics whereas voice, gait and DNA are examples of behavioral biometrics. A biometric system is user-convenient as compared to the traditional methods as it doesn't require be kept with safety or crammed all the times [3].

Hand-veins are seeking researcher's attention due to its attracting characteristics of user convenience and anti-forgery. Under the human skin, there is vein pattern which consists of various networks of blood vessels. The structure of vascular pattern in human body parts is diverse and stable over time [4]. Further it is present beneath skin which cannot be visible directly with bare eyes of humans. Related to other biometric

features, vein pattern are very difficult to recognize. Still its characteristics like uniqueness, contactless, universality and anti-forgery are strong enough to hide the challenges in its use. One of the main challenges in vein pattern based biometric system is to acquire images of vein pattern quickly without involving harmful and expensive devices which has been resolved by the near infra-red cameras.

Another major challenge is extracting the vein pattern from the images captured under poor illumination, with different hand postures and for people of different age group [5]. The thickness of hand veins vary with age and it gets affected due to some common diseases like diabetes and hypertension. In addition to these, irregular shading, optical blurring and skin scattering are some unavoidable factors which deteriorates the vein recognition and overall accuracy of the system. Though a lot of significant work has been contributed by researchers, yet an effective technique which can extract vein pattern from all types of images needs to be devised. This raises the need for finding a vein detection technique which is capable of extracting veins of varying thickness from the unclear images acquired from hands kept in different postures.

In this paper, a hybrid approach for the same is presented which is a merger of two techniques i.e. repeated line tracking method and maximum curvature point's method. The repeated line tracking method is based on tracing the dark lines by examining the cross section profile. It has proved to be effective for unclear images but faces issues for thick/thin veins. On the other hand, maximum curvature point method detects the vein based on the points having maximum curvature. This method is quite satisfactory for detecting the veins of different thickness but fails when the images are unclear. As the former one is suitable for unclear images while the latter is for thick/thin veins, so a combination of these two is proposed which has proved to be efficient for detecting veins from all types of hand images.

The main objective of this paper is to devise a vein extraction technique which can detect the veins of varying thickness from unclear images acquired in uncontrolled environment.

The rest of this paper is organized as follows. Section II discusses related work on this topic. The general methodology is described in Section III. Section IV depicts the proposed method and Section V contains simulation results. Finally, Section VI concludes the paper.

II. LITERATURE STUDY

The generalized and most basic way to detect vein pattern is by identifying the dark lines. It is done by thresholding method. A suitable threshold is selected and all the pixel values are compared with the threshold. Based on the comparison, it is categorized as background or vein. The most common thresholding method is OTSU's thresholding which has been adopted by many researchers to extract vein pattern [6]. However, the traditional method fails for images captured in uncontrolled environments in different hand postures.

Apart from thresholding, some different concept has also been devised to detect the vein patterns. The repeated line tracking method has been implemented by Miura (2004) to extract the finger veins which was able to extract veins in the presence of irregular shading [7], [8].

In 2007, Zhao applied local dynamic threshold method for vein segmentation which was based on extracting the veins by calculating mean and variance of the points in the neighborhood [9].

Zhong, Shao and Liu (2018) adopted the curvature point algorithm to extract the veins in an uncontrolled environment which was earlier deployed by Miura(2005) to extract finger vein pattern[10][11]. This algorithm is based on the concept that, in each grey profile of the grey image the distribution of grayscale values is a concave curve and the curvature value of point on the vein is greater than zero. So, the veins can be extracted by finding the pixels whose curvature values are greater than zero. The curvature point used deeper data features, which largely eliminated the effects of noise and retained more venous details [4].

Li (2010) used a local dynamic thresholding named NiBlack to segment the vein image[12]. To overcome the effect of uneven surfaces which leads to dark regions in DHV image, Shang-Jen Chuang (2017) applied global-local threshold algorithm [13].

BELEAN (2017) applied Hough transform (VP-HT) to determine the line segments in hand vein image. This technique is based on a global description of image features [14].

The majority of the methods for DHV recognition are devised on the basis of geometrical strategies and thresholding. However, optical blurring, hand vein postures and skin scattering are the major issues which may result into irregular shadings that may degrade the accuracy of recognition. A comparative analysis of existing techniques is depicted in Table 1 which clearly shows the lack of a technique that can extract vein pattern from thick/thin veins and unclear images acquired in uncontrolled environment.

All the techniques have their own significances but none of them can be used globally to extract vein pattern from all type of hand vein images due to their own applicability to particular images.

TABLE I. COMPARATIVE ANALYSIS OF EXISTING VEIN SEGMENTATION TECHNIQUES

Year	Authors	Technique Used	Remarks
2004	Miura , Nagasaka, and Miyatake	Repeated line tracking	Tested on finger veins, not suitable for thick veins
2007	Zhao et. al	Local dynamic thresholding	Not suitable for uncontrolled environment
2010	Li, Liu, and Liu	NiBlack thresholding	Not suitable for noisy images acquired in uncontrolled environment
2012	Li, Wang, and Jiang	OTSU thresholding	Not suitable for noisy images acquired in uncontrolled environment
2017	Shang-Jen Chuang	Local-global thresholding	Not suitable for noisy images acquired in uncontrolled environment
2017	Bellean et al.	Hough transform	Not tested for noisy and thick veins images
2018	Zhong, Shao and Liu	Curvature Point Algorithm	Not validated for non-uniform and unclear images

Extraction of vein pattern is an essential requirement for designing a hand vein based human authentication system. So, a poorly extracted pattern cannot result into an accurate biometric system. This raises the need for a global method of vein detection which can be used for all kind of hand images that are acquired with different postures, ages and illumination.

To achieve this, a new vein detection technique is proposed in this paper.

III. GENERAL METHODOLOGY

A general biometric system is comprised of several components as shown in fig 1. The sensor element collects the unprocessed biometric data from an individual that may be in the form of an image, video, audio or some other signal. This unprocessed data is passed through the pre-processing block where it is refined, improved and made ready to be processed and utilized [15]. The feature extraction block gathers a unique set of features to represent the signal. The extracted feature set is labeled with user identity and stored in the database during user enrolment phase. The matching block is responsible for performing comparison between the data presented for identification and the stored data. As an outcome of comparison, a matching score is generated. The decision block categorizes the presented data as identified or unidentified based on its matching score value.

Image pre-processing specifically vein extraction plays an important role in the overall system performance. The images captured in uncontrolled environment contain noise due to variation in illumination, optical blurring and bone shadows [4]. Also, it is challenging to extract veins from all types of images with different hand postures and varying vein thickness. So, it becomes mandatory to pre-process the images.

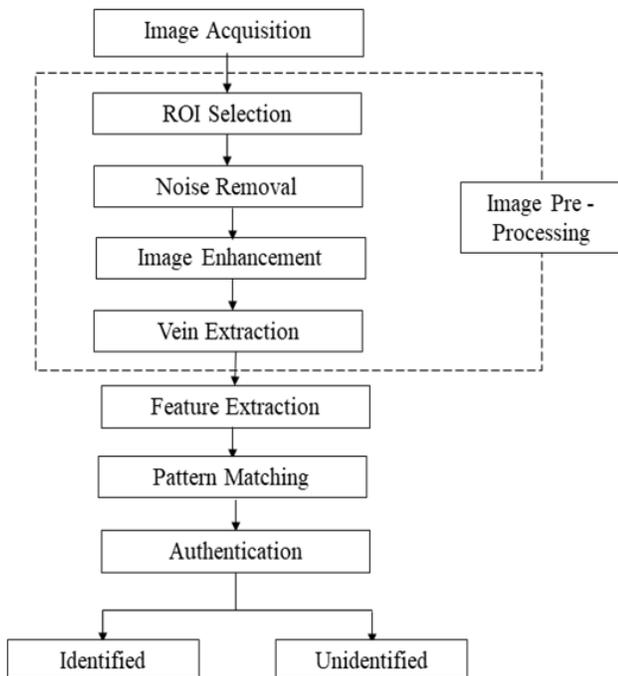


Fig. 1. Flow Chart of Hand vein based Authentication System.

For this, the images acquired by near infra-red camera are first converted to gray scale images. After this, histogram equalization is performed to make the image uniformly illuminated. Noise is removed by the application of median filter and the images are enhanced by morphological ‘open’ operations. Once enhanced hand image has been obtained, it is used to extract the region of interest (ROI) for further processing.

The ROI is a small area selected and extracted out of whole image for further processing. The size of image and the number of images in a database is very large which makes it time consuming to process the complete data as well as its memory requirements are larger. For these reasons, a particular small region is selected out of the whole image from which the features are extracted and stored for database creation and matching purpose. Generally, the ROI is selected around the centre of image so that it is present in all the images and is not affected by the position and placement of hand. In this work, ROI is extracted from the segmented hand image and is chosen such that the most affected and corrupted area is removed from the image and all the important information is retained. It comes out to be a rectangular or square region which is equidistant from all the margins in all directions. Once the region of interest is obtained, the proposed method is employed to extract the vein pattern from the selected regions.

IV. PROPOSED METHOD

Vein segmentation is an important and essential step in hand vein based authentication system as the system performance is directly dependent on it. The proposed technique of vein detection is a combination of two techniques i.e. repeated line tracking method and maximum curvature point method. The dorsal hand images are passed through

them one by one and the resultant images are obtained. Finally the resultant images obtained by the two methods are overlaid upon each other to get the vein pattern extracted.

A. Repeated Line Tracking Method

The repeated tracking of dark lines is performed to facilitate extraction of hand-vein patterns using non-uniform images. The hand veins are extracted from the non uniform images based on tracking the veins in all directions starting from a seed pixel. The line-tracking operation is initiated from any pixel in the captured image. The current position of a tracking pixel is called as the “current tracking point”, and it is moved along the dark line pixel by pixel. The vein appears as a valley in its cross sectional profile as shown in fig 2. The depth of valley is affected by the shading in the image but still it can be detected. It makes this method a robust method that can be used for extracting veins from non-uniform images. The depth of the cross-sectional profile around the current tracking point is checked to decide whether this point lies on a dark line or not. If this cross sectional profile resembles a valley, the tracking point is on a dark line. Further, the cross section profile is observed for different angles and the one with deepest profile is retained to provide the direction of tracking. After this, the current tracking point is updated with the closest pixel in that direction. If a valley is detected again at the updated tracking point, the tracking is continued in the direction of deepest valley and this result in extraction of a vein or part of a vein. Else if valley is not detected in any direction, this means that the updated pixel is not on a dark line and a new tracking operation is started at some other position. This tracking process is depicted in fig 3 showing current tracking point, direction and the cross section profile. A single line tracking operation will yield only a part of veins, so to gather and recognize all the hand veins, tracking is performed at various random starting positions across the whole image.

Sometimes, a tracking operation may also track a region of noise but there are more chances of tracking a dark line which represents a vein. A matrix named the “locus space” is created to store the number of times that each pixel has become the current tracking point. The size of the locus space is determined by the number of pixels in the captured images. Matrix elements are obtained by recording the total count in which the pixel has become the current tracking point against each pixel. Therefore, a frequently tracked element of the locus space has a higher value. The hand veins are obtained as chains of these high value positions in the locus space.

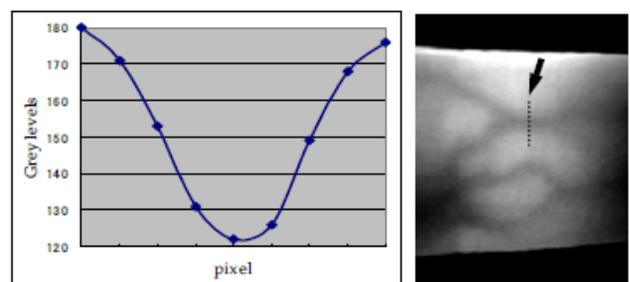


Fig. 2. Cross-sectional Profile and Pixel Position [9].

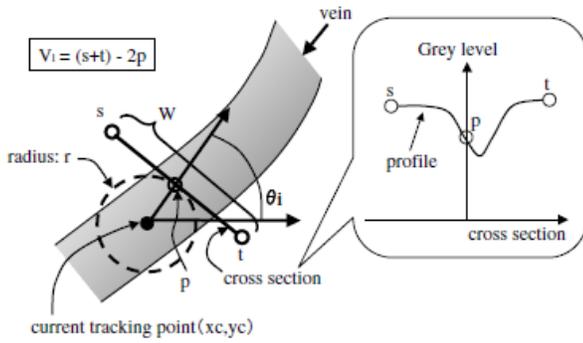


Fig. 3. Detection of Dark Line by Repeated Line Tracking [9].

The steps considered in recognizing the hand veins using repeated line tracking strategy is portrayed below:

Step 1: Discovering the starting point to move-direction attributes and start line tracking.

Step 2: Recognition of tracking direction by examining cross-sectional profile and updating tracking position along the dark line.

Step 3: Update count against a pixel in locus space whenever a pixel gets traced.

Step 4: Repeat step 1 to step 3.

Step 5: Acquiring the vein pattern via locus space.

This method has the advantage of quick processing and minimized computational time while extracting the vein pattern from unclear dorsal hand images with non-uniform illumination.

B. Maximum Curvature Point Method

The width of hand veins varies with several conditions like temperature, blood in the veins, physical conditions etc. Also with ageing, the vein thickness increases which produces difficult in tracking. To combat these issues, the maximum curvature points technique is employed to inspect the curvature of image and highlight center-lines of veins. The center-lines are determined by probing the position wherein the curvatures of a cross-sectional profile of a vein are locally maximal. These points are connected with each other to obtain the vein pattern.

The algorithm is processed with three steps which involve:

- Mining of centre positions of veins.
- Connecting centre positions.
- Image labelling.

The algorithm checks the cross-sectional profile of a hand vein image. As the vein is darker than its surroundings, the cross-sectional profile of a vein appears like a dent. It appears as a concave surface with large curvature. The curvature is large even for a thin vein. Therefore, even the thin veins get recognized by obtaining local maximum curvature points. Each local maximum position is assigned with a score which is larger for a deeper profile.

The curvature $C(z)$ of a cross sectional profile $P_H(z)$ can be represented as:

$$C(z) = \frac{d^2 P_H(z)/dz^2}{\left\{1 + \left(\frac{dP_H(z)}{dz}\right)^2\right\}^{3/2}} \quad (1)$$

Where, $P_H(z)$ is obtained from the intensity of a pixel, $H(x,y)$ in any direction and z represents a direction in profile. $P_H(z)$ is related to $H(x,y)$ by a mapping function T_{rs} [10].

The sign of $C(z)$ is observed to determine the nature of curvature. The curvature is concave for a positive sign and convex for a negative sign of $C(z)$. The points, where local maxima are obtained for a concave area, give the centre positions of the veins. All such points are stored as z'_k and a score is assigned to them. The assigned scores determine the probability that these points are on the veins. The score for any centre position is calculated as:

$$S(z'_k) = C(z'_k) \times W_r(k) \quad (2)$$

Where, $W_r(k)$ represents the width of concave region containing z'_k [10].

The centre positions of veins are joined together based on score values to obtain veins from group of points. The profiles are analyzed in all four directions to acquire the vein pattern. The vein pattern so obtained is binarized by thresholding method.

The hand veins recognized by the two methods are overlaid upon the original image and stored in the vector form represented as J . Each hand vein region is then subjected to feature extraction and classification for mining significant features and matching process respectively.

V. SIMULATION RESULTS

The proposed technique has been implemented in MATLAB and tested over Bosphorus Hand database.. The database comprises of 1575 images including both left and right hand images captured under different postures and 219 are captured with a time drop of several months.

The database comprises 642 subjects having 6 images for a person, which contains three right-hand images and three left-hand images. The images are of different illumination condition and different postures.

Moreover, 276 subjects with three left-hand images only. Amongst total 918 subjects, 160 poses hand images with time drop of several months.

The proposed method has been tested for vein detection in normal hand images, images with thick veins, rotated hand images and translated hand images. The simulation results for the same are shown in figure 4-6.

The resultant images obtained from maximum curvature point method shown in fig 4, contains dark patches due to irregular shading. While repeated line tracking fails to extract vein pattern from thick vein images as shown in fig 5. From the simulation results shown in fig 6, it is observed that the proposed method can successfully detect vein pattern from all kind of images which can be further used for feature

extraction and classification steps to develop the complete authentication system.

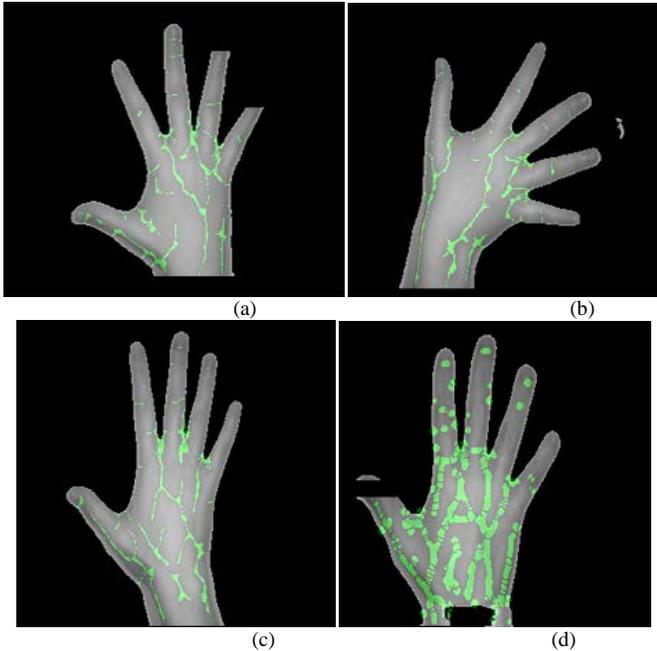


Fig. 4. Veins Extracted by Maximum Curvature Points Method for (a) Normal Image (b) Rotated Image (c) Translated Image (d) Thick Vein Image.

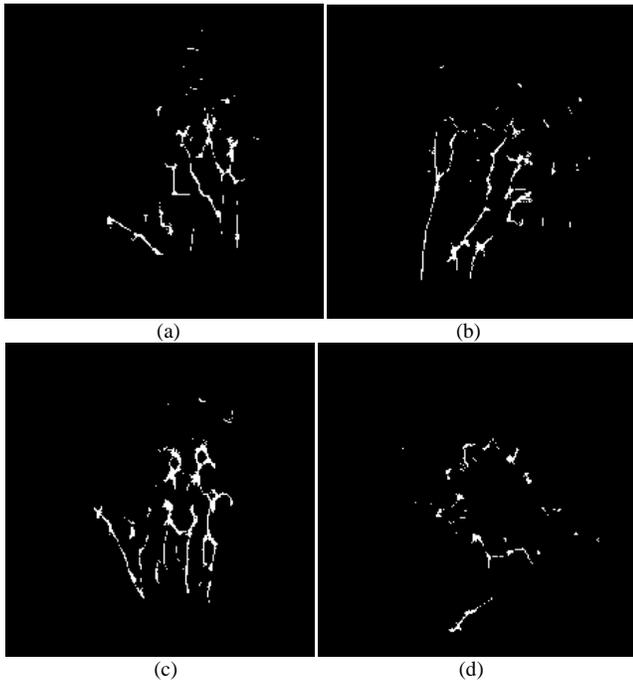


Fig. 5. Veins Extracted by Repeated Line Tracking Method for (a) Normal Image (b) Rotated Image (c) Translated Image (d) Thick Vein Image.

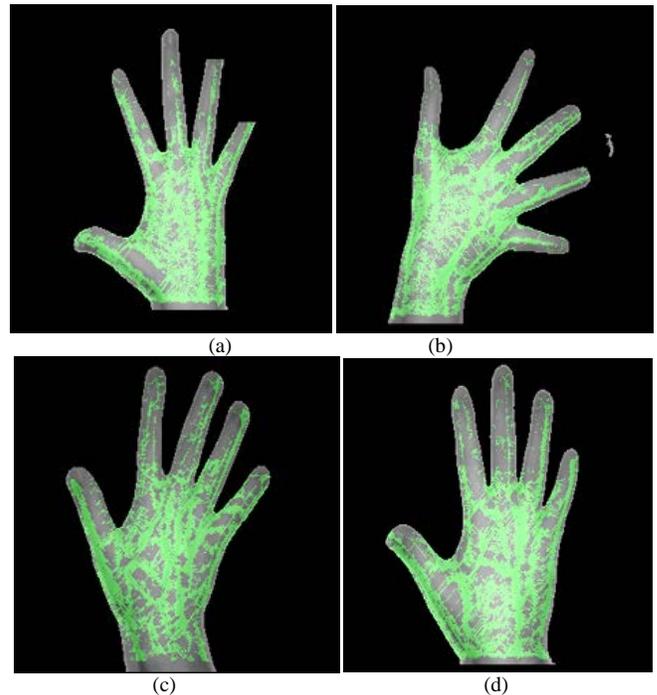


Fig. 6. Veins Extracted by Proposed Method for (a) Normal Image (b) Rotated Image (c) Translated Image (d) Thick Vein Image.

VI. CONCLUSION

It is evident from the results that the proposed method can efficiently detect the vein pattern from thick veins and unclear images acquired in uncontrolled environment. This method can be used as a global method of vein segmentation for all kind of images. The vein pattern so produced when further tested for matching purpose by using convolutional neural network (CNN) for feature extraction and deepCNN for classification yields overall accuracy of above 98% for all type of images. However, this approach needs to be validated over practical noisy images and can be extended to the design of dorsal hand vein based biometric system.

REFERENCES

- [1] K. Dharavath, F. A. Talukdar, and R. H. Laskar, "Study on biometric authentication systems, challenges and future trends: A review," 2013 IEEE Int. Conf. Comput. Intell. Comput. Res. IEEE ICCIC 2013, 2013.
- [2] N. Charaya, "Human Authentication Based On Dorsal Hand Veins : A Review," vol. 119, no. 16, pp. 2175–2185, 2018.
- [3] N.Charaya, "Biometric Systems and Attributes: A review", *Advancements & Modern Innovations in Engineering and Technology (AMIET 2020)*, pp. 250-253.
- [4] D. Zhong, H. Shao, and S. Liu, "Towards application of dorsal hand vein recognition under uncontrolled environment based on biometric graph matching," *IET Biometrics*, vol. 8, no. 2, pp. 159–167, 2019.
- [5] L. Wang and G. Leedham, "Near- and far- infrared imaging for vein pattern biometrics," *Proc. - IEEE Int. Conf. Video Signal Based Surveill. 2006, AVSS 2006*, pp. 4–9, 2006.

- [6] H. Li, Y. Wang, and X. Jiang, "Dorsal hand vein recognition method based on multi-bit planes optimization," vol. 1, Springer International Publishing, pp. 3–10, 2018.
- [7] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Mach. Vis. Appl.*, vol. 15, no. 4, pp. 194–203, Oct. 2004.
- [8] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger vein patterns based on iterative line tracking and its application to personal identification," *Syst. Comput. Japan*, vol. 35, no. 7, pp. 61–71, 2004.
- [9] S. Zhao, Y. Wang, and Y. Wang, "Extracting hand vein patterns from low-quality images: A new biometric technique using low-cost devices," *Proc. 4th Int. Conf. Image Graph. ICIG 2007*, pp. 667–671, 2007.
- [10] N. Miura, A. Nagasaka, and T. Miyatake, "Extraction of finger-vein patterns using maximum curvature points in image profiles," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 8, pp. 1185–1194, 2007.
- [11] "Miura et al. vein extraction methods-MATLAB Answers-MATLAB Central." [Online]. Available: <https://in.mathworks.com/matlabcentral/fileexchange/35716-miura-et-al-vein-extraction-methods>. [Accessed: 30-Jan-2021].
- [12] X. Li, X. Liu, and Z. Liu, "A dorsal hand vein pattern recognition algorithm," *Proc. - 2010 3rd Int. Congr. Image Signal Process. CISP 2010*, vol. 4, no. 2, pp. 1723–1726, 2010.
- [13] S. J. Chuang, "Vein recognition based on minutiae features in the dorsal venous network of the hand," *Signal, Image Video Process.*, vol. 12, no. 3, pp. 573–581, 2018.
- [14] B. Belean, M. Streza, S. Crisan, and S. Emerich, "Dorsal hand vein pattern analysis and Neural Networks for biometric authentication," *Stud. Informatics Control*, vol. 26, no. 3, pp. 305–314, 2017.
- [15] A. K. Jain, P. Flynn, and A. A. Ross, "Handbook of Biometrics Handbook of Biometrics," 2007.

An Ensemble GRU Approach for Wind Speed Forecasting with Data Augmentation

Anibal Flores¹, Hugo Tito-Chura²

Grupo de Investigación en Inteligencia Artificial
Universidad Nacional de Moquegua
Moquegua, Peru

Victor Yana-Mamani³

E.P. Ingeniería de Sistemas e Informática
Universidad Nacional de Moquegua
Moquegua, Peru

Abstract—This paper proposes an ensemble model for wind speed forecasting using the recurrent neural network known as Gated Recurrent Unit (GRU) and data augmentation. For the experimentation, a single wind speed time series is used, from which four augmented time series are generated, which serve to train four GRU sub-models respectively, the results of these sub-models are averaged to generate the results of the proposal ensemble model (E-GRU). The results achieved by E-GRU are compared with those of each sub-model, showing that E-GRU outperforms the sub-models. Likewise, the proposal model (E-GRU) is compared with benchmark models without data augmentation such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), showing that E-GRU is much more precise, reaching a difference of around 15% with respect to the Relative Root mean Square Error (RRMSE) and 11% with respect to the Mean Absolute Percentage Error (MAPE).

Keywords—Wind speed forecasting; recurrent neural networks; gated recurrent unit; ensemble GRU; data augmentation

I. INTRODUCTION

Earth's natural greenhouse effect makes life possible as we know [1]. However, human activities, such as the burning of fossil fuels and deforestation, have intensified the natural phenomenon, causing global warming [2], and due to this problem, the exploitation of renewable energies such as solar, wind, thermal energy and others have emerged as excellent alternatives for its solution.

Regarding wind energy, this is harnessed through the use of wind machines or wind motors capable of transforming wind energy into mechanical rotational energy usable for the production of electrical energy. Thus, the prediction of wind speed time series has become an essential task in wind energy farms, this helps in the planning of energy production [3] among others.

In models based on deep learning, the problem of overfitting [4], [5], [6] is usually presented due to the lack of data. Various solutions have been suggested in the literature, such as the use of dropout layers, regularization and data augmentation.

In this work an ensemble model for wind speed forecasting is proposed, it is based on the recurrent neural network known as Gated Recurrent Unit (GRU), where despite having enough historical data for the training phase [7], a data augmentation process is used with the sole objective of improving the precision of the model results, thus it is used the data augmentation technique proposed by Flores et al (2021) "in press" [8]. GRU is used instead of Long Short-Term Memory (LSTM), due to the antecedents such as [9], [10], and others where GRU presents slightly better results than LSTM.

The proposal ensemble model (E-GRU) consists of four GRU sub-models, for which four different augmented time series have been generated from a single wind speed time series. The final result is the average of the four sub-model predictions. The idea of using an ensemble model arises from the need to take advantage of the default and excess predicted values with respect to the observed or original data.

The main contribution of this study is a novel ensemble model (E-GRU) for wind speed time series forecasting based on recurrent neural networks as GRU and data augmentation.

The content of the work has been organized as follows. In the first section, the problem and the respective solution are described. In the second section, the theoretical bases are described, which are the basis of the paper's proposal. In the third section, the methodology followed for the implementation of the proposal is described. In the fourth section, the results achieved are described and discussed. In the last section, the conclusion reached at the end of the study is presented, as well as future work.

II. BACKGROUND

This section briefly describes some theoretical bases that are important for understanding the content of the paper.

A. Recurrent Neural Networks (RNN)

Just like Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), RNNs are part of the fundamental architectures of Deep Learning, which specialize in working with sequential data, hence their use in natural language processing (NLP) as well as in time series regression.

The best known RNN is probably Long Short-Term Memory (LSTM) known to overcome the vanishing gradient problem in RNNs. Several variants are generated from LSTM, including Gated Recurrent Unit (GRU), which, as mentioned above, for certain case studies, especially in time series, presents better results than LSTM.

The GRU architecture is shown in Fig. 1

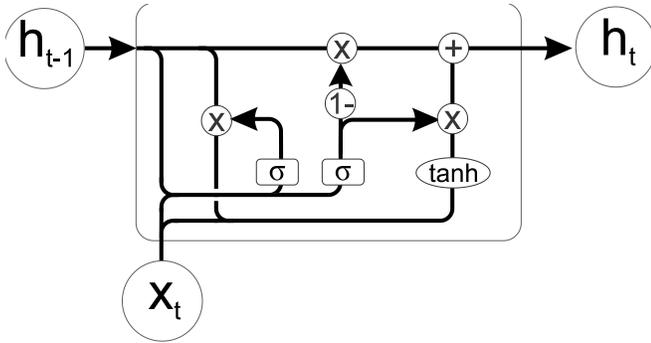


Fig. 1. GRU Architecture.

From Fig. 1, to estimate h_t it is necessary the following equations:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (1)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (2)$$

$$\tilde{h}_t = \tanh(W x_t + U(r_t \odot h_{t-1}) + b) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

Where:

W_z, W_r, W, U_z, U_r, U Matrices of parameters

b_r, b_z, b Vectors of parameters

σ Element-wise sigmoid function

\odot Element-wise multiplication

B. Data Augmentation

Data augmentation arose to solve overfitting problems in image classification [11] models like CNN and others. Many of these techniques consisted of zooming, rotation, flipping, etc. Later, the concept was transferred to time series classification, here techniques such as time-warping, rotation, scaling, jittering, etc. emerged.

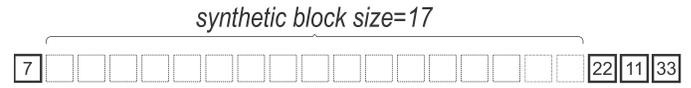
This work uses the technique proposed in "in press" [8] which is based on two basic techniques such as time-warping and jittering. The first one allows to increase the length of the original time series and the second one makes the synthetic data generated with the first one non-linear. Thus, this technique works with two parameters, the block size and the sub-block size, the first indicates the number of synthetic items to insert between each pair of the original time series and the second the number of linear synthetic items in each synthetic block. Fig. 2, shows a graphical view of this data augmentation technique.

Original time series

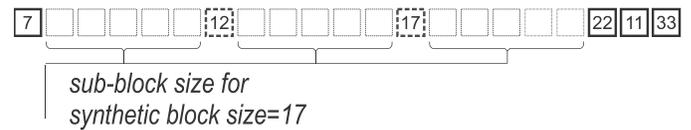


Parameters for data augmentation

synthetic block size=17 sub-block size=6



Linear synthetic values



Random non-linear synthetic values

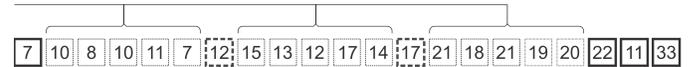


Fig. 2. Data Augmentation based on Time-warping and Jittering [8].

III. METHODOLOGY

The methodology followed for the implementation of the proposal is described below.

A. Time Series Selection

The selected daily wind speed time series is the same that was used in the work "in press" [7], and was obtained from the repository of the National Aeronautics and Space Administration (NASA) using Power Data Access Viewer with latitude: -17.6851 and longitude: -71.3515. This corresponds to a point in Ilo city in Peru that has enormous potential for wind energy.

This time series ranges from 1981-01-01 to the present, however, for the purposes of experimentation in this study, the years 1981-2016 will be used for training and the years 2017-2020 for testing.

B. Time Series Imputation

The selected daily wind speed time series does not present NA values, so the application of any data imputation technique was not necessary at this stage.

C. Data Augmentation

In this phase, the data augmentation technique based on time-warping and jittering proposed in [8] was configured according to Table I.

TABLE I. PARAMETERS OF DATA AUGMENTATION TECHNIQUE

Time series	Augmented time series	Block-Size	Sub-Block Size	Augmented ítems	Total
1981-2016	TS-1	6	3	78888	92037
	TS-2	6	3	78888	92037
Ítems 13149	TS-3	6	4	78888	92037
	TS-4	6	4	78888	92037

As can be seen in Table I, the first two augmented time series (TS-1 and TS-2) have the same parameters as well as the third and fourth (TS-3 and TS-4), but due to the randomness of the data augmentation technique different items are generated for each synthetic block, this can be seen in Fig. 3.

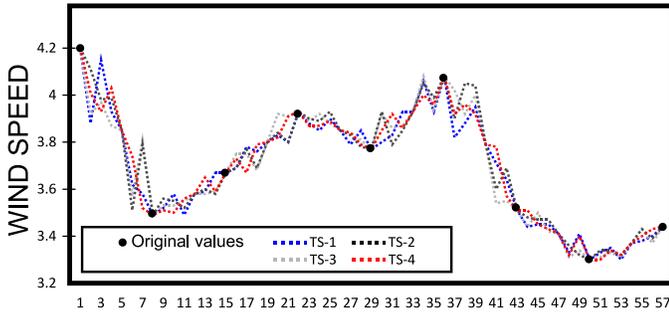


Fig. 3. Augmented Time Series for 9-first Original Items.

D. Ensemble Model Implementation (E-GRU)

At this stage, the ensemble model is implemented. Here the four sub-models have the same characteristics, which are detailed in Table II.

TABLE II. HYPERPARAMETERS OF EACH SUB-MODEL

Sub-Model	Hyperparameters	Values
GRU GRU GRU GRU	Hidden neurons	160
	Epochs	100
	Optimizer	adam
	Drop rate	0.2
	Activation function	ReLu
	Layer 1, 2, 3 y 4	(40,40,40,40)
	Batch size	40

The tools used for implementation of proposal model are Google Colab and tensorflow 2.4.1

E. Evaluation

For the evaluation of the predicted days, it is necessary to extract those corresponding to the original data since these also include predicted synthetic values. For this process, the value of the block-size parameter of the data augmentation technique is considered, which we will call z ; the predicted time series begins to be traversed and the predicted value located after the z value is extracted, then z new positions are traversed and the next value is extracted, and so on until reaching the last predicted value.

The model is evaluated through three regression metrics, these correspond to the Root Mean Square Error (RMSE), Relative RMSE (RRMSE) and Mean Absolute Percentage Error (MAPE), which are estimated through equations (5), (6) and (7) respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (5)$$

$$RRMSE = \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n O_i} * 100 \quad (6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{(O_i - P_i)}{O_i} * 100 \right| \quad (7)$$

A graphical version of the proposal model (E-GRU) can be seen in Fig. 4.

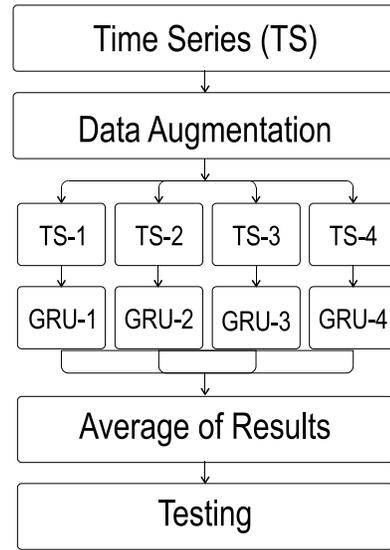


Fig. 4. Proposal Ensemble Model.

IV. RESULTS AND DISCUSSION

After experimentation, this section shows and describes the results achieved.

A. Results

According to Table III and Fig. 5, it can be seen that the ensemble proposal model E-GRU on average surpasses all the sub-models.

Regarding the RMSE, on average E-GRU is superior to all sub-models. However, for the forecast horizon of 500 days, GRU-1 (0.0284) slightly exceeds E-GRU (0.0288), this is the horizon where E-GRU reaches its worst performance.

According to RRMSE on average and in all prediction horizons, E-GRU outperforms all sub-models. It is important to highlight that according to the RRMSE achieved, E-GRU and all the sub-models can be classified as excellent since they present RRMSE < 10% [12], [13].

With respect to MAPE, like the previous metrics, on average E-GRU outperforms all sub-models. However, it is important to highlight that GRU-1 for the horizons of 50 and 100 predicted days, manages to surpass E-GRU.

According to Fig. 6, the importance of the ensemble process in the proposal can be appreciated. The data predicted by the sub-models closely approximates the original data by default and excess, and the average operation of the ensemble model makes it much closer to these, making E-GRU more accurate than the sub-models.

Likewise, it is important to highlight the importance of each sub-model, thus in Fig. 6 for the point enclosed in the circle, GRU-4, the worst of the sub-models according to Table III, is the only one that contributes to improving the proposal model precision.

TABLE III. SUB-MODELS VS MODEL RESULTS

Model/ Sub-Model	Predicted Days						Avg
	50	100	250	500	1000	1461	
GRU-1							
RMSE	0.0188	0.0235	0.0313	0.0284	0.0292	0.0298	0.0268±0.0050
RRMSE	0.5791	0.6854	0.9027	0.8102	0.8389	0.8549	0.7785±0.1297
MAPE	0.4166	0.5031	0.6041	0.5525	0.5669	0.5759	0.5365±0.0723
GRU-2							
RMSE	0.0299	0.0353	0.0348	0.0331	0.0339	0.0352	0.0337±0.0021
RRMSE	0.9219	1.0269	1.0010	0.9427	0.9723	1.0087	0.9789±0.042
MAPE	0.8045	0.8514	0.8121	0.7050	0.7335	0.7596	0.7776±0.0602
GRU-3							
RMSE	0.0206	0.0320	0.0385	0.0337	0.0349	0.0350	0.0324±0.0067
RRMSE	0.6371	0.9335	1.1091	0.9604	1.0009	1.0019	0.9404±0.1759
MAPE	0.4458	0.6675	0.7851	0.6704	0.6716	0.6774	0.6529±0.1236
GRU-4							
RMSE	0.0511	0.0434	0.0461	0.0419	0.0426	0.0433	0.0447±0.0037
RRMSE	1.5770	1.2656	1.3290	1.1929	1.2234	1.2406	1.3047±0.1537
MAPE	1.2298	1.0035	1.0455	0.9682	0.9930	1.0049	1.0408±0.1053
Proposal Ensemble Model (E-GRU)							
RMSE	0.0177	0.0230	0.0255	0.0288	0.0238	0.0247	0.0239±0.0040
RRMSE	0.5459	0.6713	0.7333	0.6498	0.6839	0.7069	0.6651±0.0691
MAPE	0.4375	0.5142	0.5210	0.4622	0.4727	0.4869	0.4824±0.0354

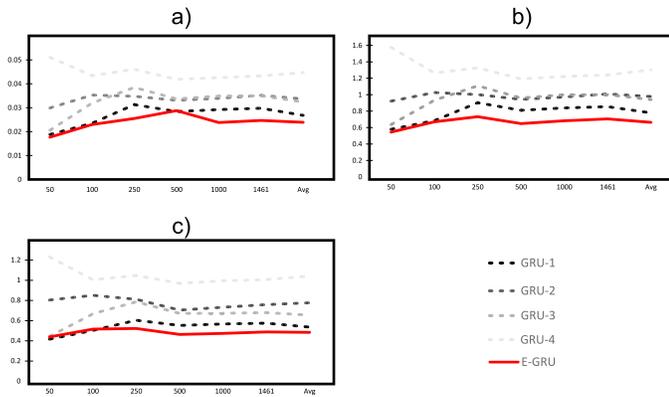


Fig. 5. Comparison of Metrics: Sub Models vs Proposal Model. a) RMSE, b) RRMSE and c) MAPE.

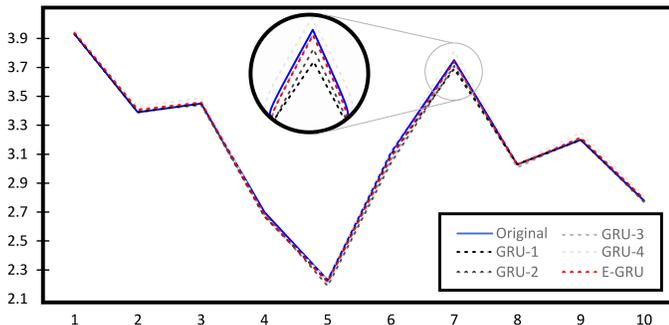


Fig. 6. Comparison of First 10 Predicted Days for Sub Models and Proposal Model.

According to Table IV, in reference to the average and all the prediction horizons it can be seen that the ensemble proposal model E-GRU far exceeds the results of the benchmark models (LSTM and GRU). Here it is important to highlight that the architecture of the LSTM and GRU models is four-layer and use the same hyperparameters as the sub-models of ensemble proposal model, but they do not use data augmentation.

Regarding the RRMSE, there is an average difference of approximately 15% between the results of the ensemble proposal model (E-GRU) and the benchmark models. Likewise, with respect to MAPE, the percentage difference is approximately 11%.

B. Discussion

In this part, the results achieved by the ensemble proposal model E-GRU are compared with those achieved by other state-of-the-art models in the prediction of wind speed time series.

Here, according to Table V, the high precision of the models proposed by Qureshi et al [14] and Flores et al [7] can be highlighted. In the first case, the authors use an architecture based on Deep Neural Networks and Meta Regression with Transfer Learning (DNN MRT), reaching an RMSE = 0.0953. In the second case, the authors use an architecture based on the recurrent neural network GRU including data augmentation, reaching an RMSE = 0.0876.

The E-GRU proposal model uses the same GRU architecture of [7] for each sub-model as well as the same data augmentation technique, the fundamental difference is that instead of using a single augmented time series, it uses four augmented time series, which are different due to the randomness of the technique and also work with different values for the sub-block size parameter.

The results show that the proposal ensemble model manages to surpass the state-of-the-art models including the techniques proposed in [14] and [7].

TABLE IV. BENCHMARK MODELS VS PROPOSAL MODEL RESULTS

Model/ Metric	Predicted Days						Avg
	50	100	250	500	1000	1461	
GRU GRU GRU GRU							
RMSE	0.4828	0.5680	0.5761	0.5181	0.5190	0.5146	0.5298±0.0354
RRMSE	14.9025	16.5702	16.592	14.770	14.896	14.744	15.4127±0.907
MAPE	13.0355	14.1669	13.929	12.124	12.314	12.276	12.9745±0.892
LSTM LSTM LSTM LSTM							
RMSE	0.4748	0.5711	0.5824	0.5224	0.5224	0.5380	0.5319±0.0392
RRMSE	14.6557	16.6608	16.772	14.881	14.994	14.843	15.4680±0.973
MAPE	0.4748	13.9701	13.886	12.040	12.164	12.115	10.7751±5.124
Proposal Ensemble Model (E-GRU)							
RMSE	0.0177	0.0230	0.0255	0.0288	0.0238	0.0247	0.0239±0.0040
RRMSE	0.5459	0.6713	0.7333	0.6498	0.6839	0.7069	0.6651±0.0691
MAPE	0.4375	0.5142	0.5210	0.4622	0.4727	0.4869	0.4824±0.0354

TABLE V. COMPARISON WITH RESULTS OF RELATED WORK

Work	Technique	Freq.	Train	Test	RMS E
Zhang et al, 2013 [15]	WTT+SAM+RB FNN	Daily	696	48	0.88
Bokde et al, 2018 [16]	EEMD+PSF	Hourly	2160	720	0.36
Mezaache et al, 2018 [17]	AE+ENN	10-minutes	26000	11000	3.0506
Khodayar et al, 2019 [18]	RBM+IPDL	10-minutes	105120	52560	11.126
Li et al, 2019 [19]	CNN+LSTM	15-minutes	3500	500	3.0012
Liu et al, 2019 [20]	GRU	Daily	811	372	0.9899
Deng et al, 2019 [21]	Bi-GRU			400	6.75
Jiang et al, 2019 [22]	VWC		2304	576	0.2557
Wang et al, 2019 [23]	EWT+KLD	Hourly	14016	3504	1.07
Qureshi et al, 2017 [14]	DNN+MRT	Hourly			0.0953
Yan et al, 2020 [24]	ISSD+LSTM-GOADB	Hourly	600	100	1.0156
Cheng et al, 2020 [25]	MSSO	10-minutes	2880	720	0.3002
Altan et al, 2020 [26]	DM+LSTM+GWO	10-hours	4397	775	0.1878
Noman et al, 2020 [27]	NARX	10-minutes	Data 2017	Data 2018	0.3590
Luo et al, 2020 [28]	DE+MOO	10-minutes	3200	800	0.2348
Flores et al, 2021 [7]	GRU	Daily	13149	1461	0.0876
Tian et al, 2021 [29]	IWOA-ESN	Hourly	800	200	0.8544
Proposal Model	GRU	Daily	13149	1461	0.0239

V. CONCLUSION AND FUTURE WORK

According to what is observed in the Results and Discussion section of this paper, it can be concluded that the proposal model allows to improve the results of the state of the art in relation to wind speed forecasting. Likewise, it is important to highlight the importance of the data augmentation process, since all the sub-models implemented for the ensemble proposal model E-GRU present excellent results according to the RRMSE evaluation. Thus, the main advantage of the proposal model with respect to the state-of-the-art models for wind speed prediction is its high precision, and the simplicity of model implementation and each of its respective sub-models.

As a future work, it should be noted that the main weakness of the proposal model lies in the computational cost involved in training 4 GRU models with 92,037 items each. Thus, the minimum amount of synthetic and historical data

could be analyzed to obtain satisfactory results. On the other hand, it could be experimented with time series with characteristics different from those of wind speed.

ACKNOWLEDGMENT

To the National University of Moquegua (UNAM), which, through mining canon funds, allowed the execution of the study and publication of its results.

REFERENCES

- [1] "What is the greenhouse effect," American Chemical Society, [Online]. Available: www.acs.org. [Accessed 24 06 2021].
- [2] B. Bose, "Global warming: energy, environmental pollution, and the impact of power electronics," IEEE Industrial Electronics Magazine, vol. 4, no. 1, pp. 6-17, 2010.
- [3] P.J. Zucattelli, E.G.S.. Nascimento, A.Á.B. Santos, A.M.G. Arce, D.M. Moreira, "An investigation on deep learning and wavelet transform to nowcast wind power and wind power ramp: A case study in Brazil and Uruguay," Energy, vol. 230, no. 120842, 2021.
- [4] J. Yeomans, S. Thwaites, W.S.P. Robertson, D. Booth, B. Ng, D. Thewlis, "Simulating time-series data for improved deep neural network performance," IEEE Access, vol. 7, pp. 131248-131255, 2019.
- [5] G. Forestier, F. Petitjean, H.A. Dau, G.I. Webb, E. Keogh, "Generating synthetic time series to augment sparse datasets," in IEEE International Conference on Data Mining, New Orleans, USA, 2017.
- [6] B.K. Iwana, S. Uchida, "Time series data augmentation for neural networks by time warping with a discriminative teacher," arXiv.org, pp. 1-9, 2020.
- [7] A. Flores, H. Tito-Chura, V. Yana-Mamani, "Wind speed time series prediction with deep learning and data augmentation "in press"," in Lecture Notes in Networks and Systems, Amsterdam, Netherlands, Springer, 2021.
- [8] A. Flores, H. Tito, H. Apaza-Alanoca, "Data augmentation for short-term time series prediction with deep learning," in Lecture Notes in Networks and Systems, London, United Kingdom, Springer, 2021.
- [9] A. Flores, H. Tito, D. Centy, "Comparison of Hybrid Recurrent Neural Networks for Univariate Time Series Forecasting," Advances in Intelligent Systems and Computing, p. 375-387, 2021.
- [10] J. Zheng, X. Chen, K. Yu, L. Gan, Y. Wang, K. Wang, "Short-term power load forecasting of residential community based on GRU neural network," in International Conference on Power System Technology, Guangzhou, China, 2018.
- [11] C. Shorten, T.M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 60, pp. 1-48, 2019.
- [12] A.N.-L. Huynh, R.C. Deo, D.-A. An-Vo, M. Ali, N. Raj, S. Abdulla, "Near real-time global solar radiation forecasting at multiple time-step horizons using the Long Short-Term Memory network," Energies, vol. 13, no. 14, p. 3517, 2020.
- [13] R. Khelifi, M. Guermoui, A. Rabehi, D. Lalmi, "Multi-step ahead forecasting of daily solar radiation components in Saharan climate," International Journal of Ambient Energy, vol. 41, no. 6, pp. 707-715, 2020.
- [14] A.S. Qureshi, A. Khan, A. Zameer, A. Usman, "Wind power prediction using deep neural network based meta regression and transfer learning," Applied Soft Computing, vol. 58, p. 742-755, 2017.
- [15] W. Zhang, J. Wang, Z. Zhao, M. Tian, "Short-term wind speed forecasting based on a hybrid model," Journal of Applied Soft Computing, vol. 92, no. 106294, pp. 1-20, 2013.
- [16] N. Bokde, A. Feijoo, K. Kulat, "Analysis of Differencing and Decomposition preprocessing methods for wind speed prediction," Applied Soft Computing, vol. 71, pp. 926-938, 2018.
- [17] H. Mezaache, H. Bouzgoud, "Auto-encoder with neural networks for wind speed forecasting," in IEEE International Conference on Communications and Electrical Engineering, El Oued, Algeria, 2018.
- [18] M.I. Khodayar, J. Wang, M. Manthouri, "Interval deep generative neural network for wind speed forecasting," IEEE Transactions on Smart Grid, vol. 10, no. 4, pp. 3974 - 3989, 2019.

- [19] L. Gang, T.F. Wang, F.X. Hu, T.C. Liu, "Algorithm considering multi-dimensional meteorological feature extraction in short-term wind speed prediction," in IEEE Information Technology, Networking, Electronic and Automation Control Conference, Chengdu, China, 2019.
- [20] M. Liu, P. Qiu, K. Wei, "Research on wind speed prediction of wind power system based on GRU deep learning," in IEEE Conference on Energy Internet and Energy System Integration, Changsha, China, 2019.
- [21] Y. Deng, H. Jia, P. Li, X. Tong, X. Qiu, F. Li, "A deep learning methodology based on bidirectional gated recurrent unit for window power prediction," in IEEE, Xi'an, China, 2019.
- [22] P. Jiang, Z. Liu, "Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting," Applied Soft Computing Journal, vol. 82, no. 105587, pp. 1-19, 2019.
- [23] J. Wang, Y. Li, "An innovative hybrid approach for multi-step ahead wind speed prediction," Applied Soft Computing Journal, vol. 78, p. 296-309, 2019.
- [24] X. Yan, Y. Liu, Y. Xu, M. Jia, "Multistep forecasting for diurnal wind speed based on hybrid deep learning model with improved singular spectrum decomposition," Energy Conversion and Management, vol. 225, no. 113456, pp. 1-22, 2020.
- [25] Z. Cheng, J. Wang, "A new combined model based on multi-objective salp swarm optimization for wind speed forecasting," Applied Soft Computing Journal, vol. 92, no. 106294, pp. 1-20, 2020.
- [26] A. Altan, S. Karasu, E. Zio, "A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer," Applied Soft Computing, vol. 100, no. 106996, 2020.
- [27] F. Noman, G. Alkaws, A.A. Alkahtani, A-S, Al-Shetwi, S.T. Tiong, N. Alalwan, E. Janaka, A.I. Alzahrani, "Multistep short-term wind speed prediction using nonlinear auto-regressive neural network with exogenous variable selection," Alexandria Engineering Journal, vol. 60, no. 1, pp. 1221-1229, 2020.
- [28] L. Luo, H. Li, J. Wang, J. Hu, "Design of a combined wind speed forecasting system based on decomposition-ensemble and multi-objective optimization approach," Applied Mathematical Modelling, vol. 89, pp. 49-72, 2021.
- [29] Z. Tian, H. Li, F. Li, "A combination forecasting model of wind speed based on decomposition," Energy Reports, vol. 7, pp. 1217-1233, 2021.

Relative Merits of Data Mining Algorithms of Chronic Kidney Diseases

Harsha Herle¹, Dr. Padmaja K V²

Electronics and Instrumentation Engineering, Department
RV College of Engineering, Affiliated to VTU, Bengaluru, India

Abstract—Early prediction of Chronic Kidney Disease in human subjects is considered to be a critical factor for diagnosis and treatment. The use of data mining algorithms to reveal the hidden information from clinical and laboratory samples helps physician in early diagnosis, thus contributing towards increase in accuracy, prediction and detection of Chronic Kidney Disease. The experimental results obtained from this work, with subjected to optimal data mining algorithms for better classification and prediction, of Chronic Kidney Disease. The result of applying relevant algorithms, like K-Nearest Neighbors, Support Vector Machine, Multi Layer Perceptron, Random Forest, are studied for both clinical and laboratory samples. Our findings show that K - Nearest Neighbour algorithm provides the best classification for clinical data and, similarly, Random Forest for laboratory samples, when compared with the performance parameters like, precision, accuracy, recall and F1 Score of other data mining analysis techniques.

Keywords—*Ultrasound images; support vector machine (SVM) k-nearest algorithm (K-NN); multilayer perceptron algorithm (MLP); random forest (RF); clinical data*

I. INTRODUCTION

In recent years more than two million people across the globe suffer from Chronic Kidney Disease (CKD) like Kidney stone, kidney transplant, blockage of urine, congenital anomalies, cyst, bacterial infection, dialysis or cancerous cells, to stay alive, of which at least only 10% of the patients need treatment to live with high health care costs [1]. Due to the increase in cost, only 2 million people are capable of receiving treatment for CKD, representing 12% of the global population [2-3]. Within developed countries, only 20% of patients are treated for CKD, and under developed countries, on an average one million die from untreated kidney failure due to financial constraints [9]. CKD can be detected using either X-rays, Ultra Sound (US), Computer Tomography (CT) and MRI or laboratory samples for medication. This work focuses on applying data mining techniques for kidney stone detection using Ultra Sound (US) technique and laboratory samples collected from database, for further treatment by medical doctors.

In detecting the CKD, the laboratory samples obtained from standard database UCI machine learning repository is subjected to data cleaning or preprocessing of the data samples and for further classification, the labels are converted to numbers. The database contains 25 attributes of which serum creatinine, blood ureas, hemoglobin, hypertension remains important in this work and others are out of the scope for our analysis [3].

Once the CKD is detected, the next step is to focus on detecting the kidney stone using US technique. The US technique has more advantages like non-ionising nature, portability, low cost and also in giving the details of real-time monitoring of patient's vital internal organs. US image is recorded by incisive technique, where a high frequency signal of order greater than 1MHz is penetrated into the human body with the help of a transducer. The US waves reflected from kidney tissues are received by transducer and displayed on a computer screen either in two-dimensional (2D) or three dimensional (3D). The obtained US images contain background information and labels that require crossing out and to enhance the quality of US image. Further, the US wave was subjected to speckle noise -multiplicative type of noise that appear as dark and bright spots resulting in trouble analysis and diagnosis interpretation. In reduction of speckle noise in US images, a method termed as speckle denoising is carried out for analysis, preprocessing and interpolation of US images [5]. In removal of speckle noise from US images many algorithms exists and differ in their basic methodologies. Preprocessing filters like, spatial and wavelet filters have proved its efficiency, in-terms of statistical parameters like increased Signal to noise ratio (SNR), Peak Signal to Noise Ratio (PSNR), decrease in Mean Squared Error (MSE), Mean Absolute Error (MAE).

Moreover in US image, the Region of Interest (RoI) is retained and other portions are removed using segmentation techniques. Segmentation is the logical implementation to find the RoI against the characteristic of the US images, thus its features are expected to find the region where kidney stone may be present. Feature extraction method aims at reducing the input data by finding the features from several input patterns, resulting in an input vector consisting of appropriate image properties, which will be given to data mining techniques for further classification.

In classification phase, the input image is classified into abnormal or normal classes depending on the statistical parameters of features obtained from clinical and laboratory data samples. Generally, classification phase is further categorized into unsupervised and supervised classification, where, in supervised classifier, the algorithm iteratively arrives at predictions on the training data and is validated by the teacher, often coined as learning with teacher. Conversely, in unsupervised learning, algorithms are left to their own devices to determine and present the interesting structure in the data, hence does not necessitate training phase for classification. The supervised learning, further classified into,

artificial/logical, Perceptron based, statistical based, and Support Vector Machines (SVM).

This work focus on statistical analysis of clinical and laboratory data samples, where data mining algorithm, K-Nearest Neighbors algorithm, SVM, Multi Layer Perceptron (MLP) and Random Forest (RF) and are used to evaluate the performance of the classifier against different parameters like sensitivity, precision, Recall and F1 score. The experimental results validated the above statistical parameter in estimating the best machine learning algorithm for CKD. Upon classification of kidney stone US images, a Graphical User Interface (GUI) is developed to assist medical doctors about the presence or absence of Kidney US images.

The preceding section in the paper foresees the literature review carried out to find the problem of interest. The research methodology for the problem defined is well stated. The description of constituent blocks and its mathematical relations are well explained and derived. Finally experimental results demonstrate the effectiveness of best suitable algorithm for kidney US Images with conclusion drawn towards it.

II. RELATED WORK

The medical imaging plays a prominent role in detection and diagnosis of diseases related to human subjects, have wide range of scope, thereby researchers and scientists have contributed significantly over decades [6]. To begin with, the laboratory data samples obtained are subjected to transformation and preprocessed or data cleaned for replacing the missing values using data mean method. The data samples are applied to classifiers to evaluate the performance of the algorithms to detect the presence of CKD in human subjects. Upon detection of CKD, the next step is to find the presence of kidney stone using US images.

The raw US image is obtained from the radiologist, unwanted details like anatomical information, are removed by binary threshold method. The US images are subjected to removal of speckle noise either using statistical or classification of model. The main objective in statistical modeling is to remove the noisy images by obtaining the statistical features from training data and later obtain the parameters of interest. Initially, statistical filters such as Weiner filter [20], adaptive filters in spectral domains, finds applicable in removal of additive noise [7]. In order to address the multiplicative noise, Jain mode is proposed [8] that works by taking the logarithmic of the image is obtained, later multiplicative noise is converted to additive noise, and Weiner filtering is applied. This process is tedious, time consuming and depends on the size of the image. A region based segmentation method is applied to kidney along with Gabor filter, resulting in reduction of speckle noise and smoothen the image signal, along with histogram method to improve the quality of the image [9]. The experimental results demonstrate reduction of speckle noise up-to to 85%, focusing on only few parameters. Thresholding methods like soft thresholding, Visu-Shrink, hard thresholding, Sure-Shrink, Bayes sure shrink and Bayes thresholding for speckle reduction were also applied for kidney US images [10]. The Visu shrink method is based on wavelet shrinkage and uses over smooth images. The Bayes shrink gives promising results in Mean

squared Error (MSE) over visu shrink, all these methods are based on soft thresholding, with the input value is shrunk to zero by the amount of thresholding. In hard thresholding, the input is retained to same value, if it is greater than threshold, else retained to zero [11]. After preprocessing, the next step is to extract the features from US images followed by applying data mining techniques to detect into normal or abnormal using Graphical User Interface (GUI) developed.

The most prominent work involves detection of absence or presence of kidney stones. The US images were segmented using intensity threshold variation that helps in identifying multiple classes to classify the images as stone, early stone stages and normal, [13]. Other methods in feature extraction can be intensity histogram features and Gray Level Co-Occurrence Matrix (GLCM) features. The kidney US images were classified into four different groups like Normal (NR), Bacterial Infection (BI), Cystic Disease (CD) and kidney stones (KS). Thus create the database for classified kidney US image for further pathological studies [12-13].

In classifying the abnormalities in kidney, statistical methods like GLCM or Run Length Matrix (RLM) are used along with SVM, reaching an accuracy of 85.8% [21]. Increased in classification accuracy up-to 98.8% can be reached by applying two level set segmentation methods with Artificial Neural Network (ANN) architecture [14]. Intensity histogram and Harlick features were used as feature extraction for segmented Region of Interest (RoI) Kidney US images A two level of classification methods is proposed, of which in the first method, a lookup table based approach was used to classify the US images into normal and abnormal, followed by second stage, where an SVM with MLP was used to classify the presence of stone or cyst in the kidney with promising experimental results reaching an accuracy up-to 98.14%. SVM was used to classify the Kidney US images for early detection of CKD for classifying the training and testing results and to evaluate the performance of SVM with accuracy of 97.6%. The prominent features were extracted from abnormal kidney US images and classified using SVM algorithm reaching accuracy up-to 83.74% [15]. Likewise, the US images were subjected top adaptive median preprocessing method and segmented using K-Means method. GLCM features were extracted and meta-heuristic SVM classifier is used to classify the US images to detect the renal calculi, and have performed better in noisy images exhibiting detection accuracy of 98.8%.

Further, the other machine algorithms like K-NN is used in classification for normal and cyst Kidney US images, with experimental results predicting up-to 92% for normal images and 85% for Cystic images [16]. In addition to this, several machine algorithms like, Logistic Regression, Elastic Net, Lasso Regression, Ridge Regression, SVM, Random Forest, Neural Network, K-NN Elastic Net were applied for kidney US images, of which K-NN alone gives prediction accuracy of 85% [17]. Extending further, a new decision support system was developed to predict and classify CKD using Artificial Neural Network, K-NN, decision tree, Random Subspace, Linear Discriminate Analysis (LDA), of which K-NN alone gives 78% prediction accuracy [19]. Recently, Hybrid classification algorithms play a prominent role in classification of US kidney images, were subjected to SVM-

KNN together, reaching the prediction accuracy up-to 99.6% [17]. Recent advancements in US imaging have revealed the way to increased interest in removal of the speckle noise from the medical images using different algorithms, without reducing much of the diagnostic information [20].

Data mining techniques are applied for laboratory data set consisting of 361 CKD patients. SVM, MLP, Radial Basis function (RBF), Probabilistic neural network (PNN) were applied to these data samples with PNN emerging as the best algorithm that can be used for physicians for further treatment [4].

Of the literature reviewed in this work, the kidney US images are subjected to preprocessing method like Median, Adaptive median, Weiner filter to remove salt and pepper noise, Neigh SURE shrink method to remove speckle noise. The resultant images were applied with GLCM and histogram method for feature extraction. This current work focus data mining technique like SVM, MLP, RF and KNN for the preprocessed Kidney US image to increase the classifier performance. Further the kidney US Images were subjected to data augmentation methods like rotations of US kidney images to enhance the number of images, thereby contributing to increase in overall classification accuracy. The experimental results of both clinical and laboratory data are compared to estimate the optimal data mining algorithm for early diagnosis.

III. RESEACRH METHOD

Fig. 1 shows the block diagram for estimating optimal data mining technique for CKD, using both clinical and laboratory samples. The laboratory data sample obtained from UCI machine learning repository consists of 400 CKD Indian patients and contains 11 numerical and 14 attributes that are categorized. The obtained datasets are preprocessed or data cleaned. The data sample are then passed to different data mining techniques to find the optimal classifier algorithm with respect to accuracy, precision, recall and F1 score. All the preprocessing methods, classification and detection area are carried out in sklearn machine learning API using python programming environment. Upon detection of CKD using laboratory data sets, the next step is to use the same classifiers to detect the presence or absence of kidney stone using Ultrasound Imaging technique.

The raw image is obtained from the radiologist, and unwanted anatomical information are removed using binary threshold method. The resultant noisy image is then filtered to suppress the Salt and Pepper noise, speckle noise using Median, Adaptive Median, Weiner and Neigh SURE shrink filtering method. The preprocessed image is segmented by threshold method and morphologically operated to detect RoI in kidney, and then the statistical features are extracted from normal and abnormal images to classify the kidney abnormalities. If the abnormalities are detected, then the Centroid of the abnormal image is estimated to find the area of kidney abnormal region. The detected abnormalities will provide information for medical doctors, to take the medication to next level. All the preprocessing methods, classification and detection of centroid area for kidney stone detection are carried out in image acquisition tool box in

MATLAB® 2018a version simulation environment for image resolution of 256×256 using high end computers.

A. Preprocessing

In preprocessing the laboratory samples, data sets are subjected to interpolation, transformation and scaling, where interpolation method includes mean, median or most frequent values. Data transformation, method converts label to number and scaling is a process to normalize the values in the data set. Likewise, for Kidney US images, the acquired image undergo different preprocessing stages to increase the quality of images [39]. The acquired images are prone to speckle noise, salt and pepper noise, which requires to be filtered out. In this work, Neigh SURE shrink, median, adaptive median, Weiner filter are used to remove the both the noise and un-sharp masking for sharpening. Entropy based segmentation is used for finding the RoI and morphological operations like erosion and dilation are also used for finding the final segmented image.

1) *Median filters*: Median filters are non-adaptive filters that adjust the coefficients based on data within a rigid moving window, and summons between the quality of speckle noise suppression and the potential to preserving image details. They are linear statistical filter which works with the principle of replacing the current pixel value in US images into the median value in a neighborhood. [19-20].

2) *Adaptive median filers*: This filter works with the three step methodology of which the primary step is to detect the noise, in US images, secondly, adaptively estimate the window size based on the number of noise pixels within the window. Finally, determine the weight of each non-noise point in filtering window and filter off noise points by means of weighted median filtering algorithm. This method is advantageous compared to other preprocessing filters, as it preserves the edges of its high frequency parts of an image. [25].

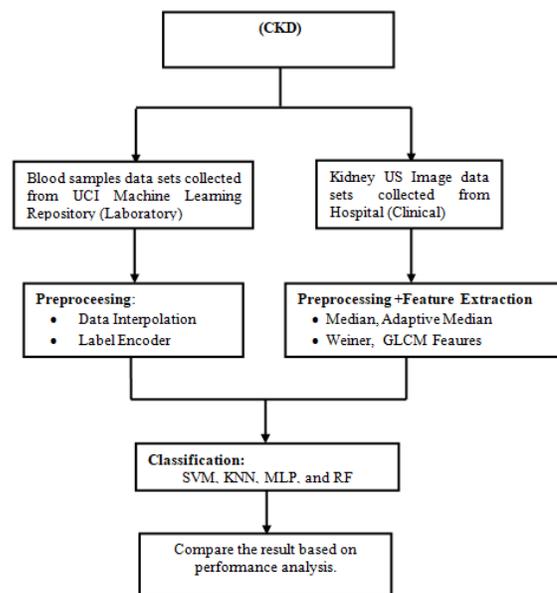


Fig. 1. Block Diagram for Estimating the Optimal Data Mining Algorithm for CKD.

3) *Weiner filters*: Another common preprocessing filter that finds application in US image is Wiener filter that works on the principle of blurring and removal of additive noise simultaneously by performing best possible substitution between inverse filter and noise smoothing [18] and is optimal in terms reduction in Mean Square Error [22].

4) *Neigh SURE shrink filters*: Neigh SURE shrink is the adaptive filter based method works on the principle of thresholding technique, combining both soft thresholding and hard thresholding, and only deals with binary values. In this method, a fixed threshold value is considered based on the window size, $[W_s]$ and variance value, is calculated using the equation 1,

$$\sigma^2 = \frac{\text{median}(|w_s|)}{0.6745} (w_s \in \text{subband LL}) \quad (1)$$

The smooth or low frequency components are approximated from father wavelet and high frequency components are resembled from mother wavelet. A 3x3 sized window is moved all over the image to be filtered, where the surrounding pixel value is averaged over the central pixel. Mean and variance is applied over the decomposed image to reduce the speckle noise from US image reduction in speckle noise is based on the thresholding value, when the pixel value crosses the threshold value then that pixel value becomes 1 or 255 and if the pixel value is below the threshold value then that pixel value becomes 0. The Neigh Shrink uses a suboptimal universal threshold and identical window size in all wavelet subbands, whereas the improved version of it determines an optimal threshold and neighboring window size for every subband by the Stein's unbiased risk estimates [23] as given in equation 2.

$$(T^s, k^s) = \arg \min_{T,k} \text{SURE}(w_s, \rho, k) \quad (2)$$

Where ρ is the threshold, k is the window size and s denotes the sub band [24].

B. Feature Extraction

Feature extraction primarily focuses on reducing the input data by determining the distinguishable features from numerous input data samples. The output of feature extraction stage results in an input vector consists of significant image properties, which will be fed into classifier to classify normal and kidney stone US images. In medical imaging analysis, texture is an important feature that provides the spatial distribution of pixels gray level in a region. Textural characteristics of particular image is extracted, to perform the identification of object regions, using many diverse algorithms like fractal based methods, Markov random field and Gabor filter. Harlick features often coined as Gray level co-occurrence matrix (GLCM) is a statistical feature extraction method, represented in the form, $N_g \times N_g$, where N_g is number of gray levels, Within MATLAB, GLCM is represented in the form of a matrix where number of rows and columns are equal to number of gray levels, G in the image. GLCM gives spatial relationships between pixels $P(i,j)$ [26].

Of various features, this work focuses on homogeneity, contrast, correlation, and energy. Another prominent feature extraction method, is Gray Level Run Length Matrix

(GLRLM), can be used to extract the higher order statistical features in a US kidney images. Unlike GLCM, the GLRLM is a two dimensional matrix of $N_g \times R$ element in which each element $P(k, l|0)$ gives the total occurrence of runs having length k of gray level L in a given direction, with R representing the longest run. All the seven statistical features namely, Short Runs Emphasis, Long Runs Emphasis, Gray Level Non Uniformity, Run Length No uniformity, Run Percentage, Low Gray level Runs Emphasis, and High Gray level Runs Emphasis are used in this work. [21].

C. Data Mining / Classifier

The principle behind the classification stage is to categorize the input image into normal or abnormal class depending on the features extracted, based on statistical parameters. The data mining techniques like SVM, KNN, MLP and RF are used in the present work for classification of clinical and laboratory data samples.

1) *Support Vector Machine (SVM)*: SVM classification is a non probabilistic linear binary classifier that analyzes the input data and predicts one of the two classes it belongs to, that are separated by hyper plane, constructed using structural risk minimization principle. The SVM is independent of the dimensionality of the feature space that outperforms as compared to other classifiers with minimum training samples. [27].

2) *K- Nearest Neighbour (KNN)*: Another important classifier used in clinical and laboratory data samples, K-NN classifier that classify the datasets according to majority of its neighbors belongs to. Selection of number of neighbours is elective and user defined. [17].

To select the K , for the data sets, the KNN algorithm is executed several times for different values of K , the optimal value of K is chosen so that number of errors is reduced. Euclidean distance is one of frequently used method for distance measures to find the K-NN and is given by,

$$\text{Euclidean} = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (3)$$

The minimum distance between the training and test samples gives best classification of the test samples.

3) *Multilayer Perceptron Algorithm (MLP)*: Another significant class of classifier algorithm is MLP, consists of an input layer, one or more hidden layers, and the output layer, of which the output in different stage is activated either using linear or non-linear activation function. This algorithm is sub classified into two stages, wherein the feed forward stage, the output is estimated by weightage average of inputs and bias. In the backward stage, the errors are minimized by updating the weights [28].

4) *Random Forest (RF)*: Random forest algorithm works with the principle of creating the decision trees on data samples obtained, accumulate the prediction from each of the samples, and provides the optimal solution through voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result.

Generally RF consists of many decision trees, and the features are randomly selected from the optional features, thus allowing the tree at each node to grow without trimming. RF algorithms reaches high accuracy even for a portion of the data is missing in the data samples [29].

IV. RESULT AND DISCUSSION

During the preprocessing stage for all the 400 laboratory samples, the patient ID is removed from the dataset and missing predictor attributes/variable values are replaced using data interpolation method. For the categorized attributes like Hypertension, Coronary Artery Disease, Appetite, Pedal Edema, Pus Cells, Pus Cells Clumps, Diabetes Mellitus, and Anemia, the digitizing label encoder was also applied. The resultant data sets are further subjected to data mining techniques as mentioned above. Each data attribute contains distinct and unique features that also help in classification of CKD, which was shown in Table I.

TABLE I. ATTRIBUTES DESCRIPTION USED IN THE ANALYSIS (PRED-PREDICTOR, NUM-NUMERICAL, NOM-NOMINAL, TAR-TARGET)

Sl.N O.	Dataset Attributes	Class	Type	Misssing	Distinct	Unique
1	Age	Pred	Num	9	76	16
2	Blood Presure	Pred	Num	12	10	3
3	Specific Gravity	Pred	Nom	47	5	0
4	Albumin	Pred	Nom	46	6	1
5	Sugar	Pred	Nom	49	6	0
6	Red Blood Cells	Pred	Nom	152	2	0
7	Pus Cells	Pred	Nom	65	2	0
8	Pus Cells Clumps	Pred	Nom	4	2	0
9	Bacteria	Pred	Nom	4	2	0
10	Blood Glucose Random	Pred	Num	44	146	65
11	Blood Urea	Pred	Num	19	118	55
12	Serum Creatinine	Pred	Num	17	84	41
13	Sodium	Pred	Num	87	34	7
14	Potassium	Pred	Num	88	40	8
15	Hemoglobin	Pred	Num	52	115	28
16	Packed Cell Volume	Pred	Num	71	42	8
17	White Cell Blood Count	Pred	Num	106	89	31
18	Red Cell Blood Count	Pred	Num	131	45	3
19	Hypertension	Pred	Nom	2	2	0
20	Diabetes Mellitus	Pred	Nom	2	2	0
21	Coronary Artery Disease	Pred	Nom	2	2	0
22	Appetite	Pred	Nom	1	2	0
23	Pedal Edema	Pred	Nom	1	2	0
24	Anemia	Pred	Nom	1	2	0
25	Class	Tar	Nom	0	2	0

The next step is to use the mentioned-preprocessing filters for speckle noise reduction, applying for both normal and kidney stone US images of clinical data samples. The different spatial filters like, Median, Adaptive median, Weiner are used in estimating statistical parameters. Further, for the different noise variance (NV) varying from 0.01 to 0.08, the statistical parameters like Peak Signal to Noise Ratio (PSNR in dB), Signal to Noise Ratio (SNR in dB), Root Mean Square Error (RME) and Mean Absolute Error (MAE) are estimated for both normal and abnormal kidney images.

Further, for the US images, the experimentation begins with removal of background and labels associated with kidney images for both normal and kidney stone US Images, followed by enhancing the contrast and sharpening the quality of image for better performance as shown Fig. 2.



Fig. 2. (a) Normal Kidney Image. (b) Removal of Background and Label of Kidney Images. (c) Contrasted Image and Sharpened Image.

Table II represents statistical analysis of noise variance for different filters used in preprocessing of US images. The NV values from 0.01 to 0.08, for different filters, the statistical parameters like PSNR, SNR, MAE and RME are calculated and only the optimal values are provided in Table II. For Kidney stone US images, the adaptive median filter with NV value of 0.01 gives PSNR as 33.51dB and SNR of 18.01dB as compared with other filters. Likewise, the same method was applied to normal kidney US images and results are tabulated. The presence of kidney stone decreases the PSNR and SNR as compared to normal US images with the variation in RME and MAE.

Fig. 3 shows the preprocessed normal US images, from (a-c), and kidney stone US images from (d-f). It is evident that the speckle noise is reduced relatively better using median filter, but image is blurred and artifacts are introduced, as compared with other filters. In case of Adaptive filter, maximum speckle noise reduction is observed, and the edges features are well preserved. In Weiner filter, image is subjected to over enhancement, resulting in reduction in speckle noise, but posing diagnosis problems. Likewise, the same US images are also subjected to Neigh SURE Shrink wavelet filters with haar, db4, sym-8, and for decomposition levels 2 and 4.

From Fig. 4, for decomposition level-4, Wavelet Neigh SURE shrink along with haar, db-4, sym8, increases the contrast resolution that leads to superior visual quality. This, in turn, enhances the quality of kidney stone boundaries, reducing the speckle noise, increasing the SNR, when compared to decomposition level-2, thus, improving the image quality for further diagnosis. Likewise, comparing the decomposition level 2 and level 4, it is observed that decomposition level 4 improves the PSNR and SNR.

TABLE II. STATISTICAL ANALYSIS OF NV AGAINST DIFFERENT PREPROCESSING FILTERS

Kidney Stone US image					
Filter	Noise variance	PSNR in dB	SNR in dB	RME	MAE
Median	0.01	31.73	16.23	6.60	1.49
Adaptive Median	0.01	33.51	18.01	5.38	1.70
Weiner	0.04	31.85	16.35	6.51	2.58
Normal Kidney US image					
Median	0.01	30.84	18.50	7.31	3.74
Adaptive Median	0.01	33.70	21.36	5.26	2.76
Weiner	0.05	35.29	23.29	4.21	2.24

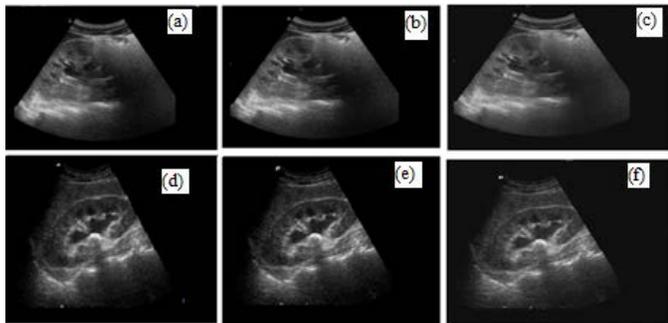


Fig. 3. (a-c) Results of Median Filter, Adaptive Median, Weiner Filter for Normal US Image (d-f) Results of Median Filter, Adaptive Median, Weiner Filter for Kidney Stone US Images.

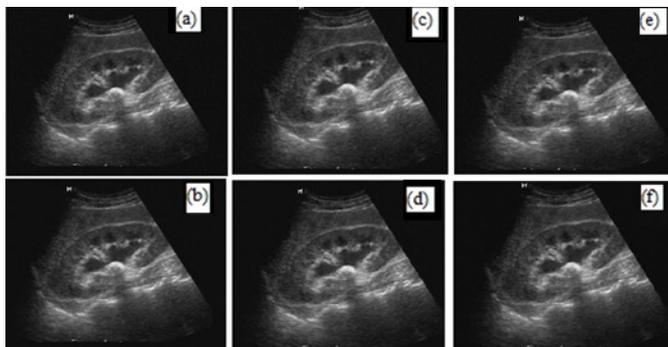


Fig. 4. Preprocessing Filters of Kidney Stone US Image, (a-b) Sys8 Level 2 and 4 Decomposition, (c-d) db4 Level 2 and 4 Decomposition, (e-f) Haar Level 2 and 4 Decomposition.

In Table III, it is observed that Neigh SURE shrink with sym8 for decomposition level 4 of kidney stone images, PSNR is found to be 41.82dB and SNR of 26.36dB, along with reduction in RME and MAE respectively. The presence of kidney stone in US images reduces PSNR and SNR values, compared to normal kidney images. Comparing Tables II and III, the increase in PSNR and SNR, reduction of RME and MAE can be observed. In Conclusion, sym8 Neigh SURE shrink filter is a far superior filter that can be employed in US Image preprocessing method.

TABLE III. STATISTICAL ANALYSIS OF NV AGAINST DIFFERENT WAVELET FILTERS

Kidney Stone US image					
Neigh SURE shrink	Decomposition Level	PSNR in dB	SNR in dB	RME	MAE
db4	4	41.62	26.12	2.11	1.34
Haar	4	41.33	25.83	2.18	1.37
sym8	4	41.86	26.36	2.03	1.30
Normal Kidney US image					
db4	4	41.70	29.36	2.09	1.46
Haar	4	40.10	27.75	2.51	1.72
sym8	4	42.21	29.87	1.97	1.39

Fig. 5 depicts the GUI developed using MATLAB, to detect the presence/absence of the kidney stone in US images Upon GLCM and GLRLM feature extraction, the next step in image processing is to classify the kidney US Images into normal and kidney stone US images. The normal and kidney stone US Images together are used in training and validation process for classification. A data mining/Classifier performance measure is based on accuracy, precision, recall and F-1 Score, i.e. number of sample that can be into normal or abnormal classes, hence this work proposes the application of SVM, KNN, MLP and RF classifier for the classification US images with tenfold cross validation to train and validate the classifiers.

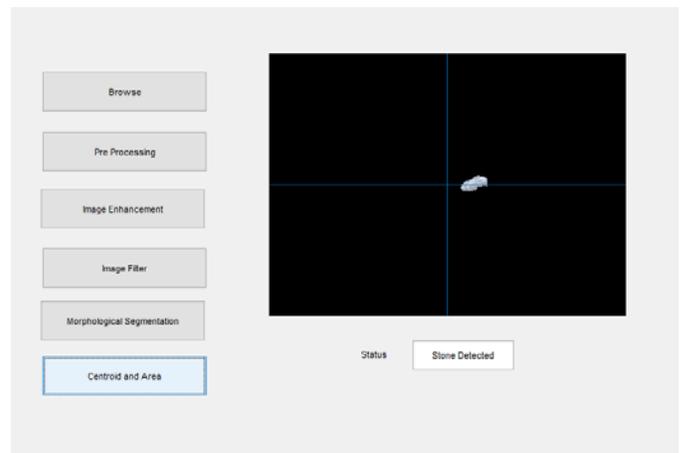


Fig. 5. RoI along with Graphical User Interface (GUI) to Detect the Presence of Kidney Stone.

True positive, True Negative, False Positive and False Negative are the confusion matrix features that are used for measuring the accuracy, precision, Recall, and F-1 Score of the classifier system. All these parameters are characterized by below relation.

$$\text{True Positive (TP)} = \frac{\text{Number of Image with kidney stone}}{\text{Total Number of kidney Images}}$$

$$\text{True Negative(TN)} = \frac{\text{Number of Images without kidney stone}}{\text{Total number of kidney Images}}$$

$$\text{False Positive(FP)} = \frac{\text{Number of Images falsely detected as kidney stone}}{\text{Total Number of kidney Images}}$$

$$\text{False Negative(FN)} = \frac{\text{Number of Images having kidney stone but not detected}}{\text{Total Number of kidney Images}}$$

$$\text{False Negative(FN)} = \frac{\text{Number of Images having kidney stone but not detected}}{\text{Total Number of kidney Images}}$$

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{F1 Score} = \frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})}$$

Table IV depicts the classifier category of clinical and laboratory samples using different data mining techniques. From Table IV it is observed that, the KNN algorithm yields better results of 100% precision, accuracy against other classifier for clinical sample or kidney US images. Subsequently increase in accuracy precision, Recall and F-1 Score are also noticed in below table. Likewise, for laboratory samples, RF algorithm reached maximum accuracy of 100% along with other performance parameters against different data mining techniques.

TABLE IV. PERFORMANCE EVALUATION FOR CLINICAL AND LABORATORY DATA SAMPLES

Performance Evaluation (%)				
Clinical data sets- US images				
Classifier	Precision	Recall	F1 Score	Accuracy
Support Vector Machine	0.77	0.62	0.58	0.62
K Nearest Neighbour	1.0	1.0	1.0	1.0
Multi Layer Perceptron	0.58	0.57	0.56	0.57
Random Forest	0.88	0.87	0.87	0.87
Laboratory data sets- Blood sample				
Support Vector Machine	0.61	0.61	0.57	0.61
K Nearest Neighbour	0.88	0.88	0.88	0.88
Multi Layer Perceptron	0.83	0.71	0.70	0.71
Random Forest	1.0	1.0	1.0	1.0

Furthermore, the experimental results obtained from the work are compared with the results obtained from similar works and are tabulated as depicted in Table V. By optimal selection of nose variance value from the results obtained, it is observed that the machine learning algorithm KNN when applied to US images reaches a maximum accuracy, whereas the Random Forest algorithm yields an accuracy of 87%. Extending further, for the laboratory samples, the experimental results obtained for Random forest with that of enhanced decision tree algorithm, with KNN reaching an accuracy of 88%. Of all the comparative analysis carried out, it is evident from the experimental results that KNN best suits for detecting the Kidney stone using US technique.

TABLE V. COMPARATIVE ANALYSIS OF DATA MINING METHODS FOR CKD

Ref.No.	Kidney category	Feature Extraction	Classifier	Accuracy (%)
A. Clinical Data samples				
[30]	Normal, abnormal	PCA	SVM, KNN	84, 89
[21]	Normal, abnormal	Gradient features, GLCM	SVM	85.8±3.1
[31]	Normal, abnormal	Statistical features GLCM	SVM	85.7
[32]	Normal, abnormal	GLCM	ANN	87.5
[33]	Normal, abnormal	Gobar features	ANN	87.06
[34]	Normal, abnormal	SVD Features	SVM	90
[13]	Normal, abnormal, bacterial infection, cystic diseases	GLCM and intensity histogram	ANN	-
[35]	Normal, abnormal	Statistical wavelet based features+ PCA	ANN	97
[36]	Normal, abnormal	GLCM+GLRLM+ Differential Evolution (DE)	SVM	86.3
[37]	Normal, abnormal	Statistical features	SVM	84
[38]	Normal, abnormal	Gabor wavelet features	ANN	-
[39]	normal, kidney stone, cystic, kidney tumor	GLCM+PCA	ANN	77.8
[40]	Normal, abnormal	GLCM+PCA	SVM,KNN	84,89
[41]	Normal, abnormal	K-Means and GLCM features	Meta-Heuristic SVM	98.8
[42]	Normal, abnormal	wavelet energy features	ANN	96.8
This work	Normal, abnormal	GLCM+GLRLM	KNN, RF	100,87.0
B. Laboratory data				
[5]	Normal, abnormal	Physiological and Clinical Attributes	PNN, RBF, SVM,MLP	96.7,87 60.7, 51.5
[43]	Normal, abnormal	Physiological and Clinical Attributes	Ant Colony based Optimization (D-ACO)	95
[44]	Normal, abnormal	Physiological and Clinical Attributes	Decision Tress.	99.5
[45]	Normal, abnormal	Physiological and Clinical Attributes	Enhanced Decision Tree	100
[46]	Normal, abnormal	Physiological and Clinical Attributes	SVM,KNN	97.75, 98.5
This work	Normal, abnormal	Physiological and Clinical Attributes	RF,KNN	100,88.0

V. CONCLUSION

The main focus of this work is to find the optimal data mining algorithms for both clinical and laboratory data sets. This work also characterizes the NV parameter varying from 0.01 to 0.08 against the statistical parameters like SNR, PSNR, RME and MAE for different preprocessing filters like spatial and wavelet filters for US images, and find the best filter can be used for preprocessing the US Images. From the experimental results it is evident that, sym8 decomposition level 4, provide increased SNR (29.87dB), PSNR (42.21dB), with reduction in RME (1.97) and MAE (1.3921) as compared other filtering methods. Further, morphological operations like erosion and dilation were applied to segment the filtered US image. By applying entropy based segmentation and morphological operations, feature extraction, RoI and exact area of kidney stone can be located for kidney stone US image. In classifying clinical and laboratory data sets different data mining techniques such as SVM, KNN, MLP and RF are used. The RF and KNN reaches good accuracy up to 100% for laboratory and clinical data sets against other techniques discussed above.

The notable limitation of this work is direct comparison of experimental result obtained with result obtained from similar work is beyond scope of this work, as kidney US Images obtained vary between hospitals. To author knowledge this is the first work to use NV from 0.01 to 0.08 against statistical parameters. Further increase in noise value may not directly improve the PSNR, SNR or decreases in RME and MAE.

This work recommends the use of KNN for kidney US Images and RF for laboratory samples that can be used for physicians in order to eliminate diagnostic and treatment errors. To assist the medical doctor for further treatment, a GUI was developed, that helps in detection of kidney stone and its area with minimum effort.

REFERENCES

- [1] William G Couser, Giuseppe Remuzzi, Shanthi Mendis, Marcello Tonelli, "The contribution of chronic kidney disease to the global burden of major non communicable diseases", *Kidney Int*, 2011, , vol. 80, no. 12, December 2011, DOI: 10.1038/ki.2011.368.
- [2] Vivekanand Jha, Guillermo Garcia-Garcia, Kunitoshi Iseki, Zuo Li, Saraladevi Naicker, Brett Plattner, Rajiv Saran, Angela Yee-Moon Wang, Chih-Wei Yang "Chronic kidney disease: global dimension and perspectives", *Lancet*, Vol 382, pp.260-273, July 2013.
- [3] Andrew S Levey, Josef Coresh, Ethan Balk, Annamaria T Kausz, Adeera Levin, Michael W Steffes, Ronald J Hogg, Ronald D Perrone, Joseph Lau, Garabed Eknoyan, National Kidney Foundation, "National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification", *Annals of internal medicine*, Vol.139, No.2, pp. 137-147, July 2003.
- [4] Dua, D. and Graff, C. (2019). "UCI Machine Learning Repository" [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
- [5] El-Houssainy A.Rady, Ayman S.Anwar, "Prediction of kidney disease stages using data mining algorithms", *Elsevier Journal, Informatics in Medicine Unlocked*: Vol. 15, pp.1-7, April 2019.
- [6] R.F. Wagner, S.W. Smith, J.M. Sandrik, H. Lopez, "Statistics of Speckle in Ultrasound B-Scans", *IEEE Transactions on Sonics and Ultrasonics* Vol 30, No.3, pp.156-163, May 1983.
- [7] Kunio Doi, Ph.D., "Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential", *Comput Med Imaging Graph.*, Vol 31, Issue (4-5), pp.198-211, 2007.
- [8] J. Portilla; V. Strela; M.J. Wainwright; E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain" *IEEE Transactions on Image Processing*, Vol 12, Issue 11, pp-1338-1351, 2003,
- [9] Carlos A. Duarte-Salazar, Andrés Eduardo Castro Ospina, Miguel A. Becerra, Edilson Delgado-Trejos, "Speckle Noise Reduction in Ultrasound Images for Improving the Metrological Evaluation of Biomedical Applications: An Overview" Vol 8, pp 15983 – 15999, 2020.
- [10] Tanzila Rahman, Mohammad Shorif Uddin, "Speckle noise reduction and segmentation of kidney regions from ultrasound image", 2013 International Conference on Informatics, Electronics and Vision (ICIEV), DOI:10.1109/ICIEV.2013.6572601.
- [11] D.Gnanadurai, V.Sadasivam, "An Efficient Adaptive Thresholding Technique for Wavelet Based Image Denoising", Published 2006, World Academy of Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, Vol 2, No 8, pp 1703-1708, 2008.
- [12] Quan Pan, Lei Zhang, Guanzhong Dai, and Hongcai Zhang, "Two Denoising Methods by Wavelet Transform", *IEEE transactions on signal processing*, Vol. 47, No. 12, pp 3401-3406, December 1999.
- [13] Hafizah, Wan Mahani and Supriyanto, Eko and Yunus, Jasmy, "Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix", *In Proceedings - 6th Asia International Conference on Mathematical Modelling and Computer Simulation*, IEEE, New York, USA, pp. 115-120, 2012.
- [14] Wan Mahani Hafizah, Eko Supriyanto, "Automatic Region of Interest Generation for Kidney Ultrasound Images", *International Journal of Biology and Biomedical Engineering*, Vol 6, Issue 1, pp 26-34, Oct 2012.
- [15] K. Viswanath R. Gunasundari, "Modified Distance Regularized Level Set Segmentation Based Analysis for Kidney Stone Detection", *International Journal of Rough Sets and Data Analysis (IJRSDA)*, Vol 2, Issue 2, pp 18, 2015.
- [16] Yan Chen, Ruming Yi, Patrick Flynn, Shira Broschat, "Aggressive region growing for speckle reduction in ultrasound images", *Pattern Recognition Letters Elsevier*, Vol 24, Issue 4-5, pp 677-691, 2003.
- [17] Prema T. Akkasaligar, Sunanda Biradar, "Classification of Medical Ultrasound Images of Kidney", *IJCA Proceedings on International Conference on Information and Communication Technologies*, 2014.
- [18] Jing Xiao, Ruifeng Ding, Xiulin Xu, Haochen Guan, Xinhui Feng, Tao Sun, Sibozhu, Zhibin Ye, "Comparison and development of machine learning tools in the prediction of chronic kidney disease progression", *Journal of Translational Medicine*, Vol 17, pp 119, 2019.
- [19] Amit GARG, Vineet Khandelwal, "Speckle Noise Reduction in Medical Ultrasound Images Using Modelling of Shearlet Coefficients as a Nakagami Prior", *Advances in Electrical and Electronic Engineering*, Vol 16, No4, pp-538-549, 2018.
- [20] Ramandeep Kaur; Akshay Girdhar; Jappreet Kaur, "A New Thresholding Technique for Despeckling of Medical Ultrasound Images", *Fourth International Conference on Advances in Computing and Communications IEEE*, 2014.
- [21] M. B. Subramanya, Vinod Kumar, Shaktidev Mukherjee & Manju Saini, "SVM-Based CAC System for B-Mode Kidney Ultrasound Images", *Journal of Digital Imaging, Elsevier*, pp 448-458, Dec 2015.
- [22] A.D. Hillery and R.T. Chin, "Iterative Wiener filters for image restoration", *IEEE Transactions on Signal Processing*, Vol 39, No. 8, pp. 1892 – 1899, 1991.
- [23] Zhou qin-wu et.al,(2002) "Denoise and contrast enhancement of ultrasound speckle image based on wavelet", *6th International Conference on Signal Processing*, Beijing, China, 2002.
- [24] Q.Pan, "Two denoising methods by wavelet transform", *IEEE Transactions on Signal Processing*, Vol. 47, pp.3401-3406, 1999.
- [25] T. Loupas, W.N. McDicken, P.L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images", *IEEE Transactions on Circuits and Systems*, Vol. 36, No.1, pp-129-135, January 1989.
- [26] Pallavi Vaish, R Bharath, P Rajalakshmi, U. B. Desai, "Smartphone Based Automatic Abnormality Detection of Kidney in Ultrasound

- Images”, IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom),2016.
- [27] Chih-Wei Hsu, Chih-Jen Lin “A comparison of methods for multiclass support vector machines”, *IEEE transactions on Neural Networks*, Vol. 13, No. 2, pp. 415-425, March 2002.
- [28] Vinoth R, Anusha C, SaiPriya D, Sridhar BS, ManiKantaPavan C, Aditya G. “A Novel Approach on Ultrasound Kidney Images Classification by using Micro Deep Neural Network Algorithm”, *International Journal of pure and applied Mathematics*, Vol. 119, No. 12, pp.15047-15058., 2018.
- [29] Angshuman Paul, Dipti Prasad Mukherjee, Prasun Das, Abhinandan Gangopadhyay, Appa Rao Chintha, Saurabh Kundu, “Improved Random Forest for Classification”, *IEEE Transactions on Image Processing*, Vol.27, No.8, pp. 4012 – 4024, Aug. 2018.
- [30] J. Verma, M. Nath, P. Tripathi, and K. Saini, “Analysis and identification of kidney stone using kth nearest neighbour (KNN) and support vector machine (SVM) classification techniques,” *Pattern Recognition and Image Analysis*, vol. 27, no. 3, pp. 574–580, 2017.
- [31] K. Sharma and J. Virmani, “A decision support system for classification of normal and medical renal disease using ultrasound images: A decision support system for medical renal diseases,” *International Journal of Ambient Computing and Intelligence*, vol. 8, no. 2, pp. 52–69, 2017.
- [32] M. Pawar and A. Mulla, “Design and analysis performance of kidney cyst detection from ultrasound images,” *International Journal of Engineering Development and Research*, vol. 5, no. 4, pp. 911–917, 2017.
- [33] S. Chaitanya and A. Rajesh kumar, “Classification of kidney images using particle swarm optimization algorithm and artificial neural networks,” *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 4, pp. 526–530, 2019.
- [34] S. Sudharson and P. Kokil, "Abnormality classification in the kidney ultrasound images using singular value decomposition features," 2019 IEEE Conference on Information and Communication Technology, 2019, pp. 1-5.
- [35] Attia MW, AbouChadi FE, Moustafa HE, Mekky N. Classification of ultrasound kidney images using PCA and neural networks. *International Journal of Advanced Computer Science and Applications*. 2015 Apr 1; 6(4):537.
- [36] Subramanya MB, Kumar V, Mukherjee S, Saini M. “SVMbased CAC system for Bmode kidney ultrasound images”, *Journal of digital imaging*. 2015 Aug 1; 28(4):44858.
- [37] Rana S, Jain S, Virmani J, “SVM Based characterization of focal kidney lesions from bmode ultrasound images”, *Research Journal Of Pharmaceutical Biological And Chemical Sciences*. 2016 Jul 1;7(4):83746.
- [38] Raja KB, Madheswaran M, Thyagarajah K.,” Texture pattern analysis of kidney tissues for disorder identification and classification using dominant Gabor wavelet. *Machine Vision and Applications*”, 2010 Apr 1; 21(3):287300.
- [39] Priyanka, Dr. Dharmender Kumar, “ Feature extraction and selection of kidney Ultra Sound Images using GLCM and PCA”, *International Conference on Computational Intelligence and Data Science*, *Procedia Computer Science* 167 (2020) 1722–1731.
- [40] Jyoti Verma, Madhwendra Nath, Priyanshu Tripathi, and K. K. Saini, “Analysis and Identification of Kidney Stone Using Kth Nearest Neighbour (KNN) and Support Vector Machine (SVM) Classification Techniques” , *Pattern Recognition and Image Analysis*, Vol. 27, No. 3, pp. 574–580 (2017).
- [41] S. Selvaran & P. Rajendran, “ Detection of Renal Calculi in Ultrasound Image Using Meta-Heuristic Support Vector Machine” *Journal of Medical Systems* 43: 300, (2019).
- [42] Akkasaligar, P. T., and Biradar, S., “ Diagnosis of Renal Calculus Disease in Medical Ultrasound Images” . *IEEE International Conference on Computational Intelligence and Computing Research*, 978-1-5090-0612-0/16 (2016).
- [43] Elhoseny, M., Shankar, K. & Uthayakumar, J. “ Intelligent Diagnostic Prediction and Classification System for Chronic Kidney Disease”, *Sci Rep* 9, 9583 (2019). <https://doi.org/10.1038/s41598-019-46074-2>.
- [44] I.A. Pasadana et al , “ Chronic Kidney Disease Prediction by Using Different Decision Tree Techniques”,2019 *J. Phys.: Conf. Ser.* 1255 012024.
- [45] Chaudhuri, A.K., Sinha, D., Banerjee, D.K. et al. “ A novel enhanced decision tree model for detecting chronic kidney disease” ,*Netw Model Anal Health Inform Bioinforma* 10, 29 (2021).
- [46] Alaiad A, Najadat H, Mohsen B, Balhaf K ,” Classification and association rule mining technique for predicting chronic kidney disease.”, *J Inf Knowl Manag* , 2019(01):2040015.

Comparing the Balanced Accuracy of Deep Neural Network and Machine Learning for Predicting the Depressive Disorder of Multicultural Youth

Haewon Byeon

Department of Medical Big Data
College of AI Convergence, Inje University, Gimhae 50834, Republic of Korea

Abstract—Multicultural youth are more likely to experience negative emotions (e.g. depressive symptoms) due to social prejudice and discrimination. Nevertheless, previous studies that analyzed the emotional aspects of multicultural youth mainly compared the characteristics of multicultural youth and those of the other youth or identified individual risk factors using a regression model. This study developed models to predict the depressive disorders of multicultural youth based on the Quick Unbiased Efficient Statistical Tree (QUEST), Classification And Regression Trees (CART), gradient boosting machine (G-B-M), random forest, and deep neural network (deep-NN) using epidemiological data representing multicultural youth and compared the prediction performance (PRED PER) of the developed models. Our study analyzed 19,431 youths (9,835 males and 9,596 females) aged between 19 and 24 years old. We developed models for predicting the self-awareness of health of youths by using QUEST, CART, G-B-M, random forest, and deep-NN and compared the balanced accuracy of them to evaluate their PRED PER. Among 19,431 subjects, 42.9% (5,838 people) experienced a depressive disorder in the past year. The results of our study confirmed that deep-NN had the best PRED PER with a specificity of 0.85, a sensitivity of 0.71, and a balanced accuracy of 0.78. It will be necessary to develop a model with optimal PRED PER by tuning hyperparameters (e.g., number of hidden layers, number of iterations, and activation function, number of hidden nodes) of deep-NN.

Keywords—Quick unbiased efficient statistical tree; gradient boosting machine; deep neural network; classification and regression trees; balanced accuracy

I. INTRODUCTION

South Korea has shifted to a multicultural society at a very rapid pace over the past 30 years. As a result, the number of multicultural youth has been increasing rapidly. The Ministry of Education (2019)[1] reported that the number of multicultural youth students was 137,225 people (2.5% of all students) in 2019, which is a 200% increase // an increase by 200% from 67,806 people in 2014. Particularly, the total fertility rate (TFR) decreased to 0.918 as of 2019 [2] due to the low fertility trend in Korean society, and the proportion of children with multicultural backgrounds in the school-age population is predicted to increase continuously [1].

Nevertheless, most of the policy interests targeting multicultural families (MUC-fam) in Republic of Korea have emphasized employment and welfare aspects, and there are not enough surveys on the mental health of MUC-fam [3,4]. In

addition, studies conducted on multicultural youth have mainly focused on academic achievement or school adaptation, and emotional characteristics such as depression have been studied rarely [5].

Multicultural youth are more likely to experience negative emotions (e.g. depressive symptoms) due to social prejudice and discrimination [6]. Moreover, a national survey (epidemiological survey) [7] reported that the youth in low socioeconomic status such as low household income experienced a depressive disorder more. It was also implied that the multicultural youth were more likely to be exposed to negative emotionality since the proportion of MUC-fam in rural areas was higher than that in cities and only 9.7% of multicultural households made KRW 30 million or more per year. Above all, Byeon (2014)[8] revealed that approximately 15% of multicultural youth aged between 19 and 24 years old experienced social discrimination and children of international marriage families born in South Korea and immigrant children of international marriage families had difficulties in social adaptation due to the social characteristics of MUC-fam and rapid changes in adolescence. Since multicultural youth are highly likely to have multiple factors associated with a depressive disorder (e.g., sociodemographic factors and personal characteristics), they are believed to experience a depressive disorder more than the other youth. Nevertheless, previous studies [9,10,11,12] that analyzed the emotional aspects of multicultural youth mainly compared the characteristics of multicultural youth and those of the other youth or identified individual risk factors using a regression model.

It is necessary to develop a prediction model based on big data to grasp the characteristics of a depressive disorder because emotional issues are induced by multidimensional factors such as stress, social support, and environment. In recent years, machine learning (Mach Learn) has been widely used in various fields as a means to overcome the limitations of the Mach Learn regression model. This study developed models to predict the depressive disorders of multicultural youth based on quick unbiased efficient statistical tree (QUEST), classification and regression trees (CART), gradient boosting machine (G-B-M), random forest, and deep neural network (deep-NN) using epidemiological data representing multicultural youth and compared the prediction performance (PRED PER) of the developed models.

II. METHODS AND MATERIALS

A. Source of Data

This study analyzed the raw data of the National Survey of MUC-fam on MUC-fam residing in South Korea in 2012 jointly conducted by the Ministry of Health, Welfare and Family Affairs (MHWFA), the Ministry of Justice (MJ), and the Ministry of Gender Equality (MGE). The National Survey of MUC-fam was carried out to grasp the living conditions and welfare needs of MUC-fam to develop policies customized for MUC-fam [2]. This national survey consists of economic level, general characteristics, marital life, employment, health and health care. The National Survey of MUC-fam was conducted July 20 to October 31, 2012, and targeted all 154,333 marriage immigrants by using the status of foreign residents and the basic status of MUC-fam by the Ministry of Public Administration and Security (MPAS). The selection criteria for MUC-fam were based on the MUC-fam Support Act: this study targeted (1) families composed of marriage immigrants and South Korean citizens and (2) those composed of foreigners who obtained South Korean citizenship by acknowledgment or naturalization and South Korean citizens who obtained Republic of Korean citizenship by naturalization, birth or acknowledgment. This study analyzed 19,431 youths aged between 19 and 24 years old (9,835 males and 9,596 females) among the children of marriage immigrants (MI).

B. Measurement of Variables

A depressive disorder, the outcome variable, was divided into “experienced a depressive disorder” and “did not experience a depressive disorder” based on the question, “Have you ever felt sad or despair that lasted two weeks or longer in a row in the past year?”, of the Diagnostic and Statistical Manual of Mental Disorders-five [13], a diagnostic criterion for major depressive disorder. Explanatory variables included gender, education, residence area (rural or urban), economic activity (yes or no), social discrimination experience (yes or no), career counseling experience (yes or no), Korean language education experience (yes or no), the experience of using a support center for multi-cultural families (yes or no), level of Korean reading (good, average, or poor), level of Korean writing (good, average, or poor), level of Korean listening, level of Korean speaking, and experience of Korea society adaptation training (yes or no) by referring to previous studies [9,10,11,12] related to a depressive disorder.

C. Developing Models for Predicting a Depressive Disorder:

QUEST

QUEST [13] selects a significant variable among variables entered first, performs a quadratic discriminant analysis based on the selected variable, and selects a predictor for reducing the variable selection bias. The variable selection is made (1) by ANOVA F-statistics for continuous variables and (2) by choosing the various with the smallest probability as a classification variable for categorical variables using the chi-square test. It is characterized by selecting a threshold by finally conducting quadratic discriminant analysis for the classification variable selected in this process [14].

The QUEST algorithm for predicting and classifying variables is performed by the following procedure [15]. First,

the significance probability (p-value) of an ordinal or continuous predictor is calculated by the F-test of ANOVA. Second, when a categorical predictor is selected, the p-value of the cross-validation is calculated in the contingency table of predictor and target variables. If the p-value is smaller than the adjusted Bonferroni's threshold, the corresponding variable is selected as a classification variable. If there are more than 3 categories, it is replaced with a binary classification based on two-means clustering. Third, if a separation variable cannot be selected by the above two procedures, the p-value of the Levene F-test is calculated for an ordinal or continuous classification variable, and it is compared with the adjusted Bonferroni's threshold to finally select the variable corresponding to the p-value as the classification variable. This procedure is repeated until the condition is fully satisfied. For data composed of categorical variables, if the categories of outcome variables are three or more, the two-mean clustering consisting of two categories is performed. At this time, quadratic discriminant analysis is used to find the optimal threshold of explanatory variables, and child nodes are formed by finding explanatory variables that classify the outcome variables best.

D. CART

CART measures impurity using the Gini Index (Gini impurity) and performs a binary split in which only two child nodes are formed from a parent node [16]. The Gini impurity refers to the probability that two elements randomly extracted from n elements belong to different groups. When the decrement of the Gini Index is calculated, the predictor and the optimal threshold that decreases the Gini Index most are selected as child nodes, as the final step.

E. G-B-M

G-B-M is a Mach Learn algorithm that creates a strong learner (S-learner) by combining weak learners (W-learner) of decision trees (Decis Tree) by using an ensemble technique [17]. This method generalizes models by generating models for each step and optimizing the loss function (loss-func) that can be arbitrarily differentiated, like other boosting methods. Even if the accuracy is low, a model is created, and the error of the generated model is complemented by the next model. Through this process, more powerful learner), than the previous a learner, is created. The basic principle is to increase accuracy by repeating this process. The concept of G-B-M is presented in Fig. 1.

Algorithm	Gradient boosting algorithm.
Input:	Input data $(x, y)_{i=1}^N$ Number of iterations M Choice of the loss-function $\Psi(y, f)$ Choice of the base-learner model $h(x, \theta)$
1:	Initialize \hat{f}_0 with a constant
2:	for $t = 1$ to M do
3:	Compute the negative gradient $g_t(x)$;
4:	Fit a new base-learner function $h(x, \theta_t)$;
5:	Find the best gradient descent step-size p_t
	$p_t = \operatorname{argmin}_p \sum_{i=1}^N \Psi[y_i, \hat{f}_{t-1}(x_i) + p h(x_i, \theta_t)];$
6:	Update the function estimate:
	$\hat{f}_t \leftarrow \hat{f}_{t-1} + p_t h(x, \theta_t);$
7:	end for

Fig. 1. The Concept of G-B-M.

F. Random Forest

Random forest is a decision tree-based ensemble method that generates a large number of random samples (Rand Sam) from a training dataset through bootstrapping (random sampling with replacement of the same size from a given data), learns an independent Decis Tree for each sample (Eac Sam) set, and creates a final model by synthesizing the results. The structure of it is presented in Fig. 2.

G. Deep-NN

Deep-NN is an algorithm composed of an composed of independent variables, an composed of dependent variables, and two or more hidden layers between them. Independent node (Ind node)s are arranged in each layer of the input layer, the hidden layer, and the output layer. A group of nodes in each layer is connected by weighted neurons (connecting lines) (Fig. 3).

This study used H2O's deep-NN among various deep learning types. H2O's deep-NN is based on a multi-layer feedforward artificial neural network (A-NN) that is trained with stochastic gradient descent (Grad Des) using back-propagation (Back-propag). This study set two hidden layers with ten nodes in each layer (20 nodes in total) and five epochs (number of iterations). This study used the Rectifier Linear Unit (ReLU), default value, as the activation function of deep learning (Fig. 4).

H. Validation of the Models

This study developed models for predicting the subjective health of youth by using QUEST, CART, G-B-M, random forest, and H2O's deep-NN and compared the balanced accuracy of them to evaluate their PRED PER. Balanced accuracy is an index that considers sensitivity and specificity and presents the same value as the area under the curve (AUC), indicating the accuracy of a classification model. This study used 10-fold cross-validation (Cro-Valid) to validate the developed models.

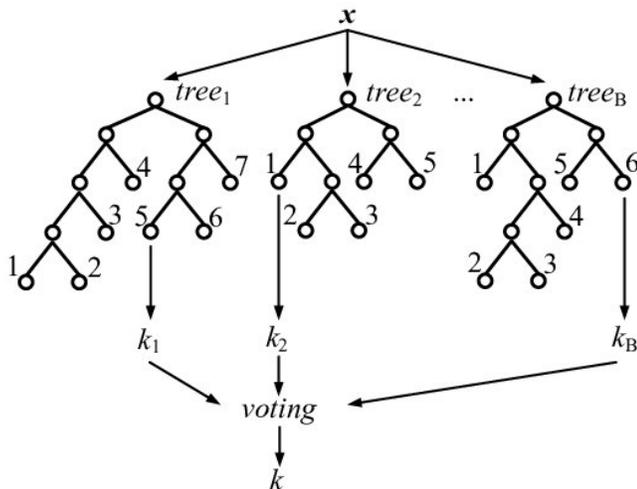


Fig. 2. Structure of Random Forest Model [18].

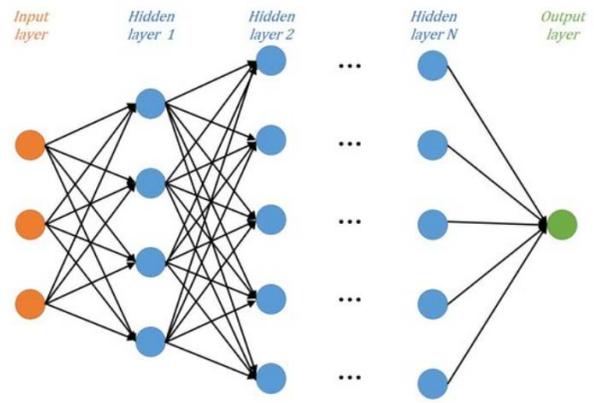


Fig. 3. Layers of Deep-NN [19].

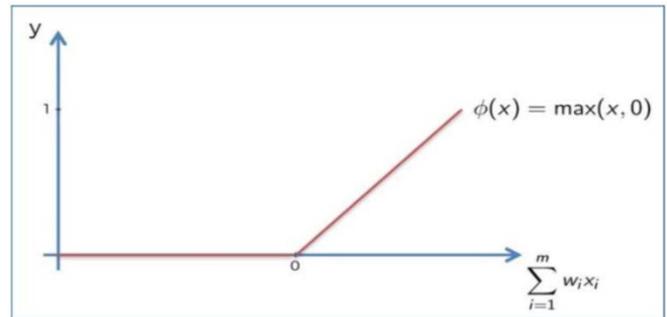


Fig. 4. The Concept of Rectified Linear Unit Activation Function [20].

In this study, a model containing randomness, such as random forest, was developed while fixing the seed to “435435”. This study defined the model with the highest accuracy as the model with the best PRED PER by comparing the PRED PER of these models. When the accuracy of multiple models was the same, the model with high sensitivity was defined as the model with the best PRED PER. All analyses were performed using R version 4.0.4 (Foundation for Statistical Computing, Vienna, Austria).

III. RESULTS

The general characteristics of subjects by depression experience are presented in Table I. Among 19,431 subjects, 42.9% (5,838 people) experienced a depressive disorder in the past year. The results of chi-square test showed the highest level of education, gender, economic activity, Korean language education experience, level of Korean speaking, Korean writing level, level of Korean listening, Korean reading level, Korea society adaptation training experience, career counseling experience, social discrimination experience, and experience of using a support center for MUC-fams were significantly ($p < 0.05$) different between multicultural youth with a depressive disorder experience and those without a depressive disorder experience.

TABLE I. CHARACTERISTICS OF SUBJECTS BASED ON EXPERIENCE OF DEPRESSIVE DISORDER, N (%)

Variables	Depressive disorder		p
	Yes (n=5,838)	No (n=13,593)	
Residence area			0.308
Rural	4,970 (30.2)	11,492 (69.8)	
Urban	869 (29.3)	2,101 (70.7)	
Level of education			<0.001
Elementary school or below	0	202 (100)	
Junior high school	725 (29.2)	1,754 (70.8)	
High school	4,376 (30.8)	9,812 (69.2)	
College or higher	738 (28.8)	1,825 (71.2)	
Gender			
Male	2,739 (27.8)	7,097 (72.2)	
Female	3,099 (32.3)	6,496 (67.7)	
Economic activity			0.001
Yes	2,311 (28.7)	5,732 (71.3)	
No	3,527 (31.0)	7,861 (69.0)	
Korean language education experience			<0.001
No	5,460 (30.6)	12,402 (69.4)	
Yes	378 (24.1)	1,191 (75.9)	
Korean writing level			0.012
Good	4,392 (29.7)	10,372 (70.3)	
Average	798 (29.5)	1,903 (70.5)	
Poor	648 (33.0)	1,318 (67.0)	
Korean reading level			<0.001
Good	4493 (29.7)	10,640 (70.3)	
Average	720 (28.5)	1,809 (71.5)	
Poor	625 (35.3)	1,144 (64.7)	
Korean speaking level			0.005
Good	4,456 (29.5)	10,659 (70.5)	
Average	941 (31.7)	2,024 (68.3)	
Poor	441 (32.6)	911 (67.4)	
Korean listening level			<0.001
Good	4,459 (29.0)	10,925 (71.0)	
Average	984 (35.1)	1,821 (64.9)	
Poor	395 (31.8)	847 (68.2)	
The experience of using a support center for MUC-fams			<0.001
No	5,641(30.7)	12,706 (69.3)	
Yes	197 (18.2)	887 (81.8)	
Korea society adaptation training experience			<0.001
Yes	1,810 (33.9)	3,524 (66.1)	
No	4,028 (28.6)	10,069 (71.4)	
Social discrimination experience			<0.001
No	4,196 (25.6)	12,200 (74.4)	
Yes	1,642 (54.1)	1,393 (45.9)	
Career counseling experience			<0.001
No	5,273 (29.0)	12,941 (71.0)	
Yes	565 (46.4)	653 (53.6)	

The balanced accuracy of five models (QUEST, CART, G-B-M, random forest, and H2O's deep-NN) for predicting a depressive disorder of multicultural youth is presented in Fig. 5. The results of our study confirmed that H2O's deep-NN had the best PRED PER with a specificity of 0.85, a sensitivity of 0.71, and a balanced accuracy of 0.78. The normalized importance of variables was analyzed using H2O's deep-NN. The major predictors of multicultural youth's depressive disorder were social discrimination experience, gender, level of Korean speaking, Korean writing level, and level of Korean listening (Fig. 6). Among them, social discrimination experience was the most important factor in predicting a depressive disorder (Fig. 6).

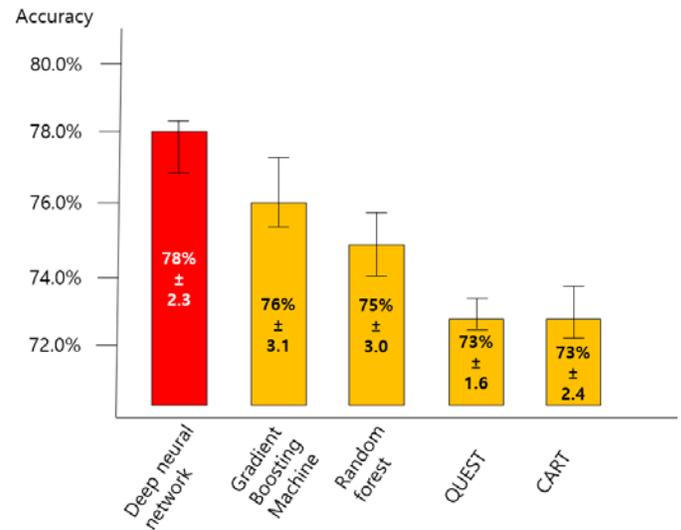


Fig. 5. Comparison of Balanced Accuracy of 5 Models for Predicting Depression, (%).

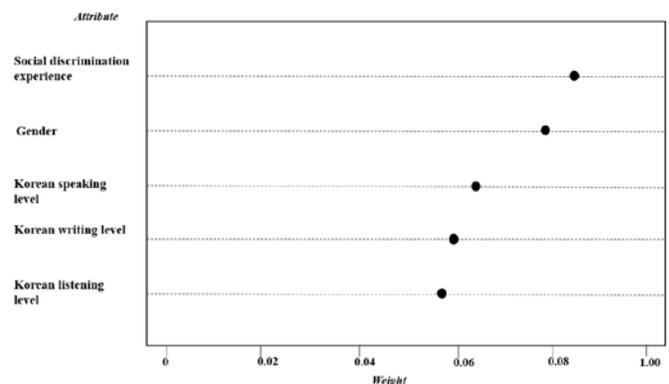


Fig. 6. Importance of Variables (Impor Var) in the Depression.

IV. CONCLUSION

Our study compared the balanced accuracy of models for predicting the depressive disorder of multicultural youth and confirmed that H2O's deep-NN was the model with the best PRED PER among QUEST, CART, G-B-M, random forest, and H2O's deep-NN. On the other hand, tree-based Mach Learns such as QUEST and CART had relatively low balanced accuracy compared to other Mach Learn methods. These results agreed with da Krauss et al. [21], which showed that the

prediction accuracy of deep-NN was better than gradient-boosted-trees or random forest. On the other hand, Grolinger et al. [22] reported that H2O's deep-NN had worse PRED PER than support vector regression and took longer computation time. These results imply that the most optimal algorithm with the best PRED PER may vary depending on the data type. Therefore, more follow-up studies are needed to compare the performance of Mach Learn and deep-NNs using various datasets with different characteristics.

Currently, various open-source platforms (e.g., Torch, MXNet, TensorFlow, Caffe, Theano, and H2O) have been developed for deep learning. Among them, the H2O platform provides access to Mach Learn algorithms through common development environments (e.g., Python, R, and JAVA), big data systems (e.g., Hadoop and Spark), and various data sources (e.g., SQL, NoSQL, HDFS, and S3). This study developed a model to predict the depressive disorder of multicultural youth using the defaults of the H2O platform (i.e., number of hidden layers, number of iterations, number of hidden nodes, and activation function). It was confirmed that the balanced accuracy of predict the depression was superior to that of other Mach Learn algorithms. It will be necessary to develop a model with optimal PRED PER by tuning hyper-parameters (e.g., number of hidden layers, number of iterations, number of hidden nodes, and activation function) of H2O's deep-NN.

Our study developed a model for predicting the depressive disorder experience of multicultural youth using H2O's deep-NN and found that social discrimination experience, gender, level of Korean speaking, Korean writing level, and level of Korean listening were related to the depressive disorder of multicultural youth. Among them, social discrimination experience was the most influential factor in predicting a depressive disorder. Consequently, it is required to prepare a legal system that can help multicultural youth overcome discrimination and prejudice and give attention to them at the social level.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07041091, 2021S1A5A8062526).

REFERENCES

- [1] Ministry of Education, 2019 Basic Education Statistics. Ministry of Education, Sejong, 2019.
- [2] National Statistical Office, Population Trend Survey. National Statistical Office, Daejeon, 2021.
- [3] H. Kim, and J. G. Lee, Multicultural society: examining Korea and China's policies for ethnic minorities and immigrants. *The Social Science Research Institute*, vol. 12, no. 1, pp. 37-67, 2019.
- [4] D. H. Lee, A study on multicultural families support act. *The Journal of the Korea Contents Association*, vol. 19, no. 7, pp. 650-658, 2019. doi: 10.5392/JKCA.2019.19.07.650.
- [5] J. Y. Oh, Role of multicultural family support center for school adaptation of elementary school children of multicultural families. *Cultural Exchange and Multicultural Education*, vol. 7, no. 3, pp. 79-99, 2018.

- [6] K. H. Han, Predictors of depressive symptom trajectories among multicultural family adolescents. *Mental Health & Social Work*, vol. 48, no. 1, pp. 56-83, 2020. doi: 10.24301/MHSW.2020.03.48.1.56.
- [7] Ministry of Health & Welfare, Korean youth obesity survey. Ministry of Health & Welfare, Seoul, 2009.
- [8] H. Byeon, The factors that affects the experience of discrimination in children in multi-cultural families using QUEST algorithm : focusing on Korean language education. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, vol. 4, no. 2, pp.303-312, 2014. doi: 10.14257/AJMAHS.2014.12.12.
- [9] M. K. Park, Y. Cha, H. Lee, The effects of using the multicultural family support centers on the bicultural identity of the multicultural youths: focusing on propensity score matching. *The Journal of Multicultural Society*, vol. 12, no. 3, pp. 107-140, 2019. doi: 10.14431/jms.2019.10.12.3.107.
- [10] R. H. Lee, Daily stress and smartphone overdependence among multicultural adolescents: comparisons by gender and age. *Journal of Digital Convergence*, vol. 16, no. 11, pp. 561-569, 2018. doi: 10.14400/JDC.2018.16.11.561.
- [11] B. Seo, Analysis of factors and actual status of delinquency among multi-cultural youth In Korea. *Contemporary Society and Multiculture*, vol. 9, no. 2, pp. 123-150, 2019.
- [12] K. H. Han, and S. K. Kahng, The relationship among acculturative stress, depression and self-esteem of multicultural adolescents-testing a theoretical model of the stress process. *Mental Health & Social Work*, vol. 47, no. 1, pp. 231-257, 2019. doi: 10.24301/MHSW.2019.03.47.1.231.
- [13] J. C. Tolentino, and S. L. Schmidt, DSM-5 criteria and depression severity: implications for clinical practice. *Frontiers in Psychiatry*, vol. 9, pp. 450, 2018. doi: 10.3389/fpsy.2018.00450.
- [14] C. L. Lin, and C. L. Fan, Evaluation of CART, CHAID, and QUEST algorithms: a case study of construction defects in Taiwan. *Journal of Asian Architecture and Building Engineering*, vol. 18, no. 6, pp. 539-553, 2019. doi: 10.1080/13467581.2019.1696203.
- [15] H. Byeon, and O. Lee, Development of prediction model for life satisfaction among Korean elderly using QUEST algorithm: findings from a community-based study. *International Journal of Applied Engineering Research*, vol. 10, no. 79, pp. 88-92, 2019.
- [16] M. M. Ghiasi, S. Zendejboudi, and A. A. Mohsenipour, Decision tree-based diagnosis of coronary artery disease: CART model. *Computer Methods and Programs in Biomedicine*, vol. 192, pp. 105400, 2020. doi: 10.1016/j.cmpb.2020.105400.
- [17] S. Touzani, J. Granderson, and S. Fernandes, Gradient boosting machine for modeling the energy consumption of commercial buildings. *Energy and Buildings*, vol. 158, pp. 1533-1543, 2018. doi: 10.1016/j.enbuild.2017.11.039.
- [18] A. Gelzinis, A. Verikas, E. Vaiciukynas, M. Bacauskiene, J. Minelga, M. Hällander, V. Ulozas, and E. Padervinskis, Exploring sustained phonation recorded with acoustic and contact microphones to screen for laryngeal disorders. In 2014 IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE), pp. 125-132, 2014. doi: 10.1109/CICARE.2014.7007844.
- [19] C. Lin, Q. Chang, and X. Li, A deep learning approach for MIMO-NOMA downlink signal detection. *Sensors*, vol. 19, no. 11, pp. 2526, 2019. doi: 10.3390/s19112526.
- [20] M. Bhurtel, J. Shrestha, N. Lama, S. Bhattarai, A. Uprety, and M. K. Guragain, Deep learning based seed quality tester. *Science, Engineering and Technology (SET) Conference*, pp.1-10, 2019.
- [21] C. Krauss, X. A. Do, and N. Huck, Deep neural networks, gradient-boosted trees, random forests: statistical arbitrage on the S&P 500. *European Journal of Operational Research*, vol. 259, no. 2, pp. 689-702, 2017.
- [22] K. Grolinger, M. A. Capretz, and L. Seewald, Energy consumption prediction with big data: Balancing prediction accuracy and computational resources. In 2016 IEEE International Congress on Big Data (BigData Congress), pp. 157-164, 2016. doi: 10.1109/BigDataCongress.2016.27.

Privacy Preserving Dynamic Provable Data Possession with Batch Update for Secure Cloud Storage

Smita Chaudhari^{1*}, Gandharba Swain²

Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur, Andhra Pradesh, India

Abstract—Cloud Server (CS) is an untrusted entity in cloud paradigm that may hide accidental data loss to maintain its reputation. Provable Data Possession (PDP) is a model that allows Third Party Auditor (TPA) to verify the integrity of outsourced data on behalf of cloud user without downloading the data files. But this public auditing model faces many security and performance issues such as: unnecessary computational burden on user as well as on TPA, to preserve identities of users from TPA during auditing, support for dynamic updates etc. Many PDP schemes creates computational burden either on TPA or Cloud User. To balance this overhead between TPA and User, this paper proposes Privacy-Preserving Dynamic Provable Data Possession (P²DPDP) scheme, which is based on ODPDP scheme. In ODPDP scheme, user relieves the burden by signing a contract with TPA regarding verification of his outsourced data. But this scheme generates computation overhead on TPA. To reduce this computation overhead of TPA, our P²DPDP scheme uses Indistinguishability Obfuscation (IO) with one-way function such as message authentication code to make a lightweight auditing process. P²DPDP scheme uses Rank-based Merkle Tree (RBMT) to support dynamic updates in batch mode which greatly reduces computation overhead of TPA. ODPDP lacks privacy which is maintained in P²DPDP using ring signature technique. Our experimental results demonstrate the reduced verification time and computation cost compared to existing schemes.

Keywords—Public auditing; ring signature; Indistinguishability Obfuscation; Rank-based Merkle Tree (RBMT)

I. INTRODUCTION

With the rapid development of cloud computing, many individuals or small-scale organizations started outsourcing their data on untrusted CS. This paradigm even though, proved to be a boon, has brought many security challenges with it. The user is not having physical ownership of data since outsourced data may be stored at any server. This outsourced data may get damaged unintentionally because of disk crashes or natural disasters. Sometimes CS may delete infrequent data intentionally to create space for new users. These incidents or data loss are not reported to cloud users to maintain reputation [1]. Provable Data Possession (PDP) is a technique that allows any user to check the integrity of data blocks outsourced on CS without downloading the entire data file. Many researchers have proposed cryptographic techniques using homomorphic authenticators [2]-[12]. Cloud users can check the integrity of outsourced data on their own but this frequent task creates an additional burden on the cloud user in terms of time and

computation cost. So, researchers have proposed solution in which user can delegate auditing work to TPA having expertise and skill. TPA generates a challenge message and CS has to produce the proof based on recent data. This proof is verified by TPA. This model has many security challenges since CS is one of the untrusted entities and TPA even if trusted, curious about auditing work.

Integrity checking during dynamic update operations of data is a major challenge in the PDP model. Authenticators have to be recalculated during insertion, deletion, and modification operations on data because most of the authenticators are calculated based on indices of files. Researchers have proposed dynamic auditing schemes using Merkle Hash Tree (MHT) [24]. MHT is a hash-based data structure used for data verification. Another data structure used for dynamic data updates are Index Hash Table (IHT) [5],[6] Dynamic Hash Table (DHT) [18]. These techniques need some additional information to store which may increase storage and computation cost. Guo et al. [21] proposed Multi-leaf-authenticated (MLA) scheme for dynamic data using Rank based Merkle Tree (RBMT). This scheme authenticates multiple leaf nodes at once without storing height and status value. With this scheme, multiple dynamic update operations can be performed in batch mode.

In auditing model, TPA is one of the trusted entity but still curious and may compromise data and user privacy. During auditing, TPA may infer user information who have signed the blocks. Researchers have given multiple approaches to address this issue. Some privacy-preserving protocols [39],[40] are proposed based on aggregate signature [14] or hash-based commitment [15] but auditing is not mentioned in these schemes. Huang et al. [18] proposed privacy-preserving scheme using group signature and blockchain. Tian et al. [20] also proposed a privacy-preserving scheme that addresses data privacy using random masking to blind the data proof. Identity privacy is preserved using a modification record table to record operation information. Different variations of signature algorithms, such as blind signature, random sampling, ID-based privacy, and attribute-based signature are used by multiple researchers during auditing [22]-[28]. Attribute-based signatures are based on attribute-based cryptography [16],[17]. Ring signature is another variation of group signature in which ring or group is formed with multiple users. User sign the blocks and share it among the group members. Verifier determines that block is signed by one of the group members

*Corresponding Author

but can't reveal who has signed the block. Certificateless authentication [29],[30] is one of the cryptographic techniques for authentication. Some authors proposed privacy-preserving integrity verification scheme [31],[32] using certificate-less ring signature which greatly reduces the computation cost during auditing. Li et al. [19] proposed integrity checking of group shared data using certificateless signature. Ni et al. [33] proposed lightweight ID-P²DPDP scheme in which privacy is achieved through zero-knowledge proof.

Many of the above auditing schemes are based on homomorphic authenticators which incur high computation cost and time during auditing. There is a need to propose a lightweight auditing process which can create a little burden on TPA as well as on cloud user. Indistinguishability Obfuscation (IO) is a modern cryptographic technique that uses one-way function for implementation of different cryptography constructs [35]. Researchers [36]-[38] proposed an efficient public verification scheme using IO combined with one-way function MAC which makes this scheme lightweight by generating very little burden on TPA. Tu et al. [13] proposed user-focus auditing which try to reduce the overhead of user by pre-generating challenges for TPA before auditing. Guo et al. [21] also proposed outsourced auditing scheme in which after each verification, TPA generates an audit data log which is checked later by the user. These schemes reduce the burden of verification on user. There is no need for a user to be available during verification, as per his convenience he can check the audit log.

Most of the auditing schemes create computational burden of tag generation and verification either on TPA or user. If we try to reduce the burden of TPA, it will increase the burden on user or vice-versa. Zhang et al. [36] scheme reduces the computation overhead of TPA but cloud user is actively involved in auditing process. Guo et al. [21] scheme proposes auditing scheme where user is not involved in auditing process but it creates computation burden on TPA because of Efficient Homomorphic Verifiable Tags (EHVT).

The remainder of this paper is described as follows: Section II describes related work, problem identification and contributions by authors. Section III elaborates basic building blocks of P²DPDP scheme. Section IV explains the P²DPDP scheme. Section V discusses the evaluation of P²DPDP scheme in terms of security and performance.

II. RELATED WORK, PROBLEM IDENTIFICATION AND RESEARCH CONTRIBUTION

A. Related Work

A good number of solutions have been proposed by many researchers [1]-[10] for integrity verification of outsourced data. In most of these schemes, cloud user and TPA are actively involved during verification phase. This may create an additional burden on cloud users as well as on TPA. Guo et al. [21] proposed ODPDP scheme which relieves user's verification overhead by migrating frequent auditing tasks to TPA. In this scheme, a contract takes place between user and TPA regarding the frequency of verification task. TPA generates challenges based on this contract. After each verification, a log file is generated at TPA which contains an

audit data log. User as per his convenience can check the log and make sure the integrity of his data as well as working of TPA. This scheme greatly reduces the overhead on user side. This scheme uses Multi-Leaf Authentication (MLA) solution with RBMT for dynamic data updates which greatly reduces storage cost as well as allows verification of multiple dynamic operations in batch mode.

To minimize the burden of TPA in terms of computations, it is necessary to develop a verification scheme which is lightweight in terms of computation. Zhang et al. [36] proposed lightweight auditing technique using IO and one-way function MAC. This greatly reduces the computation overhead of TPA during verification since TPA has to just calculate MAC each time and verify it with the received proof from CS. In this scheme, during outsourcing data at CS, a user has an additional overhead of generating circuit (audit program), obfuscating with MAC key, and send it to CS. But this is only a one-time cost to generate obfuscated program since it would not change along with the modification of public parameters and challenge message. This scheme uses Merkle Hash Tree (MHT) to support dynamic data updates on file.

In auditing model, generating privacy-preserving auditing technique is a major challenge because of the curious but trusted nature of TPA. Thokcham [34] proposed a privacy-preserving auditing technique using CDH based ring signature. This ring signature scheme is unforgeable and completely anonymous. Any ring member can sign a message using his private key and public keys of other members. So not every member needs to be present during the signing process. Using this scheme, anyone can check whether a signature is generated by a valid member of the group but at the same time not revealing the user's identity who has signed that message. Thus, preserving the identity privacy of user during verification from TPA.

ODPDP scheme reduces user overhead but increases burden on TPA in terms of computation. The computation cost on TPA in ODPDP during auditing is $(l+s+1)Exp_G + 2Pair$. Where l is number of challenged blocks, s denotes number of sectors per block, Exp_G is exponentiation operation on group G and $Pair$ is pairing operation. Compare to ODPDP scheme, Zhang et al. [36] scheme create less burden on TPA during auditing i.e., $2 Hash_{Z_p}$ where Hash is hashing operation on Z_p . It means TPA has to compute only 2 hash functions for verification. So proposed P²DPDP scheme modifies the ODPDP by using IO and MAC proposed by Zhang et al. scheme instead of EHVT of ODPDP for integrity verification. ODPDP scheme not proposing any solution for identity privacy. We extend this scheme by using CDH based ring signature to achieve identity privacy proposed by Thokcham [34]. P²DPDP also uses RBMT of ODPDP which allows verification of multiple dynamic operations in batch mode compared to MHT in Zhang et al. [36].

B. Problem Identification

To balance the computation overhead between user and TPA, this paper proposes Privacy-Preserving Dynamic Provable Data Possession (P²DPDP) scheme for cloud storage. In this scheme, the main purpose of using IO is to reduce the computation burden of TPA while maintaining security. Since

TPA only needs to validate the commitments generated by CS, user's data will not be revealed to TPA which preserve data privacy during auditing process. To avoid the continuous involvement of user during auditing process, Cloud user and TPA sign a contract which includes starting address of the block, the frequency at which auditor launches a challenge, and number of challenged blocks. TPA verifies the outsourced data based on contract and generates a log file which can be verified by user as per his convenience. For the support of dynamic updates, RBMT is used that can perform multiple update operations in a batch way. P²DPDP support user groups and preserve identity privacy by using CDH based ring signature scheme.

C. Research Contribution

Specifically, the contribution of our scheme P²DPDP is as follows:

- Guo et al. [21] proposed ODPDP scheme which uses Effective Homomorphic Verifiable Tag (EHVT) for integrity verification and RBMT for dynamic data processing. To make the auditing process lightweight and to reduce computation overhead of TPA, we modify the integrity verification scheme of ODPDP by using indistinguishability obfuscation of Zhang et al. [36].
- Guo et al. [21] scheme not offering a solution for identity privacy during auditing. We extend this ODPDP scheme to achieve group user privacy using CDH based ring signature.
- We describe a concrete P²DPDP scheme to be secure and lightweight by modifying ODPDP scheme. Experimental results certify the performance of our scheme.

III. PRELIMINARIES

This section introduces cryptographic building blocks used in P²DPDP scheme such as IO for integrity verification, MLA for dynamic updates and CDH based ring signature scheme for privacy-preserving.

A. Indistinguishability Obfuscation (IO)

Indistinguishability obfuscation is a notion that obfuscates any two distinct (equal size) programs that implement identical functionalities but computationally indistinguishable from each other [35].

Assume $\{C_l\}$ is a circuit class with security parameters l . A uniform PPT algorithm iO having input l , circuit $C \in \{C_l\}$ and outputs a circuit C' is called indistinguishable obfuscator if the following conditions are fulfilled:

1) For all security parameters l , Circuit C , and input x , we have probability as:

$$\Pr[C'(x)=C(x)]=1, \text{ where } C'=iO(l,C) \quad (1)$$

Equation (1) satisfies the completeness property of IO. It states that circuit C' must behave exactly same as circuit C if C' is generated by an independent invocation of iO on C .

2) For any (not essentially uniform) PPT adversaries D , for all security parameter $l \in \mathbb{N}$, for all pairs of circuits $C_0, C_1 \in C_l$, there exists a negligible function Negl such that if $C_0(x) = C_1(x)$ for all inputs x , then.

$$|\Pr[D(iO(l, C_0))=1] - \Pr[D(iO(l, C_1))=1]| \leq \text{Negl}(l) \quad (2)$$

Equation (2) satisfies the indistinguishability property of IO. It states that the secrets embedded in obfuscated program cannot be extracted by D .

Zhang et al. [36] proposed integrity verification using IO and one-way function MAC.

B. Multi-Leaf Authentication (MLA)

ODPDP scheme [21] proposed MLA for dynamic updates. This scheme uses RBMT instead of MHT to support authentication of indices of leaf nodes. In RBMT, no need to store height value as in MHT. Each node contains only two fields (r, h) where r is the rank of a node which is the number of leaf nodes reachable from node ω and h is a hash value of that node. Mainly, the rank of a leaf node is 1 i.e., $r=1$. The second element h is defined as in (3):

$$h = \begin{cases} H_2(m_i), & \text{if } \omega \text{ is the } i\text{th leaf node} \\ H_1(r || \omega.\text{left}.h || \omega.\text{right}.h), & \text{if } \omega \text{ is a nonleaf node} \end{cases} \quad (3)$$

where H_1 and H_2 be a secure collision-resistant hash function and $||$ denotes concatenation. The outsourced file F is divided into n blocks such as $F = \{f_1, f_2, \dots, f_n\}$. The i^{th} element f_i is bind to the i^{th} leaf node of RBMT by storing the hash value of f_i at node ω_i using H_2 in (3). Thus, leaf nodes are already sorted from left to right by their indices. For each non-leaf node, $\omega.\text{left}.hash$ and $\omega.\text{right}.hash$ indicates the hash value of the left node and right node respectively calculated using H_1 in (3). RBMT can be constructed for given n data blocks. A Merkle root h_{root} is sufficient to check the integrity of dynamic updates in a tree because of the dependency of all data blocks. Fig. 1 shows the RBMT constructed over 14 data blocks. In Fig.1, when multiple leaf nodes are challenged using MHT such as ω_3 and ω_7 , the proofs generated are $\Omega_3 = \{\omega_{26}, \omega_{22}, \omega_{15}, \omega_4, \omega_3\}$ and $\Omega_7 = \{\omega_{26}, \omega_{21}, \omega_{17}, \omega_{18}, \omega_7\}$. During verification using MHT, some repeated and unnecessary nodes have to be retrieved and processed that incur large computation costs. But using MLA solution, if multiple challenged nodes are $(3,7,8,10,13)$, the corresponding multi-proof \sqcup_p is:

$$\sqcup_p = \{\omega_3, \omega_7, \omega_8, \omega_{10}, \omega_{13}, \omega_{16}, \omega_{18}, \omega_{19}, \omega_{21}, \omega_{22}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{26}\}$$

where every necessary node appears just once which reduces computation cost as well as support multiple updates in batch mode.

C. CDH based Ring Signature Scheme

To achieve privacy during auditing, Thokcham [34] used CDH based ring signature which is one of the unforgeable and anonymous technique. No centralized entity is involved i.e., no concept of group manager. This scheme comprises two algorithms: Ring_sign and Ring-verify.

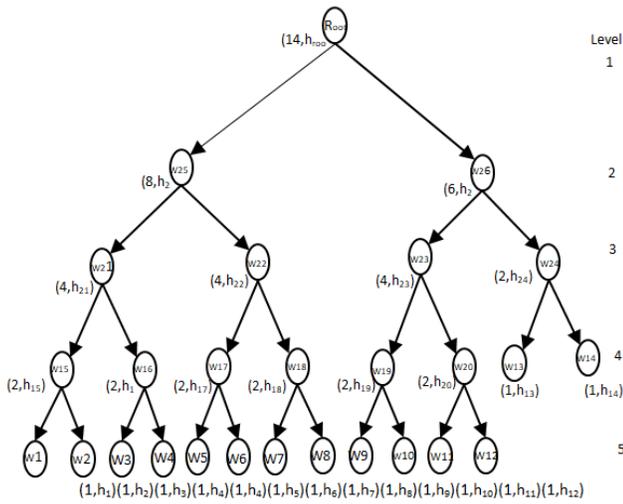


Fig. 1. RBMT Authentication Tree.

Ring-sign: This algorithm takes as input given message M . For a size of ring n , each group member chooses a secret key $Sk = x_i$ which belongs to Z_p and public key $Pk = g^{x_i}$.

- Signer t uses global parameter $d, u_0, u_1, \dots, u_l \in G_1$ of l random elements.
- Signer t will choose random $r_i \in Z_p$ for all other members of the group and generates $s_i = g^{r_i}$. Signer t again computes signature on behalf of group

$$s_t = (d \cdot \prod_{i=1, i \neq t}^n Pk_i^{r_i} (u_0 \cdot \prod_{j=1}^l u_j^{m_j})^{-r_{n+1}})^{1/x_t} \quad (4)$$

In (4), signer t computes signature s_t using his private key X_t with different parameters such as: public keys of n group members PK , the global parameter d, u_0, u_1, \dots, u_l , message M divided into l elements (M_1, \dots, M_l) , random number r . Final signature is $\sigma = (s_1, s_2, \dots, s_{n+1})$.

Ring-verify: As per (5), verifier verifies the signature using received ring signature σ , message M , and public keys PK of all members. The verifier checks the following equality.

$$\prod_{i=1}^n e(s_i, Pk_i) \cdot e(s_{n+1}, u_0 \prod_{j=1}^l u_j^{m_j}) = e(g, d) \quad (5)$$

In (5), using public keys PK , signature $(s_1, s_2, \dots, s_{n+1})$, and received message M , recomputes $\prod_{i=1}^n e(s_i, Pk_i) \cdot e(s_{n+1}, u_0 \prod_{j=1}^l u_j^{m_j})$. This recomputed part is verified using global parameter d and g .

IV. PROPOSED SCHEME

A. System Model

The framework for our P²DPDP scheme is as shown in Fig. 2. It consists of three entities: Cloud User, CS, and TPA.

- Cloud User: Cloud user is one of the members of user group who can share a file in a group. Users can check the integrity of shared files through an audit log generated by TPA.
- CS: an entity having the capability of computation and storage at its end. It is having the responsibility to maintain and manage outsourced files.

- TPA: an external entity that works on behalf of users and expertise in verifying the integrity of outsourced data.

The workflow for P²DPDP scheme from Fig. 2 is as follows:

- 1) Any cloud user from a group outsources data file on CS. Before outsourcing, user calculates the tag for each block, signs the blocks using a ring signature scheme.
- 2) The user constructs RBMT tree using hash values of file blocks. User generates circuit for auditing program, obfuscates it, and sends to CS. Shares MAC key to TPA as well as sign a contract with auditor regarding verification activity.
- 3) Based on the frequency mentioned in the contract, TPA generates challenges and performs auditing activity. During an audit, using CDH based ring signature, TPA verifies group signature using public keys of all members in a group.
- 4) Generates log file for each activity.
- 5) Users can check the log entry at any time to verify the integrity of an outsourced file.
- 6) For dynamic updates, user sends the update command uc to the TPA.
- 7) TPA updates the RBMT tree according to uc and sends updated proof to CS for verification. If verification successful, CS sends signed proof to user.
- 8) User verifies proof and if successful, send updated data blocks u_i to CS. CS updates data blocks accordingly.

B. Design Goals

To achieve privacy-preserving during integrity verification of outsourced data, proposed scheme P²DPDP should satisfy the following design objectives:

- 1) *Public verification*: to allow an external auditor to verify the integrity of outsourced data without downloading the data file.
- 2) *Privacy-Preserving*: data or user identity must not be revealed to TPA during auditing.
- 3) *Data Dynamic Support*: integrity verification process must support dynamic updations on outsourced data such as insert, delete and modify operations.
- 4) *Lightweight*: verification process must create minimum communication and computation overhead on user and TPA.

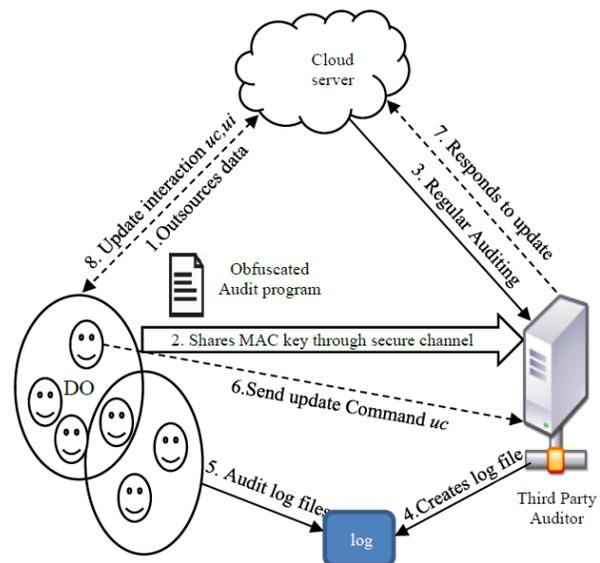


Fig. 2. Architecture of P²DPDP Scheme.

C. P²DPDP Scheme

Our proposed P²DPDP scheme works in four phases: Setup, Store, AuditData, and AuditLog.

Setup: Let G and G_T be two multiplicative groups produced by g with order p contains bilinear map $e: G \times G \rightarrow G_T$. User

U selects a signing key pair (ssk, spk) , α, v where $\alpha \rightarrow Z_p$ and $v = g^\alpha \in G$. U picks random elements u_1, u_2, \dots, u_s and fixes pseudorandom permutation, function key $\pi_{key}(\cdot)$ and $f_{key}(\cdot)$ respectively. The secret and public parameters are $sk=(\alpha, ssk)$ and $pk=(v, spk, u_1, u_2, \dots, u_s)$. Group members randomly select private key as $x_i \in Z_p$ Using key generation of CDH based ring signature scheme and $Pk_i = g^{x_i} \in G$ as a public key.

Store: Initially according to (6), U divides the file F into blocks n and sector s as in.

$$F = \{f_{i,j} \mid 1 \leq i \leq n, 1 \leq j \leq s\} \quad (6)$$

Tag Generation:

U chooses random element name for file and computes file tag as

$$\tau = \text{name} \parallel n \parallel u_1, u_2, \dots, u_s \parallel \text{sig}_{ssk}(\text{name} \parallel n \parallel u_1, u_2, \dots, u_s)$$

and data tag as in (7).

$$\sigma_i = (H(i \parallel \text{name}) \cdot \prod_{j=1}^s u_j^{f_{ij}})^\alpha, i \in [1, n] \quad (7)$$

In (7), User calculates data tag for each block i of file F using random elements of each sector u_1, u_2, \dots, u_s and hash value of block number and file name. Here H is any secure hash function. U generates processed data M as $M = \{M, \phi\}$ where

$$\phi = \{\sigma_i, \tau\}_{i \in [1, n]}$$

Constructing RBMT:

U first calculates hash values for each block of file F using H_2 .

$$h_i = H_2(m_i) \text{ where } 1 \leq i \leq n$$

Then generate tree TR using RBMT on ordered hash values. Each leaf node w_i stores the corresponding hash value h_i .

Ring Signature Generation:

U has to sign a block on behalf of a group using CDH based ring signature scheme. U randomly chooses $u_0, r_i \in Z_p$ and compute signature for all other group members except U denoted by j in (8).

$$S_i = g^{r_i} \text{ for } i = \{1, 2, \dots, n+1\} \setminus \{j\} \quad (8)$$

Where, n - number of members in a group

j - serial number of the member in the signature who

is signing it

U computes signature for every member using respective random number r of that user. Then computes $h = H(\phi \parallel T)$ where T is a timestamp. U again computes S_j using (4).

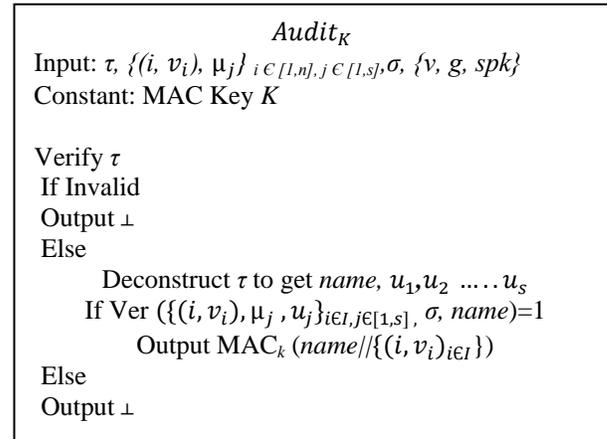
The signature at time T is

$$\phi^T = (S_1, S_2, \dots, S_{n+1}) \quad (9)$$

U outsources M and ϕ^T calculated as in (9) on CS. In (9), S_1, S_2, \dots, S_{n+1} is the signature generated for n users by U on behalf of group at time T .

Outsourcing Auditing Task:

During this task, U chooses a MAC key k and passes it to TPA using a secure network. U also produces a circuit **Audit_k** as described below.



This circuit is similar to auditing program which generate the MAC using embedded MAC Key K based on given input. Uniform PPT algorithm iO proceeds with audit circuit **Audit_k** as input and generates public parameter P as $P = iO(\text{Audit}_k)$.

$\text{Ver}(\{(i, v_i), \mu_j, u_j\}_{i \in [1, n], j \in [1, s]}, \sigma, \text{name})$ in circuit denotes checking of (10):

$$e(\sigma, g) = e(\prod_{i \in [1, n]} H(i \parallel \text{name})^{v_i} \cdot \prod_{j=1}^s u_j^{\mu_j}, v) \quad (10)$$

- U outsources RBMT tree TR with signature $\text{Sig}_{ssk}(TR)$ to TPA. If verification is successful, TPA accepts TR else rejected considering the malicious author.

Agreeing Parameters:

- All group members need to sign on a public parameter $P_{pub} = \{q, h_{root}\}$ where q is the number of data blocks and h_{root} is a Merkle root of TR .
- Contract CT is established between U and TPA as

$CT = \{BI, Fr, b\}$, where

BI- block index from which auditing work will start.

Fr- Frequency at which TPA launches a challenge.

b- number of challenged data blocks for checking.

AuditData Protocol: This protocol mainly deals with checking the integrity of outsourced data and log generation by TPA.

Auditing Phase:

- TPA generates a challenge message $Q^{(b)} = \{b, K_1^{(b)}, K_2^{(b)}\}$ using data blocks b to be audited. TPA

Generates pseudorandom permutation and function keys $\{K_1^{(b)}, K_2^{(b)}\}$ respectively and send to CS.

- CS computes $i = \pi K_1^{(b)}(\xi)$ and $v_i = f K_2^{(b)}(\xi)$ where $\xi \in [1, b]$ and b is the size of I (Input blocks to be audited). Based on public parameters and corresponding σ_i calculated using (7), $\tau, f_{i,j}$ and b , CS computes:

$$\sigma^{(b)} = \prod_{i \in I} \sigma_i^{v_i}, \mu_j = \sum_{i \in I} v_i f_{i,j} \text{ and}$$

$$PRF^{(b)} = P(\tau, \{(i, v_i), \mu_j\}_{i \in I}, \sigma^{(b)}, \{v, spk\}) \quad (11)$$

Equation (11) shows the proof PRF calculated by CS for challenged blocks b . PRF is calculated by executing audit circuit $Audit_K$ i.e. $P = iO(Audit_K)$ using input parameters $file\ tag\ \tau, \{(i, v_i), \mu_j\}, \sigma, \{v, g, spk\}$. CS send this $PRF^{(b)}$ and $Sig_{sk_{CS}}(PRF^{(b)})$ to TPA for verification.

- Using CDH based ring signature process, TPA verifies the group signature based on input signature ϕ^T , public keys $(Pk_1, Pk_2, \dots, Pk_n)$ of all members in a group, F^T and public parameter u_0 . TPA first calculate $h = H(\phi || T)$. Then verifies signature using (5).
- To verify the integrity of data, TPA computes $i = \pi K_1^{(b)}(\xi)$ and $v_i = f K_2^{(b)}(\xi)$ and compare MAC with $PRF^{(b)}$ calculated by CS using (11).

$$PRF^{(b)} = MAC_k(\text{name} || \{(i, v_i)_{i \in I}\}).$$

Log Generation:

After every audit, either successful or fail, TPA creates a log record of his auditing work.

$$L^{(b)} = \{t, Q^{(b)}, PRF^{(b)}, Sig_{sk_{CS}}(PRF^{(b)})\}$$

and saves in his local file Log_File .

AuditLog Protocol:

- U generates challenge using random subset B of file block indices and sends it to TPA. For each $b \in B$, TPA finds challenge $Q^{(b)}$, proof $PRF^{(b)}$ from his log file. Computes i, v_i from $Q^{(b)}$. TPA generates multi_audit proof \sqcup_p using h_{root} of TR and generates the proof of appointed log for subset B using (12).

$$PRF^B = \{ \sqcup_p, i, v_i, PRF^{(b)} \} \quad (12)$$

In (12), B indicates the challenge generated by U during $AuditLog$. Elements i, v_i denotes the challenge retrieved through log file for blocks b and $PRF^{(b)}$ indicates proof retrieved from log file for blocks b . \sqcup_p is a multi-audit proof generated from RBMT.

- TPA send a signed proof with his signature $Sig_{sk_{TPA}}(PRF^B)$ to U for verification. After verifying the signature, U computes new PRF as.

$$PRF_{new} = MAC_k(\text{name} || \{(i, v_i)_{i \in B}\})$$

- U compares if $PRF_{new} = PRF^B$. If matched, verification of outsourced data is successful else verification fails. U also verify \sqcup_p using h_{root} of TR .

D. Support for Dynamic Updates

P^2DPDP scheme support three types of update operations such as: deletion, insertion, and modification on blocks. If we perform these updates one by one, it will incur a large computation overhead at the auditor side to generate and verify the hash tree. To reduce this overhead, P^2DPDP is based on a MLA scheme using RBMT proposed by ODPDP which can handle updates in batch instead of one by one.

Initially, U computes all the hash and tag values of the new file block in $Store$ phase, generates the RBMT tree, and set public parameter as $P_{pub} = \{q, h_{root}\}$. U sends the update command uc to CS and TPA. U also generates audit circuit same as basic scheme but with modified verification function $VerD(\{(i, v_i), \mu_j, u_j, \sigma, name\}_{i \in I, j \in [1, s]})$ which consists of checking following equation:

$$e(\sigma, g) = e(\prod_{i \in I} H(\sum_{j=1}^s m_{ij})^{v_i} \cdot \prod_{j=1}^s u_j^{\mu_j}, v)$$

After receiving uc command, TPA updates leaf nodes, other affected nodes and generate an updated tree TR^* and its Merkle root h_{root}^* . TPA then sends the updated signed proof up to CS. CS verifies up by executing the audit circuit. U can also later check the correctness of up. If verification is successful, U sends updated information ui to CS and CS updates the processed data.

V. EVALUATION

In this section, P^2DPDP scheme is evaluated by showing correctness proof, security analysis, performance and experiment analysis.

A. Correctness Proof

$Audit_K$ is an audit circuit generated by U during $Store$ phase. Upon execution of this audit circuit, CS generates MAC based on challenge, block tag σ and using global parameters. Audit circuit contains verification function which denotes the realization of (10). Correctness proof for (10) is as follows:

$$\begin{aligned} e(\sigma, g) &= e(\prod_{i \in I} \sigma_i^{v_i}, g) \\ &= e(\prod_{i \in I} (H(i || \text{name})). \prod_{j=1}^s u_j^{f_{ij}})^{v_i}, g) \\ &= e(\prod_{i \in I} (H(i || \text{name})^{v_i}). \prod_{i \in I} \prod_{j=1}^s u_j^{f_{ij} v_i}, g) \\ &= e(\prod_{i \in I} (H(i || \text{name})^{v_i}). \prod_{j=1}^s u_j^{f_{ij}}, v_i) \end{aligned}$$

B. Security Analysis

Storage Correctness:

Theorem 1: If the CS successfully passes the verification from an auditor, then data outsourced on CS must be intact.

Proof: Assume user U outsourced data at CS using $Store$ protocol. But due to some problems, data at CS accidentally corrupted or deleted. With P^2DPDP scheme, malicious CS can't pass its verification. We prove this by game sequence as below:

- 1) Based on a contract signed between U and TPA, TPA generates a challenge using $AuditData$ Protocol.
- 2) For the challenged blocks, CS computes i, v_i . Using

- 3) Public parameters and (k_1, k_2) , CS computes \overline{PRF} and send to TPA.
- 4) TPA computes $i = \pi k_1(\xi)$ and $v_i = f k_2(\xi)$ where $\xi \in [1, c]$ and c is the size of I (Input blocks to be audited).
- 5) Using MAC_k , TPA calculates PRF and compares it with \overline{PRF} .
- 6) CS wins if TPA passes the verification even if $PRF \neq \overline{PRF}$.

But in above game, it's very difficult for malicious CS to cheat auditor because of HMAC scheme during verification.

Liability:

Theorem 2: An honest auditor can demonstrate that he did his work correctly in case of any disputes.

Proof: To prove the liability of the auditor, we consider two situations: when an auditor is honest or an auditor is dishonest. Consider first the auditor is honest. As per the contract between auditor and User, the auditor generates a challenge $Q^{(b)}$. User can reconstruct the challenge since the contract consists of number of data blocks to be audited. User can check the value of PRF by recalculating (11). Honest auditor generates a log file named Log_File which is the evidence for all the auditing work completed by auditor. So honest auditor can prove his liability by this Log_File.

Compare to this, if the auditor is malicious and not doing his work properly, user can use AuditLog Protocol to verify the behavior of auditor. By regenerating challenge, user can check the AuditLog file anytime and auditor can't deny his misbehavior.

Privacy-Preserving:

Theorem 3: From the server's response to the challenge message, TPA not able to infer any information such as data and identity of user.

Proof: During verification, user U first generates an audit circuit (which is nothing but an auditing algorithm program which is supposed to be originally executed by TPA). U obfuscates this circuit by embedding MAC key K and send to CS. For each verification, CS computes the inputs based on the challenge message and executes the obfuscated program. CS generates the MAC tag and sends it to TPA. TPA has to only verify the MAC tag to check the integrity of outsourced data. TPA needs to calculate i and v_i based on challenged blocks using the HMAC scheme. So, it is computationally infeasible for TPA to infer any information or user data using P^2DPDP .

P^2DPDP uses CDH based ring signature scheme to share any data among group members. In this scheme, user who want to share a file, computes signature on this data with his own private key using (4). During verification, TPA can verify the

signature with public keys of all users. Using this scheme, TPA can check whether the signature is computed by a valid user of group or not but scheme can't reveal individual user identity to verifier. Thus because of CDH based ring signature scheme and IO, P^2DPDP proved to be privacy-preserving.

C. Performance Analysis

We first evaluate the performance of P^2DPDP scheme which shows the privacy-preserving, lightweight auditing process. Also, we compare the performance of P^2DPDP with existing schemes.

The main important functionalities which we have considered for this work are: public auditing, dynamic data operations, privacy-preserving, and group support. Table I shows the functionality comparison of P^2DPDP scheme with existing schemes.

To evaluate the performance of P^2DPDP scheme, we evaluate the communication cost between CS and TPA during the proof generation and verification phase of *AuditData* protocol. Communication cost between the user and CS is not important since user uploads the data entirely to CS initially and user can verify the integrity of outsourced data during *AuditLog* protocol. During proof generation, TPA generates a challenge message $\{k_1, k_2\}$ for b number of blocks where k_1 and k_2 are transformed keys of HMAC. After receiving challenge message, CS generates proof PRF by calculating HMAC through the obfuscated program. So, communication overhead for proof generation is bH where H is any secure hash operation. After generating the proof PRF using HMAC, CS sends it to TPA. During verification, TPA checks the integrity of outsourced blocks using MAC key and verify

$$PRF = MAC_k(\text{name} || \{(i, v_i)_{i \in I}\})$$

So TPA has to calculate only HMAC. TPA also verifies the user signature by verification algorithm of CDH based ring signature. Using (5), TPA verifies the users n . So, the total communication overhead during verification is $H_{Z_p} + n$ where H_{Z_p} is hashing operation into Z_p .

User can check the integrity of outsourced data or performance of TPA during *AuditLog* protocol. So total Communication overhead during *AuditLog* is also H_{Z_p} . Table II shows the comparison of P^2DPDP scheme with the existing scheme. Fig. 3 shows the comparison between ORUTA, CORPA, and our P^2DPDP scheme for communication overhead in KB with respect to group size.

We can't compare this cost with ODPDP scheme since it doesn't support user groups. The comparison demonstrates the noticeable and constant performance of communication cost between CS and TPA during auditing in P^2DPDP scheme.

TABLE I. THE COMPARISON OF AUDITING FUNCTIONALITIES

Scheme	Third-Party Auditor	Dynamic Data Operation	User Group Support	Privacy Preserving
ORUTA [5]	Homomorphic Authenticators (HA)	Index Hash Table	Ring Signature	Homomorphic Authenticable Ring Signature (HARS)
Tian et al. [20]	Homomorphic Verifiable Authenticators (HVA)	Dynamic Hash Table	Group Signature	Random Masking
ODPDP [21]	Efficient HVA	Rank Based Merkle Tree (RBMT)	NA	NA
CORPA [23]	HA	Merkle Hash Tree (MHT)	Group Signature (GS)	HA+GS
Thokcham and Saikia[34]	Vector Commitment	MHT	CDH based Ring Signature	CDH based Ring Signature
Zhang et al. [36]	Indistinguishability Obfuscation (IO)	MHT	NA	NA
Proposed Scheme P ² DPDP	Indistinguishability Obfuscation (IO)	Rank Based Merkle Tree (RBMT)	CDH based Ring Signature	CDH based Ring Signature

TABLE II. COMMUNICATION COST

Scheme	Proof Generation	Proof Verification	
		AuditData	AuditLog
ORUTA [5]	$(s+nb)E+nbM+bsM+sH$	$(2s+b)E+(2s+b)M+nM+bH+(n+2)p$	NA
ODPDP [21]	$(2q+s+1)E$	$(b+s+1)E+2p$	$3E$
CORPA [23]	$H+M$	$2E+P+bE+bM$	NA
Proposed Scheme P ² DPDP	bH	$H+n$	$H+n$

s : Total no. of sectors
 n : Total no. of users in a group
 b : No. of challenged blocks
 q : Total no. of blocks
 E : Exponentiation operation
 M : Multiplication Operation
 H : Hash Operation
 P : Pairing Operation

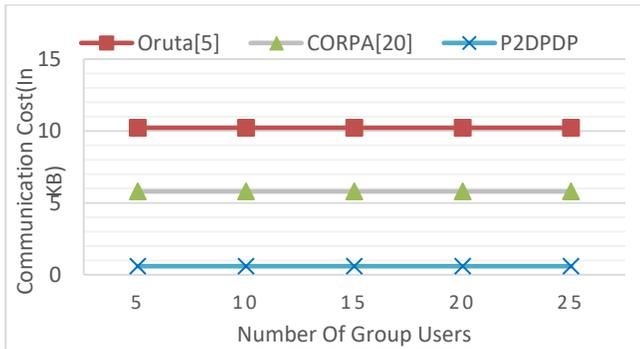


Fig. 3. Communication Cost w.r.t. Group Size.

D. Experimental Results

This section proves the performance of P²DPDP system in terms of different experiments. We deployed our P²DPDP scheme on a system comprising Windows 8.1 with an Intel Core i5-5200U CPU functioning at 2.20 GHz, 4.0 GB RAM. Python is used for module implementation of P²DPDP scheme. The hash algorithm is instantiated using SHA256. Fig. 4 shows the impact of number of users in a group on verification time during AuditData protocol. It shows that verification time is independent of the number of users. By creating a group of 25 users, we have compared the results with Oruta and CORPA

scheme. We have not considered ODPDP scheme for comparison since it doesn't support user groups. All results are average of 5 runs. For 20 users in P²DPDP scheme, the Verification time is 0.15 seconds. While Oruta, and CORPA are 2.24, and 1.75 seconds respectively. Result proves the effectiveness and lightness of our scheme because of reduced and constant verification time at auditor side during AuditData protocol. Since our P²DPDP scheme is based on ODPDP scheme, it is mandatory for us to compare results with this scheme. In both the schemes, verification is performed during AuditData and AuditLog protocol. Initially we compare the results of both schemes during AuditData where TPA performs audit based on contract and generate log entries. Fig. 5 shows the proof generation and verification time in seconds with respect to challenged data blocks during AuditData protocol.

To compare the results with ODPDP scheme, in P²DPDP scheme, the block size is kept fixed i.e., 16KB. Total outsourced data is 1GB. Fig. 5(a) shows the constant proof generation time in P²DPDP scheme as compared to ODPDP. Fig. 5(b) shows the gradual increase in proof verification time as number of challenged blocks is increasing in P²DPDP as compared to ODPDP scheme. Fig. 6 shows the performance of P²DPDP during AuditLog. It presents the computation time and communication cost required by user to verify the past work of TPA under the number of checked log entries.

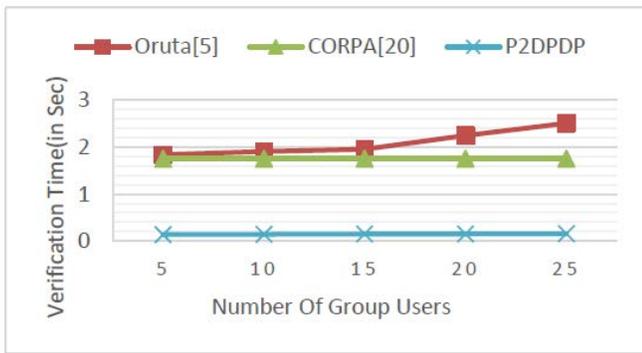


Fig. 4. Impact of Number of Group users on Verification Time.

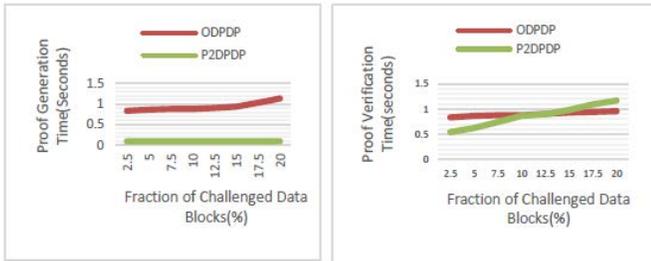


Fig. 5. (a) Proof Generation Time during AudiData Protocol w.r.t. Challenged Data Blocks; (b) Proof Verification Time during AudiData Protocol w.r.t. Challenged Data Blocks.

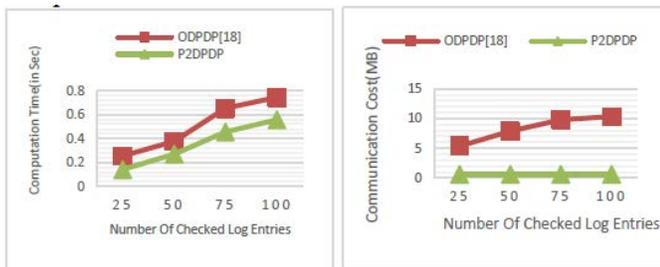


Fig. 6. (a) Computation Time for user to Check Log Entries w.r.t. Number of Checked Log Entries; (b) Communication Cost for user to Check Log Entries w.r.t. Number of Checked Log Entries.

For experiments, we have analyzed the computation time of our scheme up to 100 log entries. As expected, computation time is increasing linearly with number of checked log entries. But computation time of P²DPDP scheme is reduced as compared to ODPDP scheme.

Fig. 6 shows the communication cost during AuditLog protocol. Result shows that P²DPDP scheme is giving better performance in terms of communication cost compared to ODPDP.

VI. CONCLUSION

In most of the previous auditing scheme, Cloud user and TPA are actively involved during verification process which may create additional burden in terms of time and cost. This paper proposes P²DPDP scheme for cloud storage in which there is no need of user during verification process. TPA generates challenges based on the contract signed between TPA and user. TPA also generates log which can be audited by user as per his convenience. P²DPDP scheme also create the light-

weight verification process so as to reduce the computation burden of TPA using new cryptographic technique, Indistinguishability Obfuscation and MAC. P²DPDP support and manages user groups using CDH based ring signature scheme. CDH based ring signature is anonymous scheme which preserves the identity of users from TPA during auditing. P²DPDP scheme supports dynamic updates in batch mode using MLA solution proposed by ODPDP scheme which is based on RBMT.

Security analysis and experiments show that P²DPDP scheme is secure, lightweight and privacy-preserving. Communication cost during auditing between CS and TPA is almost constant and reduced compared to Oruta and CORPA since TPA has to just calculate MAC and compare it with MAC received from CS. Verification time is also reduced and constant compare to existing schemes. Experimental results reveal that verification time is independent of number of group users. Results of AudiLog protocol shows that P²DPDP scheme is performing better in terms of communication time and cost as compared to ODPDP scheme. CDH based ring signature generates certificates which need to be processed during verification leads to increase computation time. In terms of future work, we plan to modify P2DPDP scheme using certificateless signature schemes to reduce computation time.

REFERENCES

- [1] S. Chaudhari, S. K. Pathuri, "A Comprehensive Survey on Public Auditing for Secure Cloud Storage," International Journal of Engineering and Technology, vol. 7, no. 2.7, pp. 564-569, 2018.
- [2] C. Wang, S. Chow, Q. Wang, K. Ren and W. Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage", IEEE Transactions on Computers, Vol. 62, no.2, pp.362-375, 2013.
- [3] B. Wang, H. Li and M. Li, "Privacy-Preserving Public Auditing for Shared Cloud Data Supporting Group Dynamics", in Proc. IEEE International conference on Communications, pp. 1946-1950, 2013.
- [4] K. Yang and X. Jia, "An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing", IEEE Transactions on Parallel and Distributed Systems, Vol. 24, No. 9, pp. 1717-1726, 2013.
- [5] B. Wang, B. Li and H. Li, "Oruta: Privacy-Preserving public Auditing for Shared Data in the Cloud", IEEE Transactions on Cloud Computing, Vol.2, No. 1, pp. 43-56, 2014.
- [6] Y. Yu, Y. Mu, J. Ni, J. Deng and K. Huang, "Identity Privacy-Preserving Public Auditing with Dynamic Group for Secure Mobile Cloud Storage", in Proc. International Conference on Network and System Security, LCNS 8792, pp. 28-40, 2014.
- [7] Y. Luo, M.Xu, K. Huang, D. Wang and S. Fu, "Efficient Auditing for Shared Data in the Cloud With Secure User Revocation and Computations Outsourcing", Computers and Security, Vol.73, pp.492-506, Mar.2018.
- [8] H. Tian, Y. Chen, C. Chang, H. Jiang, Y. Huang, Y. Chen and J. Liu, "Dynamic Hash Table based Public Auditing for Secure Cloud Storage", IEEE Transactions on Service Computing, Vol. 10, No.5, pp. 701-714, 2017.
- [9] G. Yang, J. Yu, W. Shen, Q. Su, Z. Fu and R. Hao, "Enabling Public Auditing for Shared Data in Cloud Storage Supporting Identity Privacy and Traceability" Journal of Systems and Software, Vol. 113, pp. 130-139, 2016.
- [10] M. Ma, J. Weber and J. Berg, "Secure Public-Auditing Cloud Storage Enabling Data Dynamics in the Standard Model", in Proc. IEEE Conference on Digital Information Processing, Data Mining and Wireless Communications, pp. 170-175, 2016.
- [11] B. Kang, J. Wang and D. Shao, "Attack on Privacy-Preserving Public Auditing Schemes for Cloud Storage", Mathematical Problems in Engineering, vol. 2017, Article ID 8062182, 2017.

- [12] M. Krishna, B. Harika, P. SasiPriya, V. Nikitha, K., and Bharath M, "Homomorphic cryptography", Journal of Advanced Research in Dynamical and Control Systems, Vol.10, No. 4, pp.129-136, 2018.
- [13] T. Tu, L. Rao, H. Zhang, Q. Wen and J. Xian, "Privacy-Preserving Outsourced Auditing Scheme for Dynamic Data Storage in Cloud", Security and Communication Networks, vol. 2017, Article ID 4603237, 2017.
- [14] C. Xu, X. Shen, L. Zhu and Y. Zhang, "A Collusion-Resistant and Privacy-Preserving Data Aggregation Protocol in Crowdsensing System", Mobile Information Systems, vol. 2017, Article ID 3715253, 2017.
- [15] X. Cao, H. Li, L. Dang and Y. Lin, "A Two-Party Privacy-Preserving Set Intersection Protocol against Malicious Users in Cloud Computing", Computer Standards and Interfaces, Vol. 54, No. 1, pp. 41-45, 2017.
- [16] L. Tamma, and S. Ahamad, "A novel chaotic hash-based attribute-based encryption and decryption on cloud computing", International Journal of Electronic Security and Digital Forensics, Vo. 10, No.1, pp. 1-19, 2018.
- [17] N. Vurukonda, S. Trijan Kumar, J. R. Reddy, A., Adithya, and S. Boddu, "A secure attribute-based encryption scheme in cloud computing", International Journal of Engineering and Technology (UAE), Vol. 7, No. 2, pp. 90-92, 2018.
- [18] L.Huang, G. Zhang and A. Fu, "Privacy-Preserving Public Auditing for Non-Manager Group Shared Data", Wireless Press Communications, pp. 1277- 1294, 2018.
- [19] J. Li, H. Yuan and Y. Zhang, "Certificateless Public Integrity Checking of Group Shared Data on Cloud Storage", IEEE Transactions on Services Computing, vol. 14, no. 1, pp. 71-81, 1 Jan.-Feb. 2021.
- [20] H. Tian, F. Nan, H. Jiang, C. Chang, J. Ning and Y. huang, "Public Auditing for Shared Cloud Data with Efficient and Secure Group management", Information Sciences 472, pp. 107-125, 2019.
- [21] W.Guo , H. Zhang , S. Qin, F. Gao , Z. Jin , W. Li and Q.Wen , "Outsourced Dynamic Provable Data Possession with Batch Update for Secure Cloud Storage, Future Generation Computer Systems 95, pp.309-322, 2019.
- [22] Y.Ming and W. Shi, "Efficient Privacy-Preserving Certificateless Provable Data Possession Scheme for Cloud Storage", IEEE Access, vol 7, pp. 122091-122105, 2019.
- [23] R. Rabaninejad, M. Attari, M. Asaar and M. Aref, "A Lightweight Auditing Service for Shared Data with Secure User Revocation in Cloud Storage", IEEE Transactions on Services Computing, 2019.
- [24] W. Shen, J. Qin, J. Yu, R. Hao and J. Hu, "Enabling Identity-based Integrity Auditing and Data Sharing with Sensitive Information Hiding for Secure Cloud Storage", IEEE Transactions on Information Forensics and Security", Vol. 14. No.2, pp. 331-346, 2019.
- [25] R. Rabaninejad, M. Attari, M. Asaar and M. Aref, "A Lightweight Identity-based Provable Data Possession Supporting User Identity Privacy and Traceability", Journal of Information Security and Applications, Vol. 51, 2020.
- [26] Y. Yu, Y. Li, B. Yang, W. Susilo, G. Yang and J. Bai, "Attribute-based Cloud Data Integrity Auditing for Secure Outsourced Storage", IEEE Transactions on Emerging Topics in Computing, vol. 18, No.2, pp. 377-390, 2020.
- [27] Y. Zhang, J. Yu, R. Hao, C. Wang and K. Ren, "Enabling Efficient User Revocation in Identity-based Cloud Storage Auditing for Shared Big Data", IEEE Transactions on Dependable and Secure Computing, Vol. 17, No. 3, pp. 608-619, 2020.
- [28] J. Gudeme, S. Pasupuleti and R. Kandukuri, "Attribute-based Public Integrity Auditing for Shared Data with Efficient User Revocation in Cloud Storage", Journal of Ambient Intelligence and Humanized Computing 12, pp. 2019-2032, 2020.
- [29] N. B Gayathri, G. Thumbur, V. P. Reddy, and Md, Z. U. Rahman, "Efficient Pairing-free Certificateless Authentication Scheme with Batch Verification for Vehicular Ad-hoc Networks", IEEE Access, Vol. 6, pp. 31808-31819, 2018.
- [30] K. H. Vamshi, and G. Swain, "Identification and avoidance of malicious nodes by using certificate revocation method", International Journal of Engineering and Technology (UAE), Vol. 7, No. 4.7, pp.152-156, 2018.
- [31] K. Zhao, D. Sun, G. Ren and Y. Zhang, "Public Auditing Scheme with Identity Privacy Preserving based on Certificateless Ring Signature for Wireless Body Area Networks", IEEE Access, Vol. 18, pp. 41975-41984, 2020.
- [32] X. Yang, M. Wang, T. Li, R. Liu and C. Wang, "Privacy-Preserving Cloud Auditing for Multiple Users Scheme with Authorization and Traceability", IEEE Access, Vol 8, pp. 130866-130877, 2020.
- [33] J. Ni, K. Zhang, Y. Yu and T. Yang, "Identity-based Provable Data Possession from RSA Assumption for Secure Cloud Storage", IEEE Transactions on Dependable and Secure Computing, 2020.
- [34] S. Thokcham and D. Saikia , " Privacy Preserving Integrity Checking of Shared Dynamic Cloud Data with User Revocation" , Journal of Information Security and Applications , pp.2214-2126 , 2020.
- [35] A. Sahai and B. Sahai "How to Use Indistinguishability Obfuscation: Deniable Encryption and More, in Proc. 46th Annual ACM Symposium on Theory of Computing, pp.475-484, May.2014.
- [36] Y. Zhang, C. Xu, X. Linag, H. Li, Y. Mu and X. Zhang, "Efficient Public Verification of Data Integrity for Cloud Storage Systems from Indistinguishability Obfuscation", IEEE Traction on Information Forensics and Security, Vol.10, No.3, pp.676-688, Mar.2017.
- [37] S. Chaudhari, G. Swain and P. Mishra, "Secure and Verifiable Multiparty Computation using Indistinguishability Obfuscation", International Journal of Intelligent Engineering and Systems, Vol.13, No.5, pp.277-285, Jul.2020.
- [38] S. Chaudhari and G. Swain, "Efficient and Secure Group based Collusion Resistant Public Auditing Scheme for Cloud Storage", International Journal of Advanced Computer Science and Applications, Vol.12, No. 3, pp. 472-481, 2021.
- [39] B. Dhote, & Krishna Mohan G., "Trust and security to shared data in cloud computing: Open issues", Advances in Intelligent Systems and Computing, Vol. 870, pp.117-126, 2019.
- [40] B. Tirapathi Reddy, C. Rao and M. V. P., "Privacy preserving proof of ownership for data in cloud storage systems" International Journal of Engineering and Technology (UAE), Vol. 7, No. 2.8, pp.13-17, 2018.

Classification Model Evaluation Metrics

Željko Đ. Vujović

Boulevard Save Kovačevića 20/6
81000 Podgorica
Montenegro

Abstract—The purpose of this paper was to confirm the basic assumption that classification models are suitable for solving the problem of data set classifications. We selected four representative models: BaiesNet, NaiveBaies, MultilayerPerceptron, and J48, and applied them to a four-class classification of a specific set of hepatitis C virus data for Egyptian patients. We conducted the study using the WEKA software classification model, developed at Waikato University, New Zealand. Defeat results were obtained. None of the four classes envisaged has been determined reliably. We have described all 16 metrics, which are used to evaluate classification models, listed their characteristics, mutual differences, and the parameter that evaluates each of these metrics. We have presented comparative, tabular values that give each metric for each classification model in a concise form, detailed class accuracy with a table of best and worst metric values, confusion matrices for all four classification models, and a type I and II error table for all four classification models. In addition to the 16 metric classifications, which we described, we listed seven other metrics, which we did not use because we did not have the opportunity to show their application on the selected data set. Metrics were negatively rated selected, standard reliable, classification models. This led to the conclusion that the data in the selected data set should be pre-processed to be reliably classified by the classification model.

Keywords—Classification model; classification models; evaluate classification models; worst metric values; four-class classification; metric classification; reliable classified classification models; detailed class accuracy

Subject areas—Artificial intelligence and machine learning; software engineering

I. INTRODUCTION

A specific set of data on the hepatitis C virus, consisting of 1385 instances described with 29 attributes, was considered [12]. The goal is to classify these instances into four classes, which represent hepatitis diseases: class a - Portal fibrosis, class, b - Little sepsis, class, c - A lot of sepsis, and class d - Cirrhosis.[6] This paper challenges this classification. Sources in the literature suggest that classification into five classes would be better: class a-liver inflammation, class b-fibrosis, class c-cirrhosis, class d – end-stage disease (ESLD), and class e-cancer [15].

The initial assumption is that standard, generally accepted classification models, BayesNet, NaiveBayes, Multilayer-Perceptron, and J48, are suitable for such a classification. These models exist in the WEKA software and, as such, have been applied to the selected data set. Unsatisfactory results were obtained. Available instances are classified very poorly.

That was the reason, motive, and incentive to consider why this is so? These four models were chosen at random. In this introduction, we give their generally accepted definitions.

A Bayesian network is defined as a system of event probabilities, nodes in a directed acyclic graph, in which, the probability of an event can be calculated from the probabilities of its predecessors in the graph. The nodes in the network are variable. They can be concrete values, randomly given, latent values, or hypotheses. They are characterized by the distribution of probabilities. Probability is a quantity that touches a presented state of knowledge or a state of belief. In Bayesian opinion, the probability is assigned to a hypothesis. In frequency thinking, the hypothesis is tested without assigning a probability. The result of Bayesian analysis is Bayesian inference. It updates the previous probability assigned to the hypothesis because more evidence and information have been obtained [3], [16].

Naive Bayesian classifiers are based on naive assumptions of the mutual characteristics of independence. In this way, each distribution obtained can be independently estimated as a one-dimensional distribution. This alleviates the problems arising from the "curse of dimensionality". The "curse of dimensionality" is the problematic nature of the number of variables, which can be collected from a single sample. An example of this is the need for data sets that are scaled (arranged) exponentially with many characteristics [3],[14] [16], [18].

A multilayer perceptron is defined as a system composed of a series of elements (nodes - "neurons") organized into layers. Layers process information so that they react dynamically to external inputs. The input layer has one neuron for each component, which exists in the input data. Communicates with hidden layers in the network. The entire processing of input data takes place in hidden layers. The input data are weighted (measured) by appropriate coefficients. The neuron accepts them, calculates their sum, and processes it with an activation function. It processes the processed data in a "forward" process. The last hidden layer is connected to the output layer. The output layer has one neuron for each possible output.[3], [14] ,[16], [18].

J48 is a machine learning model based on the decision tree. It was created using the ID3 algorithm (Iterative Dichtomizer 3), developed by the WEKA project development team. The decision tree presents and analyzes decision-making situations when one type of decision is derived from another type of decision. This facilitates understanding of selection problems, assessment of available versions of the decision,

and coverage of uncertain events, which affect outcomes and versions of the decision [3],[14],[16],[18].

The first idea was to consider the metrics used to evaluate the classification models used. 16 metrics used by WEKA software were reviewed, described, and explained [4]. In addition, it was stated that there are, in addition to the above, the following metrics: False discovery rate, [21] Log Loss, [22] Barrier score, [23] Cumulative gain chart, [24] Lift curve, [25] Kolmogorov-Smirnov test, [26]. These metrics were not considered because they were not contained in the WEKA software, which was used. Therefore, they could not give their ratings of the classification model on the selected data set.

The research made a significant contribution to the interpretation of the 16 mentioned metrics, elements, and parameters that each of them uses to evaluate the classification models.

A significant contribution is also the question: why did the metrics negatively evaluate the classification models used on the selected data set?

As a result of this research, other questions arose. Is the number of attributes per instance of the observed data set too large? How many attributes are needed (optimal) and what are those attributes? Is it necessary to pre-process the data of the observed set? What are the techniques for pre-processing data in a set?

Unobtrusively, the question arose as to whether the four classes for the classification of instances of the observed set were correctly determined?

II. METRICS

1) Accurately classified instances are the sum of true positive (TP) and true negative (TN).

2) Incorrectly classified instances are the sum of false positives (FPs) and false negatives (FNs).

3) Kappa statistic - Cohen's Kappa coefficient (k) is a measure of how many instances are classified model of machine learning, matched the data marked as the basic truth, controlling the accuracy of the random classifier as measured, expected accuracy. The accuracy of the Random Accuracy is $1/k$. Here k is the number of classes in the data set. In the case of binary classification $k = 2$, so the accuracy is 50%.

$$K = \frac{(p0 - pe)}{(1 - pe)}$$

$p0$ - total accuracy of the module, pe - random accuracy (random accuracy of the model).

In the problem of binary classification $pe = pe1 + pe2$; $pe1$ - the probability that the predictions agree randomly with the actual values of class 1 - "good"; $pe2$ - the probability that the predictions agree randomly with the actual values of class 2 - "accidentally". The assumption is that the two classifiers (model prediction and actual class value) are independent. In this case, the probabilities $pe1$ and $pe2$ are calculated by

multiplying the share of things in the class and the share of the predicted class.[2],[20].

4) Mean Absolute Error is the mean value of the absolute values of individual prediction errors of all instances in the test set. Each prediction error is the difference between the actual value and the predicted value for the instance.

The mean absolute error (MAE) E_i of an individual model and is calculated by the formula:

$$E_i = \frac{1}{n} \sum_{j=1}^n \left| P_{(ij)} - \sum_{j=1}^n |P_{(ij)} - T_j| \right|$$

where $P_{(ij)}$ is the value predicted by the individual model i for record j (of n records); and T_j is the target value for record j . For a perfect prediction, $P_{(ij)} = T_j$ and $E_i = 0$. Thus, the index E_i ranges from 0 to infinity, and 0 corresponds to the ideal [14] [28].

5) Root mean squared error (RMSE) - The root mean square error is relative to what it would be if a simple predictor was used. Taking the square root of the relative square error, the error is reduced to the same dimensions as the predicted size.

The root mean square error (RMSE) E_i of an individual model and is calculated by the formula:

$$E_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (P_{(ij)} - T_j)^2}$$

Where $P_{(ij)}$ is the value predicted by the individual model i for record j (of n records), and T_j is the target value for the record j . For a perfect prediction, $P_{(ij)} = T_j$ and $E_i = 0$. Thus, the index E_i ranges from 0 to infinity, and 0 corresponds to the ideal.[27].

6) Relative absolute error (RAE) is the total absolute error and normalized by dividing by the total absolute error of the simple predictor (ZeroR classifier). The relative absolute error E_i of an individual model is evaluated by the equation:

$$E_i = \frac{\sum_{j=1}^n |P_{(ij)} - T_j|}{\sum_{j=1}^n |T_j - \bar{T}|}$$

Where $P_{(ij)}$ is the value predicted by the individual model i for record j (of n records); T_j is the target value for record j , and \bar{T} is given by the formula:

$$\bar{T} = \frac{1}{n} \sum_{j=1}^n T_j$$

For a perfect prediction, the counter is 0 and $E_i = 0$. Thus, the index E_i ranges from 0 to infinity, and 0 corresponds to the ideal.

A good prediction model produces a near-zero ratio. A bad model (one that is worse than a naive model) will produce a ratio greater than one x100%.[27].

7) Root relative squared error (RRSE) reduces the error to the same dimensions as the predicted size. Relative square error is the total square error divided by the total square error of a simple predictor. The root of the relative square error E_i of an individual model j is calculated by the formula:

$$E_i = \sqrt{\frac{\sum_{j=1}^n (P_{(ij)} - T_j)^2}{\sum_{j=1}^n (T_j - \bar{T})^2}}$$

Where $P_{(ij)}$ is the value predicted by the individual model i for record j (of n records). For perfect prediction, the counter is equal to 0 and $E_i = 0$. The index E_i ranges from 0 to infinity, and 0 corresponds to the ideal [28].

8) Confusion matrix for a binary classifier (Fig. 1). Actual values are marked True (1) and False (0), and are predicted as Positive (1) and Negative (0). Estimates of the possibilities of classification models are derived from the expressions TP, TN, FP, FN, which exist in the confusion matrix [10].

Class designation		Actual class	
		True (1)	False (0)
Predicted class	Positive (1)	TP	FP
	Negative (0)	FN	TN

Fig. 1. Confusion Matrix for the Binary Classification Problem [7].

TP (True Positive) - The data point in the confusion matrix is True Positive (TP) when a positive outcome is predicted and what happened is the same.

FP (False Positive) - The data point in the confusion matrix is false positive when a positive outcome is predicted, and what happened is a negative outcome. This scenario is known as a Type 1 Error. It is like a boon in bad foresight.

FN (False Negative) - The data point in the confusion matrix is false negative when a negative outcome is predicted, and what happened is a positive outcome. This scenario is well known as a Type 2 Error and is considered as dangerous as a Type 1 Error.

TN (True Negative) - The data point in the confusion matrix is True Negative (TN) when a negative outcome is predicted and what happens is the same. The results of the binary classification shown in Fig. 2.

Confusion matrix for four-class classification (Fig. 3). Four-class classification is a problem of classifying instances (examples) into four classes. Case of four classes: class A, class B, class C, and class D [13],[17].

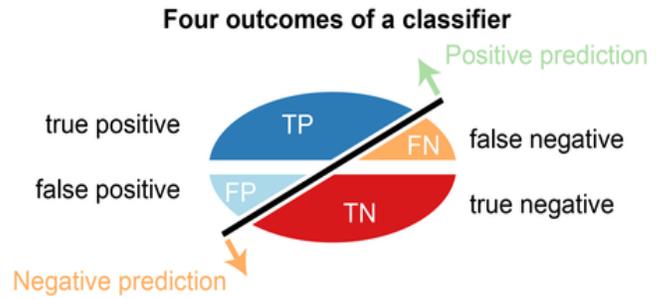


Fig. 2. Elliptical Representation of Four Binary Results of the Test Set Classification [7].

		Actual value			
		A	B	C	D
Predicted value	A	100	0	0	0
	B	80	9	1	1
	C	10	0	8	0
	D	10	1	1	9

Fig. 3. Confusion Matrix for the Four-class Classification Problem [8].

9) Accuracy is calculated as the sum of two accurate predictions (TP + TN) divided by the total number of data sets (P + N). The best accuracy is 1.0, and the worst is 0.00 (Fig. 4) [19].

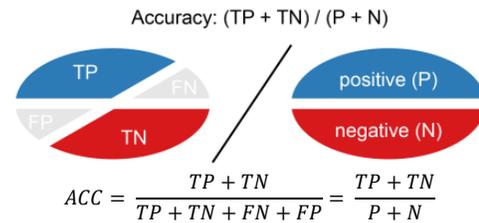


Fig. 4. Two ellipses show how accuracy is calculated [7],[11].

10)TP Rate - True Positive Rate (Sensitivity or Recall) is calculated as the number of accurate positive predictions (TP) divided by the total number of positive (P). Also called Sensitivity or Recall (REC). The best TP Rate is 1.0 and the worst 0.0 (Fig. 5) [19].

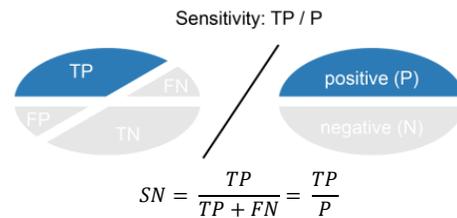


Fig. 5. Two Ellipses Show How the Sensitivity is Calculated [7].

11)FP Rate - False Positive Rate is calculated as the number of false-positive predictions (FP) divided by the total number of negatives (N). The best false positive rate is 0.0 and the worst is 1.0. It can also be calculated as 1-specificity (Fig. 6) [19].

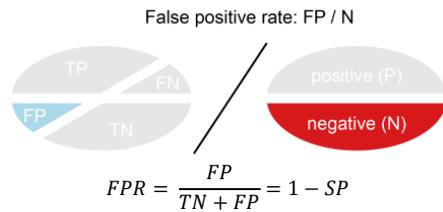


Fig. 6. Two Ellipses Show How the False Positive Rate - FPR is Calculated [7].

12) Precision is calculated as the number of correct positive predictions (TP), divided by the total number of positive predictions (TP + FP). The best accuracy is 1.0 and the worst 0.0 (Fig. 7) [19].

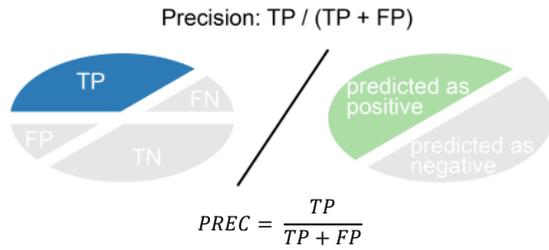


Fig. 7. Two Ellipses Show How Precision is Calculated [7],[11].

13) True Negative Rate – TNR (Specificity) - is calculated as the number of correct negative predictions (TN) divided by the total number of negatives (N). The best specificity is 1.0 and the worst 0.0 (Fig. 8) [19].

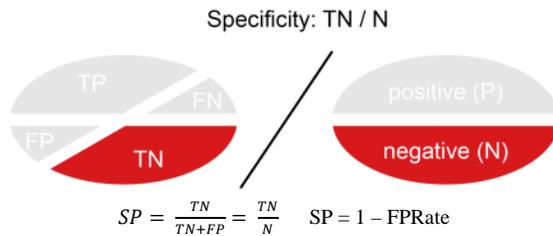


Fig. 8. Two Ellipses Show How Specificity (SP) is Calculated [7].

14) F-Measure or F-score is a measure of the accuracy of the test. It is calculated, based on precision and reminders, by the formula:

$$F\text{-Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} [7],[11],[19]$$

15) Matthews Correlation Coefficient (MCC) - is the correlation between the predicted classes and the basic truth. It is calculated based on the values from the confusion matrix.

$$MCC = \frac{TP}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

MCC is generally considered a balanced measure, which can be used even if the classes are of very different sizes [7],[11],[19].

16) ROC Area - Receiver Operating Characteristic Area - The ROC curve is a graph that visualizes the trade-off between True Positive Rate and False Positive Rate (Fig. 9)

For each threshold, we calculate True Positive Rate and False Positive Rate and plot them on one graph. The higher the True Positive Rate and the lower the False Positive Rate for each threshold, the better. Better classifiers have more curves on the left. The area below the ROC curve is called the ROC AUC score, a number that determines how good the ROC curve is [11].

The ROC AUC Score shows how good the model is in ranking predictions. Indicates the probability that a randomly selected positive instance is ranked higher than a randomly negative instance [7],[19].

17) PRC Area (Precision-Recall Curve Area) It is one number that describes the capabilities of the model. The PR AUC Score is the average of the precision scores calculated for each reminder threshold [0,0, 1,0]. The PRC curve is obtained by combining Positive Predictive Value and True Positive Rate (Fig. 10). For each threshold, Positive Predictive Value and True Positive Rate are calculated and the corresponding point of the graph is plotted. Preferably, the algorithm has high precision and high sensitivity. These two metrics are not independent. That is why a compromise is being made between them. A good PRC curve has a higher AUC. Research has shown that PRC is graphically more informative than ROC graphs when estimating binary classifiers on unbalanced sets [5],[9],[19].

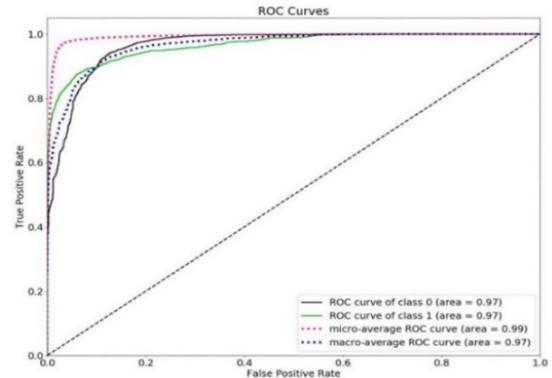


Fig. 9. ROC Curve [1],[5].

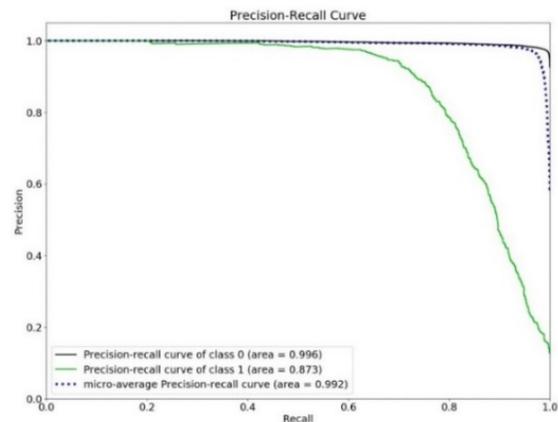


Fig. 10. Precision-Recall Curve [9].

III. EXPERIMENTAL RESULTS

The best and worst values of each general metric are used to measure the accuracy of the classifier. Metrics are in rows and values are in columns of the table.

Summary of the accuracy of the four representative classifiers expressed by general metrics. Metrics are listed in the rows of the table, and their values, for each classifier, in the columns of the table. A special number is the total number of instances, which is the same for each classifier.

The detailed accuracy of each of the four representative classifiers for each of the predictions of the class is expressed by the values of eight different metrics. Metrics are in the columns of the table, the names of the classifiers in the rows of the table, separately for each class. For each class, the weighted value of each of the eight metrics is shown. This value is the average that results from multiplying each component by a factor that reflects its significance.

TABLE I. METRICS SUMMARY

	Bayes Net	Naïve Bayes	Multilayer Perceptron	J48
Correctly classified instances	318	362	368	350
Incorrectly classified instances	1007	1023	1017	1035
Kappa statistic	-0,0287	0	0,0206	0,0029
Mean absolute error	0,3763	0,3748	0,3718	0,3751
Root mean squared error	0,4393	0,4329	0,5466	0,5814
Relative squared error	100,382%	99,9999%	99,2009%	100,0671%
Root relative squared error	101,4822%	100%	126,2575%	134,2938%
Total number of instances	1385	1385	1385	1385

TABLE II. TABLE OF BEST AND WORST METRIC VALUES FOR DETAILED CLASS ACCURACY

	The Best	The Worst
TP Rate	1,0	0,0
FP Rate	0,0	1,0
Precision	1,0	0,0
Recall	1,0	0,0
F-Measure	1,0	0,0
MCC	+1,0	0,0
ROC Area	0,9	0,5
PRC Area	1,0	0,5

TABLE III. DETAILED ACCURACY BY CLASS

		TP Rate	FP Rate	Precision	Recall	F.Measure	MCC	ROC Area	PRC Area
c	BayesNet	0,107	0,186	0,156	0,107	0,127	0,091	0,423	0,205
l	NaiveBayes	0,000	0,000	?	0,000	?	?	0,496	0,241
a	M.L.Perc.	0,193	0,254	0,196	0,193	0,195	0,060	0,453	0,220
s	J48	0,250	0,236	0,253	0,250	0,251	0,014	0,501	0,247
s(1)	Weight Av.	0,230	0,250	0,222	0,230	0,224	0,032	0,473	0,243
c	BayesNet	0,271	0,270	0,241	0,271	0,255	0,001	0,510	0,249
l	NaiveBayes	0,000	0,000	?	0,000	?	?	0,496	0,238
a	M.L.Perc.	0,577	0,236	0,271	0,277	0,274	0,041	0,527	0,249
s	J48	0,271	0,226	0,274	0,271	0,273	0,045	0,526	0,252
s(2)	Weight Av.	0,261	0,261	?	0,261	?	?	0,496	0,249
c	BayesNet	0,214	0,266	0,217	0,214	0,216	-0,052	0,457	0,234
l	NaiveBayes	0,000	0,000	?	0,000	?	?	0,496	0,255
a	M.L.Perc.	0,282	0,247	0,282	0,282	0,282	0,035	0,521	0,278
s	J48	0,231	0,245	0,246	0,231	0,238	-0,0014	0,488	0,248
s(3)	Weight Av.	0,266	0,245	0,265	0,266	0,066	0,013	0,509	0,257
c	BayesNet	0,320	0,307	0,270	0,320	0,293	0,013	0,524	0,281
l	NaiveBayes	1,000	1,000	0,261	1,000	0,414	?	0,496	0,260
a	M.L.Perc.	0,307	0,243	0,308	0,307	0,307	0,063	0,533	0,280
s	J48	0,260	0,290	0,240	0,260	0,250	0,030	0,476	0,255
s(4)	Weight Av.	0,253	0,250	0,253	0,253	0,253	0,003	0,497	0,251

TABLE IV. CONFUSION MATRIX

BayesNet				NaiveBayes				M.L.Perceptron				J48				
a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	
36	94	94	112	0	0	0	336	65	94	86	91	84	82	79	94	a= 336
61	90	85	96	0	0	0	332	79	92	86	75	67	90	80	95	b= 332
79	94	76	106	0	0	0	355	96	76	100	83	80	82	82	111	c = 335
55	96	95	111	0	0	0	362	91	78	82	111	101	74	93	94	d= 362

TABLE V. TYPE I ERRORS AND TYPE II ERRORS

	BayesNet				NaiveBayes				M.L.Perceptron				J48			
	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
Error Type I	309	215	279	246	336	332	255	0	271	240	355	251	255	242	255	270
Error Type II	195	284	277	314	0	0	0	1.033	266	248	234	249	248	238	852	300

Comparative table of four confusion matrices for all four representative classifiers. In the rows, the number is provided for each class, and in the columns the actual value of the class.

Comparative table of Type I and Type II error values for each class and each representative classifier. There are types of errors in the rows, and their size in the columns.

IV. DISCUSSION

The average value of correctly classified instances is 25.24%, and incorrectly classified instances 73.68% (Table 1).

Landis and Koch proposed the following standards for the kappa coefficient: ≤ 0 = poor, $.01 - .20$ = insignificant, $.21 - .40$ = fair, $.41 - .60$ = moderate, $.61 - .80$ = substantial, $.81 - 1$ = almost perfect [29]. In line with the above proposal, BayesNet and NaiveBaies have a poor kappa coefficient, and multilayer perceptron and J48 are negligible. It is concluded that the values of the kappa coefficients show that the instances, classified by the machine learning model, do not match the data marked as the basic truth. MAE values: 0.3763 for BaiesNet, 0.3748 for NaiveBaies, 0.3718 for MultilayerPerceptron, 0.3751 for J48 are closer to the lower limit (ideal) than the upper (worst). We, therefore, appreciate that they are acceptable (Table 1).

Anthony Ladson gave a model performance table based on the efficiency coefficient. For the case of model performance validation, the values of the efficiency coefficients describe the classification as follows: $E \geq 0.93$ - excellent, $0.8 \leq E < 0.93$ - good, $0.6 \leq E < 0.8$ - satisfactory, $0.3 \leq E < 0.6$ - transient, $E < 0.3$ - bad [30]. Based on this, values of 0.4393 for BayesNet, 0.4329 for NaiveBayes, 0.5466 for MultilayerPerceptron, and 0.5814 for J48 are in the transient group (Table 2).

Relative absolute error (RAE) can have values from 0 to infinity. Ideally, it should have a value of 0. Based on this, it is concluded that the values of 100.3802% for BaiesNet, 99.999% for NaiveBaies, 99, 2009% for MultilayerPerceptron, and 100.0671% for J48 are approximately the same as in the naive model (ZeroR classifier). The root of the relative square error (RRSE) can have a value from 0 to infinity. Ideally, it should have a value of 0. RRSE values: 101.4822% for BayesNet, 100% for NaiveBaies, 126.2775% for MultilayerPerceptron, and 134.2938% for J48 rate NaiveBaies

as a naive model, and BayesNet, MultilayerPerceptron and J48 worse than naive (Table 1).

Analysis of the detailed accuracy of the classes (Table 1 and Table 2) shows very significant results. Based on the tables of best and worst metric values for detailed class accuracy, we conclude:

1) TP Rate has extremely poor values, close to the worst, for all rated models and all classes. The exception is NaiveBaies, which has the best value of 1,000 for class 4, but the same NaiveBayes has the worst value of TP Rate, 0,000, for classes 1,2, and 3. Relatively good value of TP Rate, 0,577, showed MultilayerPerceptron for class 2. Weighted values TP Rates are consequently poor.

2) FP Rate for NaiveBayes has an optimal value of 0.000 for classes 1,2 and 3, as opposed to class 4 for which it has a maximum value of 1000. BayesNet, MultilayerPerceptron, and J48, as well as a weighted value for all four models, and all four classes are extremely bad.

3) Precision has values below a level satisfactory for all four models.

4) Recall, has the same values as TP Rate. The question is why are they separated for display in a separate column?

5) The F-Measure has values that are below levels that meet all rated models and all four classes.

6) MCC showed unsatisfactory values, which are at the level of random prediction, for all evaluated models and all classes.

7) The ROC Area showed values for all models and all classes that are on the verge of bad.

8) The value of the PRC area, for all models and all classes, is below the level that is the worst.

The metrics of detailed assessment by classes unequivocally show that the evaluated models, applied in a presented way, do not satisfy (Table III). This means that new research is needed and the answer to the question: why do metrics of detailed accuracy give poor estimates of the models used?

By comparative analysis of the confusion matrix for all four classification models and all four classes, we see that the predictions of true positive results (TP) are not good enough

REFERENCES

(Table 4). Type I and type II errors are relatively high. The goal of modeling is to reduce these errors to minimum values.

Separate consideration of type I and type II errors for the four applied models shows that NaiveBayes has a type I error value equal to 0, for class d, and type II errors for classes a, b, and c (Table 5). These data further problematize the use of this model. For the other three models, the type I and type II errors are, on average, 2.5 times larger than exactly predicted.

V. CONCLUSIONS

In this paper, we have considered in detail the 16 metrics for the evaluation of classification models, which exist in WEKA software, version 3.4.1., Developed at the University of Waikato, New Zealand. The consideration is in line with the initial assumption of the paper that classification models are suitable for solving the classification problem applied to a specific set of hepatitis C virus data for Egyptian patients.

In addition to the above 16 metrics, we found in the literature that there are other metrics: False discovery rate, Log Loss, Barrier score, Cumulative gain chart, Lift curve, Kolmogorov-Smirnov plot, and Kolmogorov - Smirnov statistics. We did not describe them because we were unable to demonstrate their application to the data set we selected. These metrics remain for display in a later review paper.

All metrics considered negatively evaluated the classification models, which we used. This has led to doubts because these are models that are generally accepted as standard and reliable. Why, metrics, do they rate them negatively on a selected data set? Is the number of attributes in the selected data set too large? How many attributes are needed and what are those attributes? Is it necessary to pre-process the data of the selected set?

The special significance of this paper is that it highlights the multitude of metrics used to evaluate each classification model. It emphasizes the diversity of these metrics and the parameters they measure to better understand the model and its features.

New questions and problems, which arose from this paper, are: What are the techniques for pre-processing data in a data set, and how should discretization, purification, reduction, and discussion of data be performed in a specific hepatitis C virus data set for Egyptian patients?

We suggest that the classification be performed in five classes, as provided in the latest professional literature: class a-inflammation of the liver, class b-fibrosis, class c-cirrhosis, class d – end-stage disease (ESLD), and class e-cancer.

ACKNOWLEDGMENT

To the editor and reviewers of IJACSA - The Science and Information (SAI) Organization. To Dejan Vujović, an engineer for the development and maintenance of application software in the Montenegrin Electricity Transmission System Podgorica.

[1] T.Fawcett, „ROC Graphs: Notes and Practical Considerations for Researchers.” Kluwer Academic Publishers, 2004.

[2] J.Sim, C.C.Wright, „The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements.” *Physical Therapy*, Volume 85, Issue 3, Pages 257 -68, 2005. <https://doi.org/10.1093/ptj/85.3.257>.

[3] J.Đ.Novaković, „Rešavanje klasifikacionih problema mašinskog učenja.” Fakultet tehničkih nauka u Čačku, 2013.

[4] R.R:Bouckaert, E. Frank, M. Hall, R. Kirkby, R.Reutmann, A. Sewald, A., D. Seuse, „WEKA Manual for Version 3-7-8.”, 2013.

[5] T. Saito, M. Rechmsmeier, „The Precision-Recall Plot is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets.”, *PLoS ONE*, 2015 doi: 10.1371/journal.phone.0118432.

[6] M.Nasr, K. Elbahanacy, M. Hamdy, S.M.Kamal, „A novel model based on non-invasive methods for prediction of liver fibrosis.”13th International Computer Engineering Conference (ICENCO), 2017.

[7] T. Saito, M. Rehmeismeier, „Basic evaluation measures from the confusion matrix.” WordPress, 2017.

[8] V. Leal, „How to build a confusion matrix for a multiclass classifier?” CrossValidated, StackExchange Inc, 2021.

[9] S. Auckland, S., „Precision-recall curves-what are they and how are they used.” Acutecuretesting, 2017.

[10] S. Narkhede, „Understanding Confusion Matrix.” Towards Data Science, 2018.

[11] A. Mishra, „Metrics to Evaluate your Machine Learning Algorithm.” Towards Data Science, 2018.

[12] D. Dua and C. Graff, „UCIMachineLearning Repository [<http://archive.ics.uci.edu/ml>].” Irvine, CA: The University of California, School of Information and Computer Science.Hepatitis C Virus (HCV) for Egyptian patients Data Set, 2019.

[13] A.Iqbal, A, Aftab2, S, Ali3, U, Nawaz4, Z, Sana5, L, Ahmad6, M, Husen7, A., „Performance Analysis of Machine Learning Techniques on Software Defect Prediction using NASA Datasets.” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 5, 2019.

[14] R.Delgrado, X-A, Tibau „Why Cohen’s Kappa should be avoided as a performance measure in classification.” *PLoS One*;14 (9), 2019.

[15] J.S.Saladi, „What Are The Stages of Liver Failure?.” Healthline, 2019.

[16] Ž. Vujović, „The Big Data and Machine Learning.” *Journal of information technology and multimedia systems*, Vol. 19, Issue 7. pp.11-19, DOI: 10.5281/zenodo.427923, 2020.

[17] S. Nandacumar, „Confusion Matrix – are you confused? (Part I and Part II).” Medium, 2020.

[18] A. Albahr1, M. Albahar2, „An Empirical Comparison of Fake News Detection using different Machine Learning Algorithms.” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 9, 2020.

[19] A. Tharwat, „Classification assessment methods.” *Applied Computing and Informatics*, Volume 17, Issue 1, 30, 2020.

[20] M. Widmann, „COHEN’S KAPPA: What It Is, When to Use It, and How to Avoid Its Pitfalls.” *The New Stack*, 2020.

[21] S. Room, „False Discovery Rate (FDR).” In Dubitzky W., WolkenHauer O., Cho KH., Yokota H., (eds) *Encyclopedia of Systems Biology*, Springer New York, NY, doi:10.1007/978-1-4419-9863-7_223, 2013.

[22] G. Dembla, „Intuition behind Log-loss score.” Towards Data Science, 2020.

[23] J. H. Orallo, P.A. Flach, C.Ferri, „Brier curves: a new cost-based visualization of classifier performance.” *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*. pp. 585–592, 2011.

[24] T. Jurczyk, „Gains vs ROC curves. Do you understand the difference?” TIBCO® Data Science, 2020.

- [25] D.S. Coppock, „Data Modelling and Mining: Why Lift?“ DM Review and Source Media, Inc., 2006.
- [26] A. Justel, D. Pena, R. Zamar, „ A multivariate Kolmogorov -Smirnov test of goodness of fit“ *Statistics&Probability Letters*, Volume35, Issue3, Pages251-259, 1997, doi: 10.1016/S0167-7152(97)00020-5.
- [27] S.Glen, „Mean Squared Error. Definition and Example.“From StatisticHowTo.com:Elementary Statistics for the rest of us! <https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/mean/squared-error/>, 2021.
- [28] Root Relative Squared Error.“GeneXproTools Online Guide, Gepsoft. Ltd., 2000-214.
- [29] L.Hartling, M.Hamm, A.Milne, et al. „Interpretation of Fliess’kappa (k) (from Landis and Koch 1977).“ *Valiability and Inter-Rater Reliability Testing of Quality Assessment Instruments [Internet]*, Rockvile: Agency for Healthcare Research and Quality (US), 2012.
- [30] A.Ladson, „Model performance based on the coefficient of efficiency.“ *Hidrology, Natural Resources, and R*, (2019).

Cluster based Certificate Blocking Scheme using Improved False Accusation Algorithm

Chetan S Arage¹, K.V.V. Satyanarayana²

Department of CSE, Koneru Lakshmaiah Education Foundation
Vaddeswaram, Guntur – 522502, Andhra Pradesh, India

Abstract—The aggregation of mobile nodes without the use of a base station is known as Mobile Ad Hoc Networks (MANETS). In nature, the nodes are moving. These networks are not connected and thus subject to security attacks due to their mobility. There are several mechanisms proposed to prevent mishaps while routing of the packets in such networks methods: The methodology outlined in Mobile Ad Hoc Networks to protect against various types of assaults is based on a recent method known as Cooperative Bait Detection Scheme. Its implementation scenario demonstrates that in the event of Sybil assaults, the packet delivery ratio and performance are low on the network. Our goal is to propose a cluster-based methodology to improve delays, packet delivery ratio, and other performance assessment criteria. Improved Cooperative Bait Detection recommends a disjointed multipath technique to avoid attacks. Until date, the dropped packet delivery ratio has been the key to preventing collaborative and Sybil assaults. In the Hybrid Cooperative Bait Detection Scheme, nodes are verified in two stages: first, on the basis of packet delivery ratio, and then, in the second stage, the exact cause of performance decline is explored to check node behavior. In order to improve security, certifying procedures must be used to clustered networks. For malevolent entities, the false accusation algorithm provided certificate revocation and blocking approaches. An algorithm is proposed that remembers false accusations for a set period of time in order to increase the number of normal nodes in the network and hence improve the system's performance. Results: With the help of NS2 simulation, the clustering approach was evaluated by considering several Sybil-attack network scenarios. When the proposed work is compared to other ways such as Cooperative Bait Detection Scheme, Improve Comparative Bait Detection Scheme, and Hybrid Comparative Bait Detection Scheme, the results show that Packet Delivery Ratio and performance are improved for Sybil attackers over the internet. In conjunction with Certifying authority, the Cluster Head in the network identifies and prevents false complaints. The results of the comparison using several performance parameters reveal that the new strategy outperforms the existing ones. As the number of normal nodes in the system grows, the system will be able to work at its best, preventing various types of attacks.

Keywords—Mobile Ad Hoc Networks; cooperative bait detection scheme; cluster; cluster head; certifying authority; certificate revocation

I. INTRODUCTION

The AdHoc Mobile Network is a less dependable wireless network for the infrastructure. [1]. Since the 1980s wireless mobile systems have been in use. We saw their developments in wireless systems of the first, second, and third generations.

With the assistance of a centralized support structure like an access point, wireless networks function. These access points help wireless users to maintain connections from one area to the other with the wireless system. The existence of a permanent support structure inhibits wireless solutions' adaptability [2]. So Bluetooth provided a new sort of wireless system called ad hoc mobile networks to develop wireless networks (MANETS).

Mobile Ad Hoc Networks (MANETS) is the aggregation without the base station of mobile nodes. The nodes are moving in nature. Due to its mobility, these networks are not wired yet vulnerable to security assaults. There are several mechanisms proposed to prevent mishaps while routing of the packets in such networks. MANET is free to roam in any direction and consequently regularly changes its connections to other devices. The goal of connecting "everywhere and every time," mobile Adhoc networks can make true. Wireless ad hoc mobile network is usually shown as Fig. 1.

A. Misbehaving Nodes in MANETS

A mobile node might be dubbed a selfish or misbehaving node to gain from other nodes. This node produces its own network connection certificate and attempts to communicate with other genuine network nodes [3]. It also acts like denying or not responding to other nodes to forward data packets to preserve its battery power.

B. Blocking of Misbehaving Node

With an unsafe node, secure communications between network nodes are interrupted. MANETS utilizes weighted bunching strategies to construct geography. Hubs inside the organization structure a bunch that is a group head (CH) along with certain group individuals (CM)[4]. Security is the critical need of MANET in light of the versatility of the hubs [3].

Fig. 2 shows the MANET working using the clustering approach. For overall study it finds that there is the number of activities in mobile Adhoc network but due to its ad-hoc nature number of unknown activity can occur which causes the problem in the network that indirectly affects the working and throughput. So it is expected to protect the adhoc network by detecting the unknown activity so can improve the performance without affecting the network communication.

The following section gives a detailed overview of each attack with the proposed approach to detect the malicious activity.

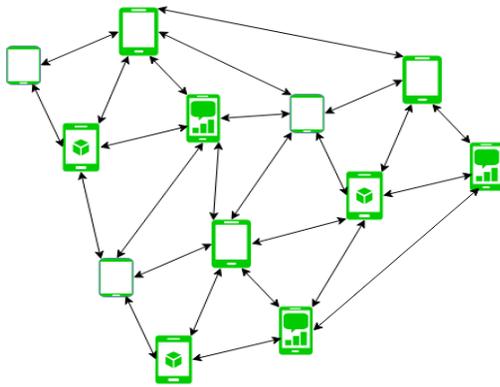


Figure - Mobile Ad Hoc Network

Fig. 1. MANET Environment.

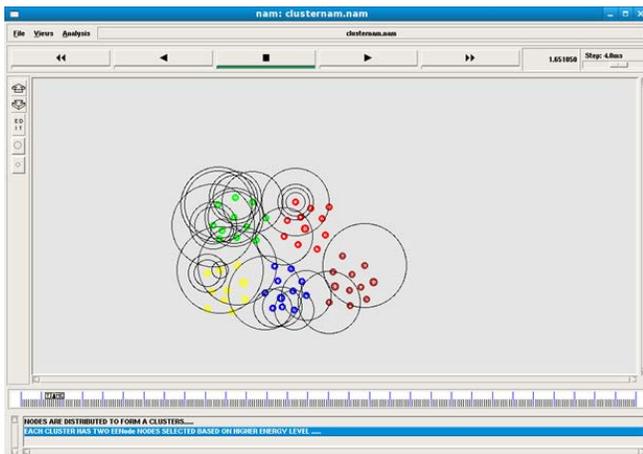


Fig. 2. MANET Working with Clustering.

Section II gives the details of MANET attacks and related works. Section III discusses the proposed approach use in-network to secure communication between the nodes. Section IV shows the simulation outcome followed by the conclusion and future scope.

II. RELATED WORK AND CLASSIFICATION

This session gives a details overview of attack and review in mobile adhoc network.

A. Data Traffic Attack

DATA traffic attacks either in nodes that discard data packets or in data packets that hold forwarding. A couple of forms of assaults are picked for dropping victim packets, while others drop them altogether independent of sender nodes. This can increase service end to end and decrease service value. It can also lead to considerable loss of data. Moreover, few nodes might be completely out of the way save for a redundant path near the irregular node.

B. Control Traffic Attack

MANET is naturally susceptible to attack because of its primary features, like open medium, conveyed hubs, self-governance for network investment (can connection and leave the organization as indicated by its will), absence of focal position (which can force net wellbeing on the organization),

disseminated coordination and communitarian activities [6]. Because of these reasons, the present routing protocols in MANET cannot be used. MANET has many different routing protocols with their individual characteristics and set of legislation [7]. The DYMO is a fast light routing technology created for Multi-Hop networks that depend upon the cooperation of each individual Node in defining the genuine routing table. But they all rely on confidence in networked nodes [14]. The first stage for a successful attack is that the node is a network component. Since there are no constraints on network membership, the hostile node can be connected to and disrupted by capturing routing tables or by dodging lawful routes.

In addition, if the node can determine itself as the quickest route to any target, it may spy on the network using unsafe routing protocols. That is why the protocol for routing must be maximally safeguarded [17]. Similar assaults are not CONTROL assaults and may be mitigated in physical security mechanisms, especially for jamming assaults.

A recent method known as Cooperative Bait Detection Scheme is the basis of the methodology described in [3] Mobile Ad Hoc Networks to defend against various forms of assaults. Its implementation scenario reveals that the delivery ratio and performance of the packet are low in the event of Sybil assaults on the network Improved Cooperative Bait Detection suggests disjoint multipath strategy to prevent assaults. Dropped Packet delivery ratio is the key for implementation to prevent collaborative and Sybil assaults till now.

In Hybrid Cooperative Bait Detection Scheme two stage verification of nodes is done where node which found malicious in first stage on the basis of packet delivery ratio, later stage exact cause of performance drop is investigated to check behavior of the node [3].

III. PROPOSED WORK

The below session explain the proposed work use in the simulation approach with malicious activity detection to improve the throughput of the overall network.

It is required to apply certifying mechanisms to clustered networks so as to enhance security measures.

Our goal is to propose a cluster-based methodology to improve delays, packet delivery ratio, and other performance assessment criteria.

The fundamental goal of our proposed work is to combine a cluster-based certificate blocking mechanism with a better false accusation algorithm, which will be discussed later in this section.

Therefore, we updated the CBDS method with the proposed approach to the clustering system. The number of hops of all nodes from the destination is considered [22]. The node receives the packet and only transfers it from that dedicated path in accordance with the CBDS scheme, otherwise, the packet will be discarded. As explained earlier this is carried out before clustering is done by integrating the CBDS algorithm with the proposed false accusation algorithm.

A. Algorithm Design

Assumptions and Abbreviations

CA=Certifying authority

CH= Cluster Head

WL= Warned List

BL= Blocked List

CH=Cluster Head

NN = Number of Nodes

SNi = Specific Node,

SCK = Specific Cluster,

MAX = Number of maximum nodes per cluster,

PD = Packet Data,

NBi = Buffer of Specific Node,

SNTCi = Specific node's trust counter (Initially set to 0 for each node),

SNTSi = Specific node's trust status (Initially set to 'F' for each node)

TV = Threshold Value (Set default is 0.8)

k = Used for Cluster Number

count = Used to count number of nodes in cluster

Algorithm

Initialize

k = 0

count = 0

Step 1: Cluster formation

for i = 1 to NN

{

Find out degree of each node ();

Find out power status of each node ();

}

while (making cluster by putting every node in to any one cluster)

{

if (SNi = = max (degree of node and power status of node))

{

Add SNi into SCi

Set count = count +1

}

if (count > MAX)

{

k++

Set count =0

}

}

Step 2: Detection of suspected nodes

while (SNTCi < TV)

{

if (SNi sends accusation message against node M = True)

{

CA updates WL and BL

}

else

{

CH sends recovery packet to CA

CA broadcast this information

SNTCi = SNTCi + 0.2

}

}

Step 3: Process for suspected nodes

Send Testing packet data RREQ to the node with TTL=1

if (receives response)

{

if (SNTCi < TV)

Set SNTSi as 'F'

}

else

{

Set SNTSi as 'D'

Go to step 4

}

Step 4: Detection of false accusation ();

Step 5: Send accusation message for time t

False accusation algorithm proposed certificate revocation and blocking techniques for malicious entities. An algorithm proposed remembers false accusation for certain amount of time to achieve increased number of normal nodes in the network and hence improves performance of the system.

B. Environment used in Proposed Approach

The network environments of 1000 m * 1000 m with various numbers of nodes are seen in Table I below. In addition, the suggested phenomenon was tested against malevolent situations in which the intruders were infected by a variety of legal nodes. In existence, the CUs became movable, where they could at any moment break from their network or combine. In addition, 802.11 was the underlying MAC layer standard, although the routers' transmission range was restricted to 250 m [14]. To quantify the protection, during the handoff and communication process, the malevolent nodes or CUs were inserted into the environment using the probability distribution.

TABLE I. PARAMETERS USED IN AD HOC NETWORK FOR SIMULATION

Simulation Time	1000 seconds
Number of nodes	50
Number of Malicious Nodes	0,5,10,15,20 (Scenario)
Network Size	1000 m X 1000 m
Transmission Range	250 m
Maximum Speed	1 m/s – 10 m/s
Mobility Model	Random Way Point
Traffic Type	CBR
Number of Source Destination Pairs	30 %
IFQ Size	NS2 Default
Channel Bandwidth	2 Mbps

C. Parameters

In our experimentation, there are multiple inputs and output attributes. The input features are the number of nodes and the time of arrival. The proposed algorithm would follow all scheduling requirements such as overall performance, packet transmission rate, reduced return time, minimum waiting time, minimum power usage, and minimum end-to-end delay in multiple cases of arrival time [12]. The algorithm will then be implemented to test the proposed process. In each case of our experiment for the assessment of results, we consider different performance measurements; some of the normal performance parameters are,

1) *Throughput*: It is a network computing efficiency parameter for the supply of packets of data from the source to the destination. This attribute reflects network performance and is crucial.

2) *End to end delay*: It is defined as the total amount of all plausible delays made due to buffering when the task of the route discovery process is ongoing and completed.

3) *Packet delivery ratio*: it gives the ratio of the total amount of packets delivered to the total amount of packets lost. This parameter signifies the efficiency of transmission.

Evaluate the proposed approach, namely improve energy-efficient resource allocation (IEERA) in cognitive radio networks using clustering. The other variables that influence this are transfer & propagation delays, transmission delays, interface queue, etc. output in various network scenarios in

terms of node movement rate, multicast group size, resource assignment group number [13]. The number of multi-cast destinations is calculated between 5 and 20 radio nodes in this simulation. The amount of traffic is 10 packets per second.

D. Energy Efficiency

Energy performance is one of the key problems in the architecture of wireless sensor node MAC protocol. In MAC-layer protocols, diverse sources contribute to energy conservation. The first energy waste source is a crash, triggered by two or more sensor nodes concurrently transmitting. The need to re-transmit a broken packet increases the consumption of electricity. The second explanation for energy depletion is lazy listening. When you hear traffic that is not being sent, a sensor node enters this mode. In many sensor network implementations, this energy-consuming silent channel monitoring can be high. The third source of energy waste is overheard when a sensor node collects packets for other nodes.

Energy Efficiency= Energy utilized by node / Total energy of the node

E. Congestion Control

Congestion takes place when traffic approaches the combined or total potential of the underlying networks. Therefore, more modern methods to eliminate, track and overcome congestion must be built-in special con-side rations. When designing certain strategies for maximal performance, the finite resources of the WSN need to be considered. Diverse techniques, including protocols for routing aided by congestion detection and control mechanisms and complex protocols for congestion control, have been adopted in the last few years.

IV. SIMULATION AND RESULTS

Cluster-based bait detection system performance review of the proposed system model in the Ad hoc Network is tested and contrasted with Boost DSR, CBDS, and HCBDS routing system based on some parameters successful the stable routing protocol applied in the presence of Sybil attack.

The clustering strategy with the help of NS2 simulation and assessed by taking into consideration various Sybil-attack network situations. The outcomes of the planned work is compared with Cooperative Bait Detection Scheme), Improve Comparative bait detection Scheme & Hybrid Comparative bait detection Scheme approaches indicate that PDRs & performance are being enhanced for Sybil attackers over the internet as compared to other existing approaches. Cluster Head in the network detects and prevents fake allegations in association with Certifying authority. The comparative result with different performance parameters shows that the proposed approach gives a better outcome as compare to the existing ones. As the number of normal nodes in the system is increased the system is able to perform at its best with achievable prevention to different kind of assaults.

Performance Analysis of proposed system model namely Cluster basted is evaluated and compared with conventional DSR and CBDS based on parameters like false positives, detection rate, energy consumption, packet loss rate. Following is some Performance analysis of cluster-based approach with proper procedure.

A. False Positives

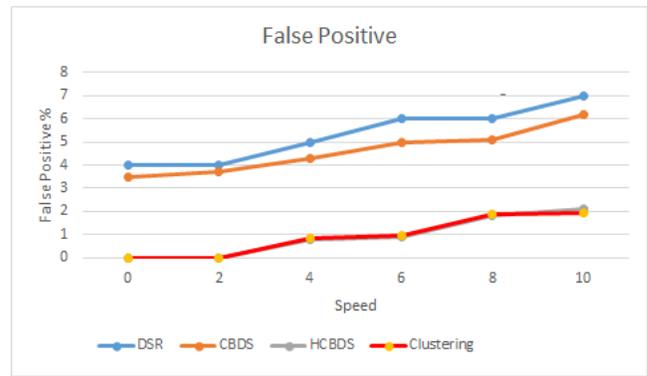
The False Positives Sum is the ratio of legal nodes considered unsafe to the total legal node numbers. We contrasted the architecture proposed with the DSR, CBDS, HCBDS system in the terminology of positives that are false, this time gripping all the frameworks together with our model of optimization, and incorporating nodes that are malicious into the network [12]. The incidence of false positives with increased speed of node movement is shown in Fig. 3. From the diagrammatic representation, it is clear that the rate of false positives decreases to a far bigger degree in our Cluster-based system compared to the other scheme. In reality, our proposed methodology better analyses the overall possible cause of an event of packet drop, and then a decision is made on the trustworthiness of the node. Overall, the statistic indicates that with an increase in growing node speed the false positive rate increases. On similar lines, Representation indicates a false positive rate along with the growing density of the node in the network, keeping the moving speed is 4 m / sec at the node is rigid. With an increased count of a node in the architecture, the source/destination number pairs are also increasing, because due to collisions, more packets are lost in the network. The number of false positives in the Cluster-based scheme is smaller relative to numerous other schemes, which considers every packet drop as an activity that is malicious [28], because the frequency of each packet drop is measured before making any judgment on the behavior of nodes and its multipath strategies.

In Fig. 3 shows the false Positives versus node Moving Speed with moving velocity and thickness on bogus positives where the x-pivot addresses the no of hubs use in-network and the y-hub addresses the False Positives rate.

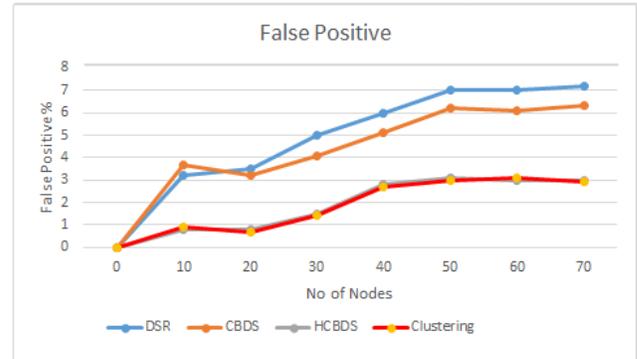
B. Detection Rate

The rate of detection with rising moving nodes under the Cluster-based scheme and the other device is shown in Fig. 2. In our Cluster-based system, the detection rate is higher, as every decision is unbiased. The detection rate is the number of true malicious nodes found in the scheme relative to the total number of malicious nodes. The other plan considers every parcel drop as malevolent and the related hub is viewed as malignant and consequently more real hubs are vindictive. As the insights show, the recognition rate for our Cluster-based plan is higher than for the other plan. Moreover, the identification rate with developing hub thickness is displayed in Fig. 2. The quantity of information associations inside the organization is additionally developing with the expanding hub thickness, which implies that more parcels are dropped on the organization because of crashes. The other scheme considers packet drops to be misbehavior of valid nodes. Therefore the detection rate in our Cluster-based scheme is higher again than in other schemes, as shown in the figure.

In Fig. 4 shows the Detection Rate vs. Node Speed with moving velocity and thickness on identification rate where the x-pivot addresses the no of hubs use in-network and the y-hub addresses the Detection rate.

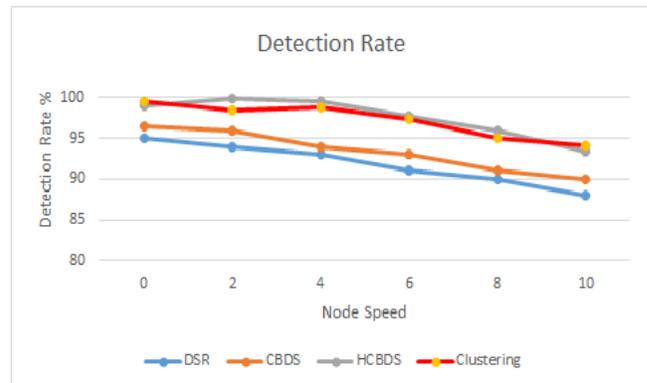


(a) False Positives vs. Node Moving Speed.

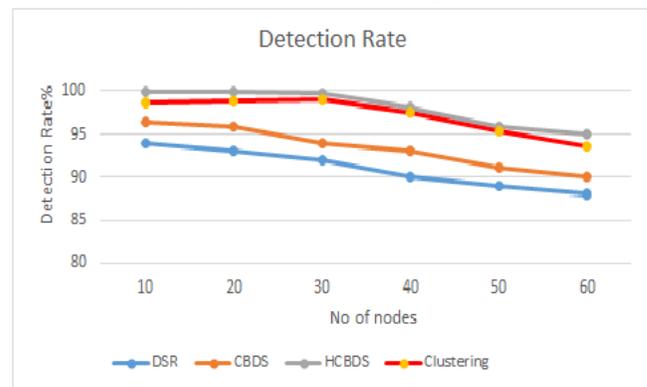


(b) False Positives vs. No Node Density.

Fig. 3. Effect of Node Moving Speed and Density on False Positives.



(a) Detection Rate vs. Node Speed.



(b) Detection Rate vs. Node Density.

Fig. 4. Effect of Node Moving Speed and Density on Detection Rate.

C. Energy Consumption

Fig. 5 shows the energy consumed under the Cluster-based system and CBDS schemes at rising node speeds. The goal of this experiment is to show that the total costs of processing and communication in the Cluster-based scheme are higher compared with the method i.e. the CBDS and other schemes. Packet transmission and receipt absorb most node resources. There is a novel energy-efficient secure routing protocol for the ad-hoc network with a Mobile sink [26]. Our Cluster-based plan doesn't build the quantity of traded messages; rather it utilizes existing directing bundles to trade data like line status and association status (already required according to routing Protocol standards). In addition, the data path continues to convey true malicious nodes.

D. Packet Loss Rate

In the cluster-based system and CBDS schema, the packet loss rate is shown in Fig. 6. The packet loss rate in our cluster system is less than in the CBDS scheme, as is seen in the figure. In reality, in the cluster-based routing path system more trustworthy nodes are chosen which leads to reduced packet losses and greater packet delivery ratios. For the CBDS & Other systems, genuine node isolation is possible, which leads to a greater drop in packets since transmission nodes to the destination are not available. In addition, the data path remains true malicious nodes, which provide them more possibilities of dropping vital data packets.

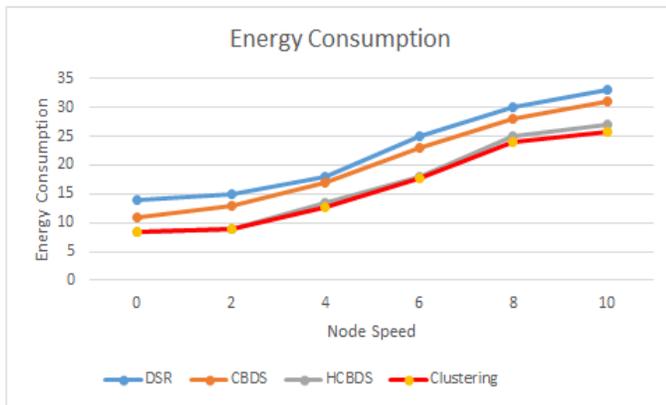


Fig. 5. Energy Consumption.

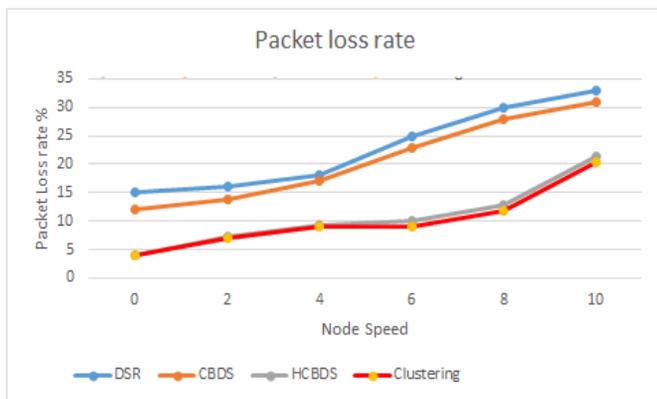


Fig. 6. Packet Loss Rate.

E. Packet Delivery Ratio (PDR)

PDR of Conventional DSR routing protocol. PDR is the presence of Sybil attack, partially secure CBDS routing protocol and proposed Cluster-based System, is shown in Fig. 7. The percentage data loss in DSR & CBDS under Sybil Attack is increased more than the Cluster-based routing protocol in all scenarios.

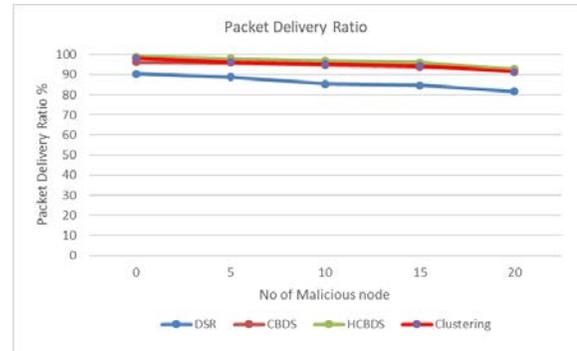


Fig. 7. Performance Analysis of Packet Delivery Ratio using DSR, CBDS, HCBDS & Clustering.

F. End to End Delay

The end-to-end delay performance of the conventional DSR, CBDS, HCBDS & clustering is the results are shown in Fig. 8. The CBDS, HCBDS producing over average end-to-end delay compared with clustering produces. From the results, it concludes that the model is flooding a minimum number of delays as compared to other.

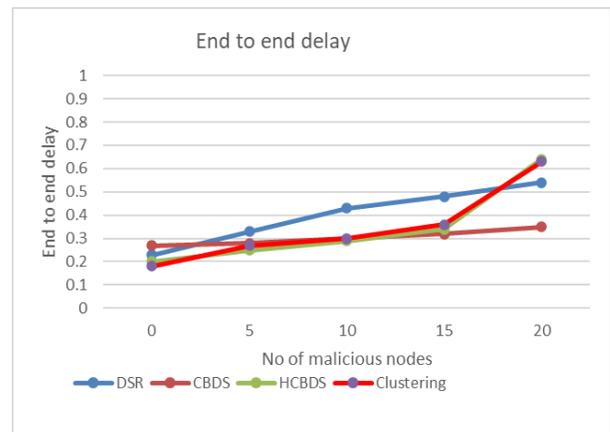


Fig. 8. Performance Analysis of End to End Delay using DSR, CBDS, HCBDS & Clustering.

V. CONCLUSION AND FUTURE WORK

Among other things, Sybil assaults are known as the most dangerous Adhoc network attacks. While several mechanisms exist for protecting Adhoc networks from such attacks, there are significant limitations and several drawbacks in the conventional approach. In addition, during the path discovery process, DSR does not detect malicious nodes and hence cannot send all data packets to the target during Sybil attacks. Many conventional approaches are inefficient to detect malicious activity. In addition, the delivery ratio of packets (PDR) under these attacks may decrease with an increase of

malicious nodes. Therefore, a new Clustering mechanism was proposed to protect ad hoc networks.

An advanced version of DSR with Clustering is first used to find and protect the malicious nodes that lead to attacks when a network is built using an NS-2 simulator. Additional encryption was done twice to boost security. It guarantees confidentiality to ahead and back. Whenever topology changes, all authenticated neighbors obtain the new neighborhood key and will be supported. In conclusion, all of these mechanisms prevent attacks of Sybil, and the proven growth in performance and an improved Packet Delivery Ratio are especially worthy of attention.

By looking at the result, the proposed system showed improved performance in terms of packet delivery ratio and output compared to the CBDS, HCBDS processes. This proposed work increases the number of normal nodes in the network and hence improves the performance. In the future, we can focus on the detection of different attacks with their performance analysis.

In future work we intend to investigate:

- 1) Varying density evaluation in the clustered network.
- 2) Other parameters of packet losses like MAC layer information.
- 3) Performance of our method under different security threats in the mobile ad-hoc network.

REFERENCES

- [1] H L Bhavyashree ,Nagarathna C R , Anusha Preetham,Priyanka R “Modified cluster based certificate blocking of misbehaving nodes in MANETs” International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering March 2019.
- [2] Chetan S Arage, K. V. V. Satyanarayana, “An experimental analysis on various techniques for malicious node detection in MANET”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-9S3, July 2019.
- [3] Arage Chetan S, Satyanarayana K V V, “Novel Routing Protocol for Secure Data Transmission in Wireless Ad Hoc Networks”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-4S2 March, 2019.
- [4] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, “VANet security challenges and solutions: A survey,” Veh. Commun., vol. 7, pp. 7–20, Jan. 2017.
- [5] I. Yaqoob, I. Ahmad, E. Ahmed, A. Gani, M. Imran, and N. Guizani, “Overcoming the key challenges to establishing vehicular communication: Is SDN the answer,” IEEE Commun. Mag., vol. 55, no. 7, pp. 128–134, Jul. 2017.
- [6] I. Ahmad, R. M. Noor, I. Ali, M. Imran, and A. Vasilakos, “Characterizing the role of vehicular cloud computing in road traffic management,” Int. J. Distrib. Sensor Netw., vol. 13, no. 5, 2017, Art. no. 1550147717708728.
- [7] Muhammad salem khan, danielle midi, majidiqbal khan, elisabertino, “fine-grained analysis of packet loss in manets”, 2169-3536 2017 ieeec.
- [8] I. Ahmad, U. Ashraf, and A. Ghafoor, “A comparative QoS survey of mobile ad hoc network routing protocols,” J. Chin. Inst. Eng., vol. 39, no. 5, pp. 585–592, 2016.
- [9] L. Li and G. Lee, “DDoS attack detection and wavelets,” Telecommun. Syst., vol. 28, nos. 3–4, pp. 435–451, 2005.
- [10] C. Buragohain, M. J. Kalita, S. Singh, and D. K. Bhattacharyya, “Anomaly based DDoS attack detection,” Int. J. Comput. Appl., vol. 123, no. 17, 2015.
- [11] S. S. Manvi and S. Tangade, “A survey on authentication schemes in VANETs for secured communication,” Veh. Commun., vol. 9, pp. 19–30, Jul. 2017.
- [12] Srinivasa Rao, Y., & Hussain, M. A. (2018). Dynamic MAC protocol to enhancing the quality of real time traffic in MANET using network load adaptation. Journal of Advanced Research in Dynamical and Control Systems, 10(7 Special Issue), 1612-1617.
- [13] Suma, P., & Hussain, M. A. (2018). Secure and effective random paths selection (SERPS) algorithm for security in MANETs. International Journal of Engineering and Technology(UAE), 7(2), 134-138. doi:10.14419/ijet.v7i2.8.10345.
- [14] A. Sinha and S. K. Mishra, “Preventing VANET from DOS & DDOS attack,” Int. J. Eng. Trends Technol., vol. 4, no. 10, pp. 4373–4376, 2013.
- [15] Boddu, N., Vatambeti, R., & Bobba, V. (2017). Achieving energy efficiency and increasing the network life time in MANET through fault tolerant multi-path routing. International Journal of Intelligent Engineering and Systems, 10(3), 166-172. doi:10.22266/ijies2017.0630.18.
- [16] Kolagani, P., Aditya, K., Venkatesh, N., & Kiran, K. V. D. (2017). Multi cross protocol with hybrid topography control for manets. Journal of Theoretical and Applied Information Technology, 95(3), 457-467.
- [17] K. Verma and H. Hasbullah, “Bloom-filter based IP-CHOCK detection scheme for denial of service attacks in VANET,” Secur. Commun. Netw., vol. 8, pp. 864–878, Mar. 2015.
- [18] Jun- zaho sun, “Mobile Ad hoc Networking: An essential technology for pervasive computing”, Machine Vision and Media Processing, 2000.
- [19] Tarunpreet Bhatia , A.K. Verma , “ Security Issues in Manet: A Survey on Attacks and Defense Mechanisms ,“ International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, pages 1-4, June 2013.
- [20] Kang, N., Shakshuki, E and Sheltami, T. , “Detecting Misbehaving Nodes in MANETs”, In Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services (iiWAS2010), ACM, pages 216-222, 2010.
- [21] Mainak Chatterjee, Sajal K. Das and Damla Turgut, “WCA: A Weighted Clustering Algorithm for Mobile Ad Hoc Networks,” Center for Research in Wireless Mobility and Networking (CRWMaN) Cluster Computing 5, pages 193–204, 2002.
- [22] P.Sakarindr and N. Ansari, “Security Services in Group Communications Over Wireless Infrastructure, Mobile Ad Hoc, and Wireless Sensor Networks,” IEEE Wireless Comm., vol. 14, no. 5, pages 8-20, Oct. 2007.
- [23] Nirwan Ansari, Wei Liu and Hiroki Nishiyama, “Cluster-Based Certificate Revocation with Vindication Capability for Mobile Ad Hoc Networks,” IEEE transactions on parallel and distributed systems, vol. 24, no. 2, pages 1-4, February 2013.
- [24] Seung Yi, Robin Kravets, “MOCA: Mobile Certificate Authority for Wireless Ad Hoc Networks,” University of Illinois at Urbana Champaign Urbana, pages 1-6, 2000.
- [25] Dr. Shaveta Ran1, Dr. Paramjeet Singh, Raman Preet, “Reviewing MANETs & Configurations of Certification Authority (CA) for node Authentication,” International Journal of Computer Science and Information Technologies, Vol. 4 (6), pages 974-978, 2013.
- [26] Kavitha. V, Ananthakumaran. S, “Novel energy-efficient secure routing protocol for wireless sensor networks with Mobile sink,” International Journal of Peer-to-Peer Networking and Applications, October, 2018.
- [27] Ananthakumaran. S, Sathishkumar. M, Bhavani. R, Ravinder Reddy. R, “Prevention of Routing Attacks using Trust-Based Multipath Protocol,” International Journal of Advanced Trends in Computer Science and Engineering, Vol. 9, No. 3, pp. 4022-4029, May-June, 2020.
- [28] Shivaprasad More, Udaykumar Naik “A Novel Technique in Multihop Environment for Efficient Emergency Message Dissemination and Lossless Video Transmission in VANETS” JCIN IEEE Journal- Volume 3, Issue 3 September 2018.
- [29] Shivaprasad More, Udaykumar Naik “Optimal Multipath routing for video transmission in VANETs” Wireless Personal Communication (WPC) , https://doi.org/10.1007/s11277-020-07740-1.

Augmented Reality based Adaptive and Collaborative Learning Methods for Improved Primary Education Towards Fourth Industrial Revolution (IR 4.0)

A F M Saifuddin Saif¹

Department of Computer Science
American International University -Bangladesh

Zainal Rasyid Mahayuddin², Azrulhizam Shapi'i³

Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM, Bangi, Selangor, Malaysia

Abstract—Existing primary education is not organized properly to fulfil the demand of fourth industrial revolution (IR 4.0) which causes lack of engagement in learning materials, lack of spatial ability and motivation towards learning by young age students. Students especially from initial education level tend to learn from seeing rather than reading or memorizing. In this context, technology like augmented reality and virtual reality bears such visual power in lieu with interactivity and engaging characteristics by bringing virtual and real world together. In addition, flexible presentation mechanisms in lieu with tagging and tracking technology with various degree of freedom provide the effective ground of augmented reality and virtual reality to integrate with current primary education or initial level of education. In this context, this research presents comprehensive and critical reviews in the previous research in terms with several aspects such as methods, research design and experimental validation. Each of the aspects is demonstrated with adequate advantages and disadvantages to design the integration of augmented reality and virtual reality with primary education methodology. In addition, this research illustrates extensive critical explanation for various challenges to integrate augmented reality and virtual reality with primary education to make students motivated towards education, effective activity and memorization with visualization. The overall investigation proposed in this research is expected to fulfil the future demand for Fourth Industrial Revolution (IR 4.0).

Keywords—Augmented reality; virtual reality; adaptive and collaborative learning; fourth industrial revolution

I. INTRODUCTION

Education is a continuous learning process which should include not only traditional class activities, learning process should be creative and interactive to meet the demand of fourth industrial revolution (IR 4.0) so that students will not feel demotivated towards learning. Besides, way of acquisition of education should be through discussion, research, training which should be fun and interesting. However, current initial level of education system seems to abandon the main idea about education and learning process to lean towards memorizing and rote learning. Traditional primary education is based on two-dimensional materials and does not provide any of these opportunities. Children aged from 5-11 are the ones who goes to primary schools and considered to be future hope, are losing their hope from learning. As a result, dropout rates are growing day-by-day and students are lacking behind

in lieu with not being able to show their creativity. To keep the children focused, motivated, creative and to keep their enthusiasm alive, changes needs to be brought in the way of delivering education to the children which demands to opt for technology integration with primary education. To integrate technology with primary education, researchers need to take account more on visualization and boosting engagement of students. Technology with such visualization power as well as interactivity is Augmented Reality, which bring virtual and real world together by putting digital information on the world. Existing research has been engaged in finding ways to integrate augmented reality with education system and achieved positive results. However, most of them were integrated with college or higher studies. Previous research on augmented reality with education medium and positive results so far gained by the previous research encourage this research to depict critical and comprehensive reviews to integrate augmented Reality with primary education system.

Augmented Reality and virtual reality is an advanced technology which connects virtual and physical world around us by superimposing digital information like sound, images or objects which can be seen with the help of regular handheld devices like mobile phones, tablets, laptops. All the positive characteristics of augmented reality make it dynamic and perfect combination for technology-based education system. Students from primary education tend to learn things by visualizing not reading or memorizing them. Augmented Reality is perfect medium in this context which provides a great deal of visualization with digital information as well as let students interact with the system so that contents can be taught in groups. AR in classroom is expected to be helpful to capture student's attention towards the learning and make them feel more engaging with other students and teacher as well [1] due to some factors, i.e. enhance environment and help with discovery-based leaning and increased student's spatial ability and motivate them towards learning. This research aims to illustrate various methods in terms with integrating AR technology with primary education level.

For the opportunities provided by augmented reality and virtual reality, AR has been used in college and higher education level in the previous research. The results from their researches illustrates that augmented reality and virtual reality has shown a promising result, better understanding of the

subject, boost engagement among students with course material, improve of interest and attention in classroom which can play motivating factors to mix augmented reality in primary education. This research presents critical and comprehensive reviews of using augmented reality in the exiting research towards using augmented reality in primary education. Section 2 presents core research background, Section 3 demonstrates comprehensive and critical reviews based on existing methods, Section 4 illustrates existing research based on frameworks, Section 5 depicts experimental validation by the previous research, Section 6 describes observation and discussion and finally, concluding remarks are presented in Section 7.

II. CORE BACKGROUND STUDY

Existing education system is not structured properly to meet the demand of fourth industrial revolution (IR 4.0) and its hampering society in a lot of ways, i.e. students are not feeling motivated, not curious about learning, tired of memorizing and reading books when they cannot grasp and there is no alternative to it. Education sharpens the way to future and it's important to work for the betterment of education. But it's not easy to improve education system in a short time where all issues need to address properly for proper structured education methodology and will provide solutions to most of the problems as well as will not become outdated within a short period of time. Making education better with technology, researchers are continuously conducting researches and they have come up with the technology called Augmented Reality [2]. For the last few decades research has been going on the possibilities of Augmented Reality in education system and most of the researches have been conducted on higher level of education to teach different subjects and measuring its success based on some key factors, i.e. students motivation towards learning, engagement with materials, concept understanding, visualization for enhancing spatial abilities, grasping deeper understanding of abstract and difficult topics [3]. The result of the research was positive and successful in terms of achieving the satisfactory level of students and teachers. Augmented Reality has provided students with new techniques of visualizing course materials, new way of interacting with course materials within the class, changing their perspective of learning and increasing their interest to the courses.

Augmented reality and virtual reality have all the characteristics needed to improve our current education system and give education the pace it needs to be mixed with the technology. Existing research showed how to implement augmented reality with education system to improve education's current situations. In Turkey it is very difficult to achieve good results in Ottoman Turkish reading but with AR technology they have found that students have less difficulty and performed well in terms of performance [30]. Medical Subjects need much deeper understanding and they are facing problems especially in generating real life scenarios but using Miracle [5], an AR system it was possible to make real life scenarios and provide students with an interactive virtual human body where they can explore for their better understanding. Similarly, there are a lot of augmented reality applications which are being used to taught different subjects

to students in different countries. Augmented Reality are used in medical education [6,7], science [8,9], engineering [10,11], object detection [12,13,14], motion modelling [15,16,17], face modelling [18,19], analysis of aerial images [20,21,22] etc. and a few mentionable Augmented reality applications are Elements 4D, Experience Chemistry teach chemistry, Anatomy 4D, Human Heart 3D and Corinth Micro Anatomy teach human anatomy, CARE an Augmented Reality application built to teach Science, Technology, Engineering and Math parts. ARVe - Augmented Reality applied to vegetal field [23], FreshAir™ teach discovery-based learning [24], Construct 3D to teach math (geometry) to college students [25], AR Physics to teach physics to college students [26]. All these AR applications are really making some differences to engaging students and their motivation. However, all of these existing AR applications are for college or higher-level study materials.

Augmented reality has also been used in primary level education but the usage is really few, yet the results are positive. Augmented Reality Popup Book application which was used to teach English Language and the course material were built according to English curriculum [27]. Realitat3 is an augmented reality system consists of six applications: skeletal apparatus, water cycle, plant development, frog metamorphosis, solar system and the senses. All these AR applications are being used to teach primary school students to improve their efficiency (academic achievement), usability and motivation [28]. Again, researcher in [27] explored the possibility of augmented reality application in teaching English to primary school children. Research in [29] was conducted to discover the geometrical disadvantages of technology for implementing augmented reality in education, the result illustrates positive feedback and promising output. The application was one mouse per children and it teaches math to students. Research in [1], opt for different methodologies to figure out which will work best for integrating augmented reality with primary education system. Other business applications which were built for primary education, i.e. Math alive teaches math's up to 3rd grade students, Animal Alphabet AR Flashcards are built to learn letters, Zookazam or Bug 3D to teach about animal species. However, there are few applications which can be incorporated with the curriculum provided for primary education level children. Existing experimental results for various AR applications are remarkable to prove enough that AR can not only be incorporated with primary education system but also it is well acceptable by teachers, students, parents as well. In addition, existing usage of augmented reality has showed better engagement with learning contents, more attentive and faster learning.

III. REVIEW BASED ON METHODS

Although currently we are living in three-dimensional era, however two dimensional materials is still provided to use for education which is failing us to provide quality education to keep our pace with modern era. To keep up pace with modern era we need to embedded technology with education and many countries are doing so, and the result indicate positive and better impact on learning. Researcher has been doing so many experiments and testing many approaches to integrate

technology with education for quality education so that they could find the optimal and best solution possible to integrate Augmented Reality with education system. However, finding organized procedure which can help us to integrate technology with our education system as well as work smoothly with our existing educational system is not an easy task but with hard work and depth research, researchers have also come up with some effective procedures which have shown how fun and amazing education can be with help of augmented reality. Existing methods in this context are shown in Fig. 1.

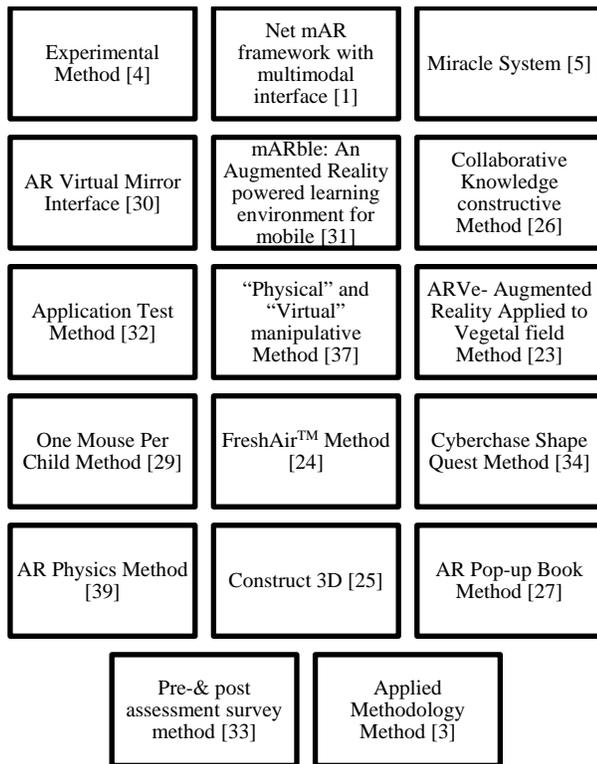


Fig. 1. Existing Learning Methods using Augmented Reality.

Augmented reality is highly praised due to its visualization capacity and interactive capacity with the system [1]. Researcher in [1] proposed multimodal interface model as mobile augmented reality framework using these dynamic characteristics of augmented reality. Their proposed system is built for mobile version in hope of increasing interactivity which will boost the engagement of students with the learning materials and increased the accessibility of augmented reality based educational material through application. However, their proposed method is still no acting as hypothesis as their proposed method is not implemented. Weakness of multimodal interface model can be improved by Construct 3D, a collaborative learning method.

Construct 3D method was introduced to overcome its drawback and flows of multimodal interface model which was theoretical hypothesis. Construct 3D method is based on the interactivity characteristics and have modified the characteristics by creating various interactive model to teach students about geometry and increase their spatial ability [25]. Construct 3D method also increased student’s interaction with system and for various model of interactivity in lieu with

providing more engagement with course materials. However, construct 3D method was designed to teach geometry only to students. To overcome the disadvantage of construct 3D method, AR mirror interface method was introduced. AR mirror interface model was proposed to study the affordance of augmented reality in education, proposing a less expensive and more user interactive tool [30]. Although, AR mirror interface model was successful to fulfil its purpose, AR mirror interface model required more engagement of students to teach some subjects than tradition educational methodologies. To fill the gap created by AR Mirror interface another method was introduced Realitat3.

Realitat3 is a method proposed for increasing interaction between system and user while keeping the affordance in AR system which consists of six AR games: skeletal apparatus, water cycle, plant development, frog metamorphosis, solar system and the senses. Realitat3 introduced AR gaming for educational purpose and more engagement while keeping student motivated, interested towards learning [28]. However, Realitat3 method is not able to teach all subjects by games which can be overcome by collaborative knowledge construction method using AR.

Collaborative knowledge construction method integrated augmented reality to help learners to develop skills and knowledge effectively, enabled co-located learners to be involved in knowledge construction by exploring the outside environment and interacting with each other through the guidance of AR-supported information [26]. However, software usability or learners perceived as learning effectiveness of the AR system needs to improved more for collaborative knowledge construction method using AR. To overcome this issue, research in [5] came up with Miracle System.

Research in [5] explored the possible areas of integrating augmented reality with medical study to give medical students a better way to learn medical studies. They focused on the visualization power of augmented reality as medical subjects especially Anatomy which needs figure to get a better understanding. Anatomy is a topic where students are taught the internal structures of human body so that they can have a deeper knowledge about human body and how organs work. To match with anatomy characteristics research in [5] introduced Miracle system which displays anatomical structures of the user’s body to a monitor and they have also proposed a gesture-based interaction to get real life experience. However, their proposed research was only for anatomy study and focused on a small number of important organs of the abdomen. For the usage of large screen and in order to solve the portability issue research in [31] proposed a portable augmented system called mARble.

Research in [31] proposed mARble which was an augmented reality powered learning environment for mobile proposed to make augmented system portable and affordable to use. As most of the students carry smartphone with them and now smartphone have enough computing power to carry out the operations of augmented reality. They tried to make students smartphone alternative of their books. They aimed for portability and making it as an alternative of books to teach

basic medicine courses. Using their system medical students performed better than those who were taught using books. Their performance was measured in terms of “the hedonic qualities simulation (HQ-S), identification (HQ-I) and attractiveness (ATT)”. However, their proposed system was used only for teaching basic courses in legal medicine which is the drawback of their method and solved by research in [27].

Research in [27] introduced AR pop-up Book to solve the content issue which was faced by most of the previous research methods. They integrated augmented reality with the book we read where users need to scan the image with the smart phone and augmented contents will appear on their mobile screens. They used AR pop-up Book to teach physics to the students which also needs deeper understanding. AR pop-up books help the students to bridge the gap between the digital and physical world. However, drawback of their method was that their proposed method was tested with 5 students only, so the results gained by AR pop-up Book may not be same for large number of users. In this context, research in [32] came with the solution by proposing Applied Test Method.

Applied Test Method experimented how augmented reality can be used in education with large number of participants [32]. Research in [32] proposed Applied Test Method to display real time results on screen, designed for educational purposes, interactive features like reactive to plain paper notebooks. However, their proposed methods provides some concerns towards privacy and ethical concerns, social acceptance concern of being constantly photographed, videotaped. Even though Applied Test Method was designed for education, however, it seems that Applied Test Method lacks effectiveness when used with high-achieving students which indicates little impact on high-achieving students. To solve this weakness, research in [33] introduced Pre-& Post Assessment Survey Method.

Research in [33] worked with how augmented reality can provide quality education in terms with delivering lessons to the students. They proposed Pre-& Post Assessment Survey

Method where they gathered students, teachers’ feedback with traditional and augmented reality based educational system. They took an assessment before and after use of augmented reality and analyzed results. According to their survey, they were able to achieve better understanding level for spatial concepts and make learning more effective and interactive. However, they implemented their research on students who were in grad 6 to 8 only. To investigate more about the spatial ability improvement with augmented reality, research in [34] proposed Cyberchase Shape Quest method.

Research in [34] introduced Cyberchase Shape Quest Method to improve spatial ability with augmented reality. Their goal was to utilize augmented reality to teach geometry as this particular subject needs more spatial ability to grasp the problem. With their proposed method they saw students became engaged with spatial memory, visualization for seeing the content and planning to solve the given task which showed that their method was more interactive, engaging and efficient. However, their method can only teach students about geometry which is a major drawback. To solve this problem, research in [4] introduced Experimental Method. Experimental method was introduced to overcome the issue and crisis defined by other methods. In this method, research in [4] tends to select a study group to gathering their feedback after using the system and monitor the results. So far experimental method gives the best output with positive result and great motivation to leaning [4]. Previously, many methods were tested to integrate AR with educational methodology to have a positive and good impact on engaging students, boosting their interactivity. Existing concerns by the researchers proves that augmented reality can be integrated with primary education system for improved and adaptive learning technique. However, existing almost all the methods needs further investigation due to the lack of experimentation on larger number of users, portability issues, user friendliness, higher engagement of the students. Previous research methods with advantages and disadvantages for integrating augmented reality to improved existing education methodology are mentioned in Table I.

TABLE I. PREVIOUS PROPOSED METHODS WITH ADVANTAGE AND DISADVANTAGE

Method Name	Advantage	Disadvantage
Experimental method [4]	So far best output with positive result and great motivation to leaning.	The experiment was conducted on 60 students and limited words. The result might be different for large group of students and syllabus
Net mAR framework with multimodal interface [1]	Using Multimodal Interface, usability of application increases as people can interact with the app.	Only theoretical method proposed.
Miracle system [5]	(a) Displayed anatomical structures on user's body, (b) New concept for gesture-based interaction.	Their method was used for only anatomy of small number of important organs of the abdomen.
AR Virtual Mirror Interface [30]	(a) Studied affordance and contains of AR technology in formal education, (b) Showed AR virtual mirror is less expensive and more users can interact with it, (c) Showed how to improve AR to integrate with national curriculum of UK Primary education	(a) Students using AR were less engaged than those using traditional resources, (b) Only used to teach science
mARble: An Augmented Reality powered learning environment for mobile[31]	Performed better than text book in terms of "the hedonic qualities simulation (HQ-S), identification (HQ-I) and attractiveness (ATT)	Applicable only for basic course in legal medicine.
AR in education [35]	(a) AR system detected students' locations and working status, (b) Provided task reminders, (c) Offered alternatives to refocus students' attention, (d) Enhance students' recognition of community of learners	(a) Students often felt overwhelmed and confused when they were engaged in a multi-user AR simulation. (b) Students require applying and synthesizing multiple complex skills in spatial navigation, collaboration, problem solving, technology manipulation, and mathematical estimation.
Collaborative knowledge construction [26]	(a) Helped learners to develop skills and knowledge effectively. (b) Enabled co-located learners to be involved in knowledge construction by exploring the outside environment and interacting with each other through the guidance of AR-supported information.	Software usability needs to be improved for effectiveness of the AR system.
Application Test Method [32]	(a) Displayed real-time result. (b) Designed for educational purposes, interactive features reactive to plain paper notebooks.	Causes privacy and ethical concerns in lieu with social acceptance of being constantly photographed, videotaped, and analyzed, less effectiveness when used with high-achieving students.
"Physical" and "Virtual" manipulative [37]	Showed that students motivation can be increased through AR technology. (b) Children who cannot yet think symbolically, AR play vital role in that context.	AR experiences may allow students to view virtual content and other people in the same mixed-reality space.
ARVe - Augmented Reality applied to Vegetal field [23]	(a) Used for students with cognitive disabled, (b) Provided feedback from the students, teachers and parents.	Worked only in the basis of visualization and matching.
Case Study of AR in Primary School [28]	An evaluation of the educational contents was made from the point of view of efficiency (academic achievement), usability and motivation	Depends on work which was conducted between 1993 and 2007
One Mouse Per Child [29]	(a) The system can be implemented even with different environment condition with minimal equipment, (b) established statistically relevant results and observed that the software proved most beneficial for the students with the lowest initial results.	Only teaches math to primary school children.
FreshAir TM [24]	Navigated the pond environment and observed virtual media information overlaid on the virtual pond	The system was built for outdoor study, there was no indication how it will integrate with indoor technology.
Cyberchase Shape Quest [34]	While using the application students engaged their spatial memory, visualization & planning skills.	Focused only on geometry.
AR Physics [39]	(a) Initial step toward understanding learners' knowledge construction process by using AR System, (b) Learners knowledge construction behaviors were qualitatively identified according to adaptive three-category coding scheme.	Focused only in the performance on learners' knowledge construction of elastic collision of physics.
Differentiate between learning theories [36]	Showed how they selected different teaching method for augmented reality application.	Provided theoretical view with no experimental justification.
Construct 3D [25]	(a) Run on heterogeneous and hybrid setups. (b) System was built for mobile platform and collaborative.	(a) Their laboratory set up cannot be fully implemented in the schools or colleges, (b) They only aimed for geometry.
AR Pop Up Book [27]	Helped students to bridge the gap between digital and physical world.	Experiment was conducted on only 5 students
Pre- & post assessment survey method [33]	Better understand spatial concepts and make learning more effective and interactive	This research is done only in grad 6-8
Applied methodology [3]	Showed different types of AR application in educational environment	No option to reproduce study.

IV. REVIEW BASED ON FRAMEWORKS

Embedding AR in education to provide quality education is considered as challenging tasks and researcher is still trying to find an optimal solution for which they tested and proposed many methods [39,40]. To prove those methods, they used some framework by which they intend to support their work which has proved to have better result in outcome. To support Multimodal Interface method researchers in [1] proposed Net mAR Framework which is divided into five phases, i.e. Phase 1: Conceptual Features to be applied on the application, Phase 2: Focus more on the software and hardware requirement to develop application, Phase 3: Prototype Development, Phase 4: Application Reliability Testing, Phase 5: Final Application. Advantages of the framework are motivating students in learning as well as helping them to retain the information longer, benefit the educator explaining more complex subjects to students and improve student’s memorization capability through visualization [1]. However, drawback of the framework have is being a theory. To overcome this drawback, another framework was introduced for collaborative augmented reality system.

Collaborative augmented reality framework was based on collaborative construct 3D method which builds for collaboration of a large number of students and supports various teacher-student interaction scenarios. Collaborative augmented reality framework had hybrid hardware setups and provided a natural setting for face to face collaboration of teachers and students, projection walls were established techniques for semi-immersive group environment, and single projector display are affordable for classroom use [25]. Having all the advantages, collaborative augmented reality framework also comes with disadvantage. Due to the number of available AR sets significantly restricts the use in large groups and the screen is shared between the active user and the observer. To overcome these drawbacks AR Virtual Mirror Interface with socio-cultural theory framework can be used as remedy.

AR Virtual Mirror Interface with socio-cultural framework was built for affordance of AR virtual mirror and its advantage over head mounted system which describes the potentiality of AR in primary education [30]. In that research, framework needs four steps to implement, i.e. AR Sessions and recordings of the sessions, class teacher interview, data analysis, comparing teacher dialogue across AR and Traditional teaching sessions. Again, AR Virtual Mirror Interface with socio-cultural framework had issues like need to train teachers so that they do not fell out of script while teaching with new technology and limited observation time. To surpass this challenge, another framework is being proposed named as collaborative knowledge constructive framework.

TABLE II. PREVIOUS FRAMEWORKS WITH ADVANTAGES AND DISADVANTAGES

Previous Frameworks	Advantages	Disadvantages
Quasi-Experimental Design framework [4]	(a)Academic Achievement shows pre-& post-test scores differ significantly. (b) Material motivation questionnaires show that motivation of the students was high in positive direction. (c) Fears of student towards the Ottoman Turkish Lesson were decreased.	(a) Obtained data are limited to the study group and miracles group, (b) AR integration was not fully developed.
Net mAR Framework [1]	(a) Motivated students in learning as well as helping them retain the information longer. (b) Benefited educator in explaining more complex objects to students, (c) Improve memorization through visualization	Still in theoretical phase, not in deployment.
AR Virtual Mirror Interface with socio-cultural theory framework [30]	(a) Affordance of AR virtual mirror and its advantage over head mounted system, (b) Describes the potentiality of AR in primary education.	(a) Need to train teachers so that they do not fell out of script while teaching with new technology, (b) Students less engaging with AR as some of them holding AR tiles and others are observing, (c) Limited time
Framework for mARble [31]	(a) Students using mARble showed better performance than textbook, (b) Improved learning rates.	Only Basic medical content were provided.
Collaborative knowledge constructive framework [26]	(a) Provided great possibilities for supporting face-to-face collaborative learning, (b) Enhanced social interactivity between or among group learners, (c) AR group learners performed significantly better in terms of their post-test scores.	(a) While constructing relations between single theoretical concepts or distinguished concepts from each other, it would confront obstacles regarding the topic of elastic collision and need support from their partners, (b) Other dimensions such as the argument and social modes dimensions needed to be analyzed.
Discovery based learning framework [38]	Showed benefits of AR in educational environments.	Limited by a number of factors, i.e. First, identified empirical studies are only informal investigations with a low number of participants. The significance of the ascertained benefits of AR applications may be unclear in these cases.
Framework for ARve [23]	Safe and highly portable	Required further experimental validation.

Collaborative knowledge constructive framework consists of various steps, i.e. Mobile AR in education, collaborative knowledge construction, enhancing learners understanding of elastic collision and momentum with technology, AR Physics system comparison with traditional 2D Physics, data analysis [26]. Collaborative knowledge constructive framework gave great possibilities for supporting face-to-face collaborative learning, enhancing social interactivity between or among group learners. AR group learners performed significantly better in terms of their post-test scores with disadvantages like while constructing relations between single theoretical concepts or distinguished concepts from each other, it would confront obstacles regarding the topic of elastic collision and need support from their partners. To overcome this challenge Quasi-Experimental Design framework was introduced.

Quasi-Experimental Design framework was based upon experimental method which takes a few steps to complete, i.e. research design, study group selection, data collection from the groups and finding representation. Material motivation questionnaires showed that motivation of the students was high in positive direction and fears of student towards the Ottoman Turkish Lesson were decreased [4]. Like all other framework, Quasi-Experimental Design framework also has some drawbacks, i.e. obtained data are limited to study group and AR technology was not fully developed. Previous frameworks with advantages and disadvantages for integrating augmented reality to improved existing education methodology are mentioned in Table II.

V. REVIEW BASED ON EXPERIMENTS

Augmented Reality is going to make education fun, increase student's motivation and will help them to improve their academic results as well. Having said that framework is the proof of the working methods, previous research conducted experiments to see if those methods or frameworks to validated their research, even if they have given some theoretical methods and frameworks, they proposed experimental steps so that if anyone implemented that framework they would know how to measure success or outcome of the experiments [41,42]. Research in [4] formed study groups, trained them with course materials, then conducted pre- & post-test to see their methods effectiveness. They used academic and motivation questionnaires to measure the success of augmented reality. From their experiment with AR, pre-test result average was 29 and post-test result average was 56.4, which illustrates significant improvement with the help of augmented reality and states that augmented reality can help students to improve their results and increase motivation as well.

Augmented Reality can provide a great way for students to interact with the provided courses material and boost the engagement of students in classroom and keep them concentrated in the topic. Research in [30], tried to prove that by proposing a method called AR Virtual Mirror Interface. For

their experiment they used web camera, AR tile, AR virtual mirror to set up the environment and by this system they taught students to prove the effectiveness of their research. They used several steps to run the experiment, i.e. AR sessions where they took AR sessions led by AR experienced teacher, own class teacher and recorded data. In the data analysis session, they used NUD*IST Quantitative analysis software to find results, they also recorded teachers interview to see teacher's reaction towards the system and use of Augmented reality in education.

Research in [32] conducted experiment to find AR contribution in improving education and integrate the technology in formal education. They made AR application to investigate how it could help students in discovery-based learning. To build their system environment, they used presentation devices, stimulus devices, tagging of known objects and computational estimators. Tagging of physical objects is what makes a device recognized a specific object in the real world. In their research, tagging of physical objects was accomplished with a system of GPS sensors, magnetic sensors, accelerometers, digital cameras, wireless sensors, digital compasses, and ultrasonic sensors. In their research, accuracy of a tagging system depended on the sensitivity of the various components where data processing components of AR systems required large amounts of RAM. Image capture was performed with a binocular or monocular lens. By analyzing their result, they found that highly accurate GPS position tracking to within 3 centimetres and accurately displayed results to users.

Research in [23] conducted their experiment to state the effectiveness for implementing augmented to visualize books. They built their system to integrate with primary school and used video projector, web cam, AR Book, webcam to scan the QR code from the AR book, and sent the findings to the video projector for visualization. In their application, participants had to complete several tasks, i.e. matching fruits, flower, seeds and leaves. Then they recorded the time to complete the task and analyzed the results. Their system was also compatible for handicapped students. Using their system 90 % of the parents considered it positively for their teaching aspects. 97% of K-2 up to K-5 students finished 4 stages with fast and less training. Their results illustrate that augmented reality has great potentiality to be used in primary education.

Research in [28] proposed Realitat3 which consist of an AR engine and several AR applications, i.e. skeletal apparatus, water cycle, plant development, frog metamorphosis and solar system. They used camera calibration, marker detection and calculation of marker position and orientation (pose estimation). They tried to find out how AR games can be used to deliver education and increase student's spatial ability. 21 Students participated in their research experimentation and 95% confidence level was gained using their proposed research. Previous research experimentation for integrating augmented reality to improved existing education methodology are mentioned in Table III.

TABLE III. PREVIOUS RESEARCH EXPERIMENTATION

Previous research methods	Parameter Used	Numeric Results
Experimental Method [4]	Academic & Motivation questionnaires'	Academic result has been increased. Average Pre-Test Score: 29, Post- Test Score: 56.4.
AR Virtual Mirror Interface with socio-cultural theory [30]	Teacher question in AR sessions clarifying the relationship between elements on the screen, Teacher question in AR sessions: children turns to hold a tile, Teacher questions in the traditional teaching sessions, Teacher interviews	96% agreement between AR & traditional teaching transcript.
Four Step Research Approach [42]	Benefits of AR	Identified 14 different benefits of AR in education and among them 20% benefits were accounted for Improved learning curve and increased motivation
Application Test Method [32]	Accuracy of a tagging system	(a) Highly accurate GPS position tracking to within 3 centimetres. (b) Accurately displays results to users.
ARVe - Augmented Reality applied to Vegetal field [23]	Accuracy to consider positivity among users.	(a) 90 % of the parents considered it positively for their teaching aspects. (b) 97% of K-2 up to K-5 students finished the 4 stages with fast & less training
Realitat3 [28]	Number of participants and confidence level.	(a) 21 Students participated in the research, (b) 95% Confidence level using application
FreshAir™ [24]	Pre-field trip training and post-field trip discussion	70 students were participated in the study
Cyberchase Shape Quest [34]	Number of usages	The game has been downloaded more than 500,000 times

VI. OBSERVATION AND DISCUSSION

Human have evolved by adapting new methods or new ways of living in the context of fourth industrial revolution, still we are doing same in every aspect of our life except for our existing education system. To keep our pace with fourth industrial revolution (IR 4.0) we need to identify the problems we are having with our education and how to integrate them with the help of technology. Augmented reality and virtual reality have proved to be a technological tool which is expected to improve current educational barrier. This research presents how previous research integrated augmented reality

and virtual reality with education and observed some problems with their methods, framework and experiment.

From the review based on the method illustrates that augmented reality and virtual reality has really been showing its capability to provide quality education and fulfil the purpose with some drawbacks. In some of the methods, i.e. Experimental Method [4], Miracle System [5], mARble [31], One mouse per children [29], AR pop-up Book [27], Pre-& Post Assessment Survey [3] researchers experimented their method for only a small proportion of population which gives them expected result but it cannot be said that the result will be the same when the experimentation for validation will be for large population. Again, in some of the methods, i.e. case study of AR in primary School [28], differentiate between learning theories [8] are only theoretical hypothesis with no experiment done. In review based on methods, i.e. Experimental Method [4], Miracle System [5], mARble [31], One mouse per children [29], Cybershape Chase Quest [34], AR physics [39] content was created for particular topic or based on small portion of the topic. In Construct 3D [25] method, proposed system setup was not idle for implementation as their setup was costly or not affordable for all. Again, in the basis of review based on methods, i.e. FreshAir™[24] proposed system was environment oriented like indoor based augmented reality system which cannot be used in outdoor and outdoor based system cannot be implemented with indoor. In some reviews for the method illustrates that users found the system confusing causes less engagement for the participants or students.

Review based on framework showed same possibility with drawbacks. In Experimental Method Framework [4], AR was not fully developed to be integrated with education system. Net mAR Framework [1] was only in conceptual theory not implemented yet. In AR Virtual Mirror Interface with socio-cultural theory [5] framework, to integrate augmented reality with education researchers need to teach teachers about the system so that they do not fell out of script while teaching. In other review based on framework, i.e. mARble [31], AR Virtual Mirror Interface [30], Application test method [32] illustrate that Augmented reality teaching materials was not fully developed which can only help to teach some particular subjects efficiently and deeply. In addition, sometimes there are some ethical problems, lack effectiveness when using with high-achieving students.

Review based on the experiment showed clear scenarios of augmented reality in education system in terms with validation. Based on the experimental results of some methods showed promising results, i.e. by using Experimental Method [4], academic results have been improved by 27.4 % and using AR Virtual Mirror Interface [30] 96% agreement has been reached between AR and traditional teaching transcript. Based on the four step research approach experiment mentioned in research [42], 14 benefits of AR found in education among them 20% benefits were accounted for improved learning curve and increased motivation. With ARve experiment [23], 90% parents were satisfied with the usage of Augmented Reality. In addition, Realitat3 [28] received 95% confidence level among students.

Augmented reality and virtual reality has been used by different countries by different method. So far, in Chile and India Desktop application has been used to teach math to primary school student. In London also AR application was used to teach science subject, AR games are used to teach and improve spatial ability of primary school students. Spain used AR to teach geometry to primary school students. All of their efforts were fruitful and have illustrated improvements in engaging students with class material, boost activity, increased motivation, and memorization with visualization. To use augmented reality in primary education system AR books and AR application need to be used together as AR book will help them to teach history, literature subjects where AR Application with AR Games can teach them science, math and other subjects as well. In this way AR can be implemented in all the subjects to teach at primary level while making students motivated towards education, boost their engagement with class materials.

VII. CONCLUSION

This research focused on integrating augmented reality and virtual reality into primary education to overcome current challenges of traditional educational method. This research illustrated thorough advantages and disadvantages for various methods to integrate augmented reality with primary education in the context of fourth industrial revolution (IR.0). Comprehensive reviews by this research show that if it is possible to integrate augmented reality and virtual reality with higher level of education, it can be integrated with primary education as well. Besides, demonstrated critical reviews by this research not only analyzed key points for the integration of augmented reality with primary education but also depicted hardware requirement to make overall research design for integration portable, cost effective, affordable which will help students to engage with the system easily. Although, augmented reality and virtual reality has been used by some countries, this research recommends to use AR books and AR application together as AR book can be used to teach one category of subjects such as history or literatures types of subjects and AR application can be used to teach science or math related subjects. This hybrid methodology is expected to improve students' motivation and engagement in class materials. Demonstrated reviews by this research are expected to play significant role in developing adaptive and collaborative education environment to fulfil the demand of fourth industrial revolution (IR 4.0).

ACKNOWLEDGMENT

The authors would like to thank Universiti Kebangsaan Malaysia for providing financial support under the "Geran Universiti Penyelidikan" research grant, GUP-2020-064.

REFERENCES

- [1] A. F. Bulagang and A. B. Baharum, "A framework for developing mobile-augmented reality in higher learning education," *Indian J. Sci. Technol.*, vol. 10, no. 39, pp. 1-8, 2017.
- [2] J. L. Bacca Acosta, S. M. Baldiris Navarro, R. Fabregat Gesa, and S. Graf, "Augmented reality trends in education: a systematic review of research and applications," *Journal of Educational Technology and Society*, 2014, vol. 17, núm. 4, p. 133-149, 2014.
- [3] K. Lee, "Augmented reality in education and training," *TechTrends*, vol. 56, no. 2, pp. 13-21, 2012.
- [4] M. F. Özcan, Â. Özkan, and N. Sahin, "The Influence of the Augmented Reality Application on Students' Performances in Ottoman Turkish Readings," *Universal Journal of Educational Research*, vol. 5, no. n12B, pp. 27-33, 2017.
- [5] T. Blum, V. Kleeberger, C. Bichlmeier, and N. Navab, "mirracle: An augmented reality magic mirror system for anatomy education," in *2012 IEEE Virtual Reality Workshops (VRW)*, 2012: IEEE, pp. 115-116.
- [6] K. S. Tang, D. L. Cheng, E. Mi, and P. B. Greenberg, "Augmented reality in medical education: a systematic review," *Canadian medical education journal*, vol. 11, no. 1, p. e81, 2020.
- [7] D. Chytas *et al.*, "The role of augmented reality in anatomical education: An overview," *Annals of Anatomy-Anatomischer Anzeiger*, vol. 229, p. 151463, 2020.
- [8] R. Salar, F. Arici, S. Caliklar, and R. M. Yilmaz, "A Model for augmented reality immersion experiences of university students studying in science education," *Journal of Science Education and Technology*, vol. 29, no. 2, pp. 257-271, 2020.
- [9] E. E. Goff, A. Hartstone-Rose, M. J. Irvin, and K. L. Mulvey, "Using Augmented Reality to Promote Active Learning in College Science," in *Active Learning in College Science*: Springer, 2020, pp. 741-755.
- [10] S. Su *et al.*, "Virtual and augmented reality applications to support data analysis and assessment of science and engineering," *Computing in Science & Engineering*, vol. 22, no. 3, pp. 27-39, 2020.
- [11] N. Arulanand, A. R. Babu, and P. Rajesh, "Enriched learning experience using augmented reality framework in engineering education," *Procedia Computer Science*, vol. 172, pp. 937-942, 2020.
- [12] A. S. Saif, A. S. Prabuwno, and Z. R. Mahayuddin, "Real time vision based object detection from UAV aerial images: a conceptual framework," in *FIRA RoboWorld Congress*, 2013: Springer, pp. 265-274.
- [13] Z. R. MAHAYUDDIN and A. S. SAIF, "A COMPREHENSIVE REVIEW TOWARDS SEGMENTATION AND DETECTION OF CANCER CELL AND TUMOR FOR DYNAMIC 3D RECONSTRUCTION," *Asia-Pacific Journal of Information Technology and Multimedia*, vol. 9, no. 1, pp. 28-39, 2020.
- [14] A. S. Saif, A. S. Prabuwno, and Z. R. Mahayuddin, "Motion analysis for moving object detection from UAV aerial images: A review," in *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, 2014: IEEE, pp. 1-6.
- [15] Z. R. Mahayuddin and A. S. Saif, "A Comprehensive Review Towards Appropriate Feature Selection for Moving Object Detection Using Aerial Images," in *International Visual Informatics Conference*, 2019: Springer, pp. 227-236.
- [16] Z. R. Mahayuddin and A. S. Saif, "A COMPARATIVE STUDY OF THREE CORNER FEATURE BASED MOVING OBJECT DETECTION USING AERIAL IMAGES," *Malaysian Journal of Computer Science*, pp. 25-33, 2019.
- [17] Z. R. Mahayuddin, A. S. Saif, and A. S. Prabuwno, "Efficiency measurement of various denoise techniques for moving object detection using aerial images," in *2015 International Conference on Electrical Engineering and Informatics (ICEEI)*, 2015: IEEE, pp. 161-165.
- [18] A. S. Saif, A. S. Prabuwno, Z. R. Mahayuddin, and T. Mantoro, "Vision-based human face recognition using extended principal component analysis," *International Journal of Mobile Computing and Multimedia Communications (IJMCMC)*, vol. 5, no. 4, pp. 82-94, 2013.
- [19] A. Saif and Z. R. Mahayuddin, "Moving Object Segmentation Using Various Features from Aerial Images: A Review," *Advanced Science Letters*, vol. 24, no. 2, pp. 961-965, 2018.
- [20] A. Saif, A. Prabuwno, and Z. Mahayuddin, "Adaptive long term motion pattern analysis for moving object detection using UAV aerial images," *International Journal of Information System and Engineering*, vol. 1, no. 1, pp. 50-59, 2013.
- [21] A. Saif, A. S. Prabuwno, and Z. R. Mahayuddin, "Moving object detection using dynamic motion modelling from UAV aerial images," *The Scientific World Journal*, vol. 2014, 2014.
- [22] A. S. Saif, A. S. Prabuwno, and Z. R. Mahayuddin, "Adaptive motion pattern analysis for machine vision based moving detection from UAV aerial images," in *International Visual Informatics Conference*, 2013: Springer, pp. 104-114.

- [23] E. Richard, V. Billaudeau, P. Richard, and G. Gaudin, "Augmented reality for rehabilitation of cognitive disabled children: A preliminary study," in *2007 virtual rehabilitation, 2007*: IEEE, pp. 102-108.
- [24] A. M. Kamarainen *et al.*, "EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips," *Computers & Education*, vol. 68, pp. 545-556, 2013.
- [25] H. Kaufmann and D. Schmalstieg, "Mathematics and geometry education with collaborative augmented reality," in *ACM SIGGRAPH 2002 conference abstracts and applications, 2002*, pp. 37-41.
- [26] T.-J. Lin, H. B.-L. Duh, N. Li, H.-Y. Wang, and C.-C. Tsai, "An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system," *Computers & Education*, vol. 68, pp. 314-321, 2013.
- [27] N. N. Mahadzir and L. F. Phung, "The use of augmented reality pop-up book to increase motivation in English language learning for national primary school," *Journal of Research & Method in Education*, vol. 1, no. 1, pp. 26-38, 2013.
- [28] G. Salvador-Herranz, D. Pérez-López, M. Ortega, E. Soto, M. Alcañiz, and M. Contero, "Manipulating virtual objects with your hands: A case study on applying desktop augmented reality at the primary school," in *2013 46th Hawaii International Conference on System Sciences, 2013*: IEEE, pp. 31-39.
- [29] C. Alcoholado *et al.*, "One mouse per child: Interpersonal computer for individual arithmetic practice," *Journal of Computer Assisted Learning*, vol. 28, no. 4, pp. 295-309, 2012.
- [30] L. Kerawalla, R. Luckin, S. Seljeflot, and A. Woolard, "'Making it real': exploring the potential of augmented reality for teaching primary school science," *Virtual reality*, vol. 10, no. 3-4, pp. 163-174, 2006.
- [31] U. Von Jan, C. Noll, M. Behrends, and U.-V. Albrecht, "mARble—augmented reality in medical education," *Biomedical Engineering/Biomedizinische Technik*, vol. 57, no. SI-1-Track-A, 2012.
- [32] G. Bitter and A. Corral, "The pedagogical potential of augmented reality apps," *International Journal of Engineering Science Invention*, vol. 3, no. 10, pp. 13-17, 2014.
- [33] P. S. Medicherla, G. Chang, and P. Morreale, "Visualization for increased understanding and learning using augmented reality," in *Proceedings of the international conference on Multimedia information retrieval, 2010*, pp. 441-444.
- [34] I. Radu, E. Doherty, K. DiQuollo, B. McCarthy, and M. Tiu, "Cyberchase shape quest: pushing geometry education boundaries with augmented reality," in *Proceedings of the 14th international conference on interaction design and children, 2015*, pp. 430-433.
- [35] H.-K. Wu, S. W.-Y. Lee, H.-Y. Chang, and J.-C. Liang, "Current status, opportunities and challenges of augmented reality in education," *Computers & education*, vol. 62, pp. 41-49, 2013.
- [36] B. Parhizkar, W. K. Obeidy, S. A. Chowdhury, Z. M. Gebril, M. N. A. Ngan, and A. H. Lashkari, "Android mobile augmented reality application based on different learning theories for primary school children," in *2012 International Conference on Multimedia Computing and Systems, 2012*: IEEE, pp. 404-408.
- [37] K. R. Bujak, I. Radu, R. Catrambone, B. MacIntyre, R. Zheng, and G. Golubski, "A psychological perspective on augmented reality in the mathematics classroom," *Computers & Education*, vol. 68, pp. 536-544, 2013.
- [38] P. Diegmann, M. Schmidt-Kraepelin, S. Eynden, and D. Basten, "Benefits of augmented reality in educational environments—a systematic literature review," *Benefits*, vol. 3, no. 6, pp. 1542-1556, 2015.
- [39] Z. Mahayuddin and A. Saif, "Efficient Hand Gesture Recognition Using Modified Extrusion Method based on Augmented Reality," *TEST Engineering and Management*, vol. 83, pp. 4020-4027, 2020.
- [40] Z. R. Mahayuddin and A. Saif, "Augmented Reality Based Ar Alphabets Towards Improved Learning Process In Primary Education System," *Journal of Critical Reviews*, vol. 7, no. 19, pp. 514-521, 2020.
- [41] A. S. Saif, A. S. Prabuwo, and Z. R. Mahayuddin, "Moment feature based fast feature extraction algorithm for moving object detection using aerial images," *PloS one*, vol. 10, no. 6, p. e0126212, 2015.
- [42] A. S. Saif, A. S. Prabuwo, Z. R. Mahayuddin, and H. T. Himawan, "A review of machine vision based on moving objects: object detection from UAV aerial images," *International Journal of Advancements in Computing Technology*, vol. 5, no. 15, p. 57, 2013.

Feasibility Study of a Small-Scale Grid-Connected PV Power Plants in Egypt; Case Study: New Valley Governorate

Mahmoud Saad¹, Hamdy M. Sultan², Ahmed Abdeltwab³, Ahmed A. Zaki Diab⁴

Dept. of Electrical Engineering
Faculty of Engineering, Minia University
Minia, Egypt

Abstract—The construction of photovoltaic power plants (PVPPs) in the right place is an important task when planning the development of the power system and choosing investors. In this paper, the technical, environmental, the economic feasibility of installing a 50kW solar power plant in different places in the New Valley Governorate in Egypt has been presented using RETScreen Expert software. The input data used in the current study are obtained from the database of the Surface Meteorology and Solar Energy Dataset of NASA. In general, five sites for the construction of 50 kW power stations were assessed which represent the five administrative regions of the New Valley governorate. The study is based on annual electricity production, greenhouse gas (GHG) emissions, and financial analysis. With the proposed PV power plant, up to 100 MWh of electricity can be produced and a minimum of 43.3 tons of GHG emission can be prevented from the exhaust into the local atmosphere annually. The obtained results from the RETScreen program proved the viability of installing the proposed 50kW photovoltaic power plant in any of the proposed locations. This study could give a piece of important information and feedback that can be utilized as a database for upcoming investments in the photovoltaic generation projects in Egypt.

Keywords—RETScreen; new valley; solar energy; energy cost; feasibility analysis; greenhouse gases

I. INTRODUCTION

Renewable Energy is the energy that can be extracted and derived from renewable natural resources or that cannot be exhausted or developed (sustainable energy). There are several fundamental differences between renewable energy sources (RES) and fossil fuels, which include "coal, natural gas, and petroleum" or "nuclear fuel that can be used in nuclear reactors". Usually, new and renewable energy does not generate harmful wastes such as carbon dioxide or harmful gases or increase the danger regarding global warming, compared to what can happen when burning fossil fuels or harmful atomic waste and waste from nuclear power reactors. Renewable energy can be produced from the sun and wind. And water [1]. It can also be produced from tidal movements and waves, or geothermal energy, in addition to some productive trees and agricultural crops.

Renewable energy has several advantages and benefits, direct or indirect, summarized in the following points: Renewable energy is not running out; It can also provide clean

energy free of waste; it aims to protect human health; it can preserve the natural environment; it has Low production cost; it can improve human livelihood and reducing poverty; it can secure new job opportunities. Moreover, renewable energy has extra advantages like Reducing harmful gaseous and heat emissions and their dangerous consequences; Reducing greenhouse gas emissions and harmful heat and their dangerous consequences [2]; Reduced numbers and risks of natural disasters resulting from global warming; Not to form acid rain that harms all crops, agricultural products, and all different life forms; offering a significant reduction in the formation and accumulation of harmful waste in all its forms (gaseous, liquid and solid); Protecting all living organisms, especially the endangered species; Protecting groundwater, seas, rivers, and fisheries from the risk of pollution; Contribute to ensuring food security; In addition to increasing the productivity of agricultural crops as a result of eliminating chemical and gaseous pollutants [3].

The sun is one of the largest sources of light and heat on the face of the earth, and this energy is distributed over the parts of the earth according to its proximity to the equator, and this line is the area that receives the largest share of that energy, and the thermal energy generated by the sun's rays is used by converting it into (energy Electricity) by panels (solar cells). There are also two methods of collecting solar energy. The first one is that the sun's rays are focused on a collector employing convex mirrors. The collector usually consists of several tubes containing water or air [4]. The sun's heat heats the air or turns water into steam. The second method, in which a level plate collector absorbs the sun's heat, uses the heat to produce hot air or steam.

The geographical location of the Arab Republic of Egypt is between latitudes 22 and 31.5 north, and as a result, Egypt is located in the depth and value of the global sunbelt. As a result, Egypt has been considered one of the world's richest countries with solar energy [4]. Where the Ministry of Electricity and Energy conducted a lot of research and studies to determine and clarify the characteristics of solar radiation in the Arab Republic of Egypt, which resulted in updating information, data, and statistics that were provided by meteorological stations and centers and adding many new stations and modern measuring devices. According to the various economic events and developments that Egypt has gone through during the modern era and their clear impact on the projects implemented

in the field of renewable energy, and as a result of the flexibility of adequate strategies to keep pace with the changes and developments. Events, its renewable energy strategy has been implemented and adjusted to target 20% of the total production of energy during 2022 [2, 5, 6]. The energy sector in Egypt (electricity, renewable energy, and petroleum) has prepared a study for the optimal combination of technically and economically for energy production until 2035 in cooperation with the European Union through the Technical Support Program for the Restructuring of the Energy Sector in Egypt (TARES) [6]. The project included several parts, the most important of which was supporting the sustainable and integrated energy strategy in Egypt until 2035.

This study includes a set of scenarios for the energy mix with different assumptions to assess the impact of the introduction of renewable energies in different proportions to the electricity generation mix from a technical and economic perspective, to choose the optimal scenario. In October 2016, the Supreme Council of Energy approved the Egyptian Energy Strategy until 2035 and chose Scenario (4-B) as the reference for energy planning in Egypt during the coming period [7], which aims to reach the percentage of renewable energy contribution to 42% of the total amount of electricity production by 2035.

RETScreen simulation program is a clean energy management software that is used for energy efficiency studies and feasibility assessment of renewable energy cogeneration projects and performing the required analysis for energy performance [8]. This program empowers professionals and decision-makers to quickly point out, evaluate and optimize the viability of renewable clean energy projects from the technical and financial sides. RETScreen software also makes it easy for managers to estimate and verify the real performance of their projects and helps them to obtain supplementary energy savings/production opportunities. RETScreen program has been utilized in many research papers. In [9] the feasibility study of a 100MW photovoltaic power station at Bati, Ethiopia has been conducted and the results showed that 2365.3 tCO₂ will be reduced to be exhausted into the environment. In [10], the assessment of solar energy potential in Algeria has been evaluated. In this study, 61 different sites have been evaluated and the results showed that the southern regions of the country have a great solar energy potential. In [11], the pre-feasibility study of nine small-scale hydropower plants in Turkey has been evaluated, and the obtained results were compared with those reported in State Hydraulic Works of Turkey. It was observed that RETScreen is capable of performing pre-feasibility studies of these small hydropower plants within a relatively short period concerning the time taken by other classic methods. In [12] cost, financial, and risk analysis for a 100 kW wind power plant, containing two 50kW wind turbines, that is planned to be installed on Taşlıçiftlik Campus in Tokat has been accomplished using RETScreen software program. In [13] an environmental and economic assessment of a grid-connected PV system established in a public institution in Brazil using the RETScreen program, the results proved that

the project is environmentally viable but still needs tax remuneration to attract investors in the country. In [14] RETScreen Expert software has been utilized to implement Measurement and Verification analysis to evaluate the energy effectiveness of an existing building in Korea based on the collected data of 12 years of gas consumption. In [15] an economic analysis based on RETScreen to evaluate the case of transition from grid electricity to PV power plants for a domestic building existing in Ado ekiti, Nigeria. The results showed that the PV system is more economic than depending on the grid electricity and the cost of energy is relatively cheap. The RETScreen software has been used in evaluating the technical potential and economic feasibility of different renewable generation projects in different countries [16-18]. Taking into consideration the great dependency on RETScreen simulation software in performing feasibility studies of renewable energy sources, the authors are motivated to evaluate the technical and economic performance of the planned PV power plants in Egypt.

The New Valley is the largest and least densely populated governorate in Egypt. Its capital is Kharga and includes the southern half of the Egyptian part of the Libyan Desert. It is bordered to the north by the governorates of Minya, Giza, Matrouh, and to the east by the governorates of Assiut, Sohag, Qena, and Aswan, and to the south by the borders with Sudan and to the west by the border with Libya. The New Valley Governorate is characterized by the presence of vast land areas that can be used in the establishment of various projects in all industrial, service, agricultural, and tourism fields. The New Valley Governorate is considered one of the best governorates in Egypt in terms of solar radiation intensity. The average annual number of hours of sunshine per day ranges from 9 hours to approximately 11 hours in the southern desert of Egypt, which means greater investment opportunities in the domain of different solar energy applications [4, 19]. Therefore, we find that there is a trend, whether personally by farmers, to operate many water wells with solar energy or by government agencies recently. Fig. 1 presents the Global Horizontal Irradiance (GHI) that appears in all regions of Egypt [20].

In this paper, a study of the potential of a PVPP installation in New Vally is carried out. In general, five sites for the construction of 50 kW power stations were assessed according to their technical potential. The selected locations represent the five administrative regions of the New Vally governorate. The analysis of the feasibility of the project is carried out using the RETScreen software, taking into account the peculiarities of electric power production at the proposed power plant, economic feasibility, and reduction of greenhouse gas emissions.

The Rest of the paper is organized in the following manner: Section II introduces the climatic conditions in the proposed sites; Section III provides the technical specifications of the PV module used in this study; Section IV presents the obtained results and discussion; and finally, Section V is conducted for the conclusion and future work.

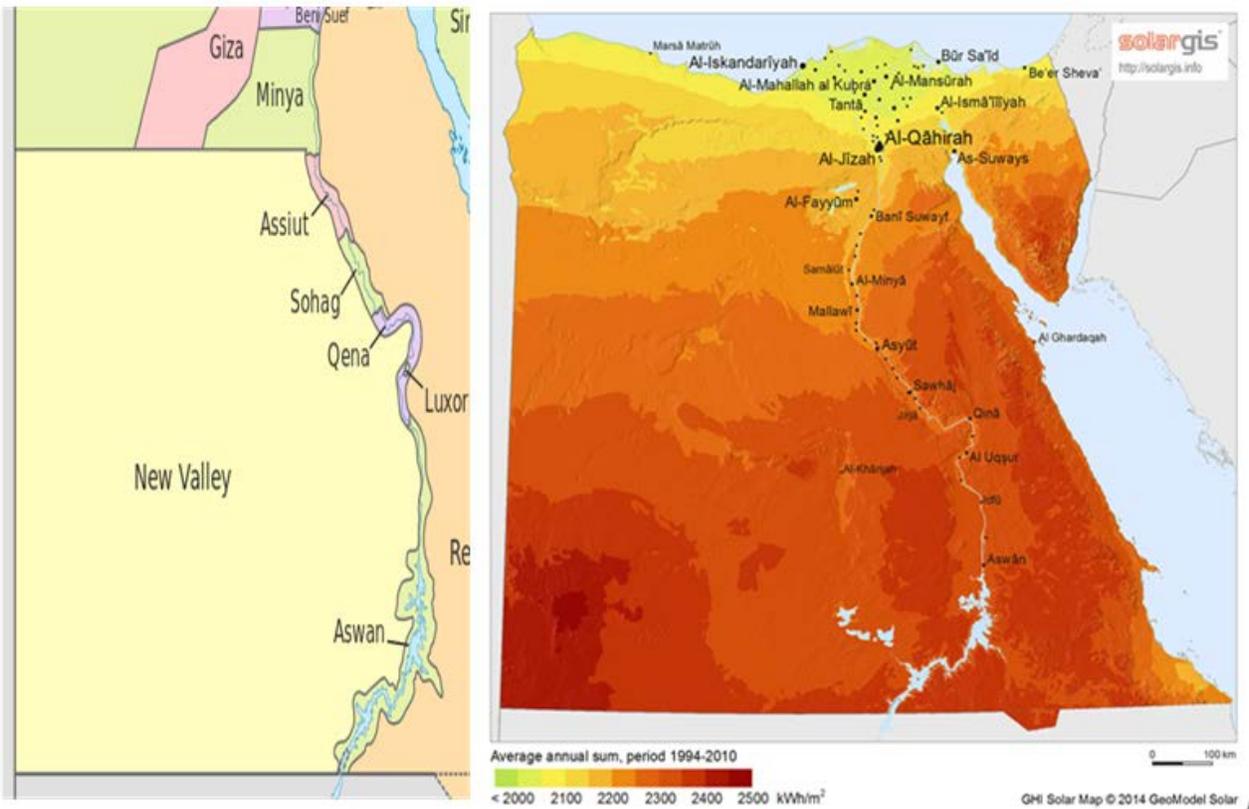


Fig. 1. Global Horizontal Radiation in Egypt.

II. CLIMATIC CONDITIONS OF THE PROPOSED REGION

From Cairo to the far south, Egypt receives radiation exceeding 6 kWh/m²/day, and the days on which clouds appear for most hours of the day are less than 20 days a year, and the total radiation increases from north to south, reaching a value of 5 kWh/m²/day "near the northern coast and exceeding 7 kWh/m²/day at the farthest point in southern Egypt, while the number of hours of sunshine exceeds 4000 hours annually. These numbers are among the highest in the world. The New Valley Governorate is considered one of the best governorates in Egypt in terms of the intensity of solar radiation and the average number of hours of sunshine. In the next subsections, the meteorological data of the five administrative regions of the New Valley governorate are presented.

A. Solar Radiation Intensity

The mean value of the solar radiation on the selected region for all months of the year is displayed in Fig. 2. From the shown figure, it has become clear that during the summer months the greatest value of the solar radiation is observed in Frafra while during winter the greatest value of the solar radiation is obtained in Paris and Kharga regions. The average, greatest, and least values of the solar radiation in the five locations are presented in Fig. 3. From this figure, it is noticed that Frafra region is characterized by the greatest and least values of solar radiation in New Valley. The mean value of the solar radiation in Paris and Kharga is the best within the five locations. Fig. 4 shows the variation of the mean value of the solar radiation, calculated by the mean value of the solar radiation over the five regions, over the year. It is obvious from

the figures that the least solar radiation value is 5.88 kWh/m²/day, and it is found in Frafra and the greatest solar radiation value is 6.28 kWh/m²/day, and it is found in both Kharga and Paris.

B. Air Temperature

The manners of air temperatures have been presented in Fig. 5, 6, and 7 for the different locations. The monthly values of the air temperature over the selected sites are shown in Fig. 5, while the greatest, least, and average values of the air temperature in the proposed regions are shown in Fig. 6. It is obvious from the figures that the greatest air temperature value is 33.4°C and it is found in both Kharga and Paris. The least air temperature value is 11.3°C and it is found in Frafra. The lower values of air temperature were detected during winter months while the greater values in winter as expected.

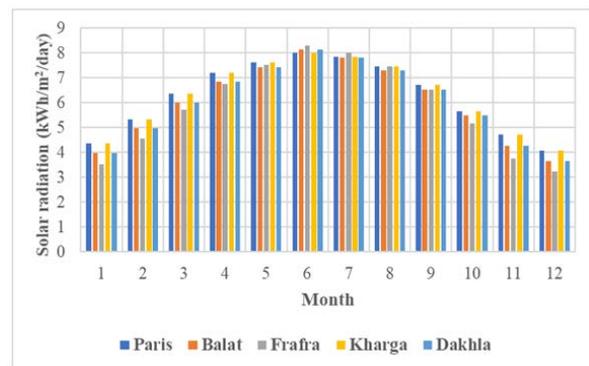


Fig. 2. The Monthly Values of Solar Radiation of Different Locations.

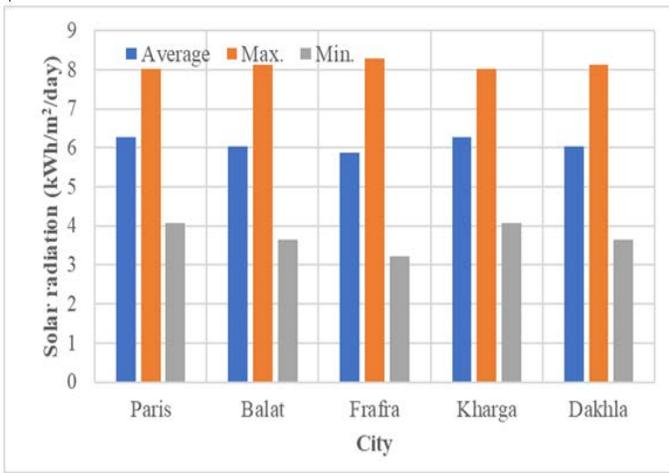


Fig. 3. Long-Term Average Solar Radiation.

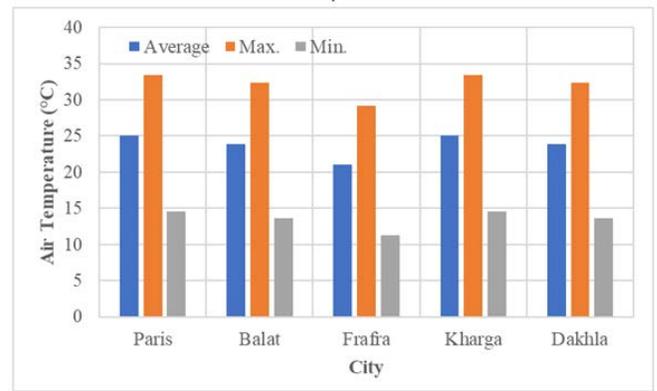


Fig. 6. Long-Term Averaged, Maximum and Minimum Air Temperature.

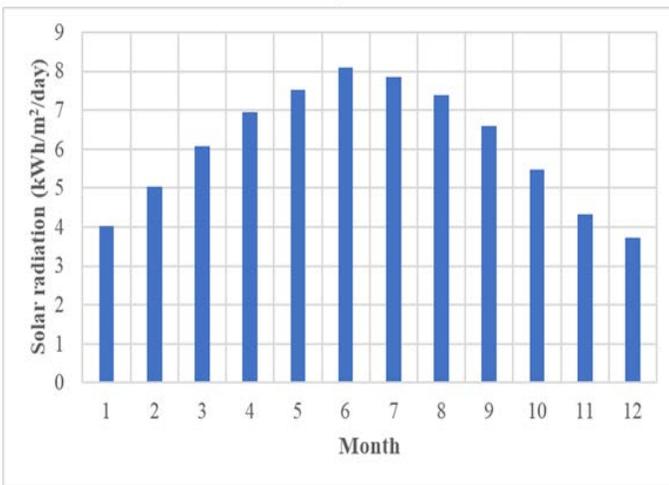


Fig. 4. The Total Variation in All Seasons of Global Solar Radiation in the New Valley.

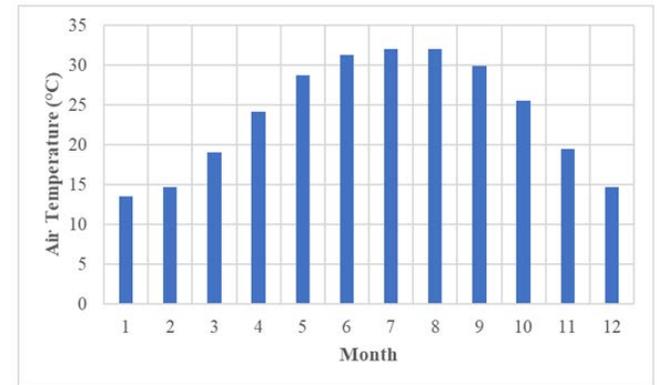


Fig. 7. The Total Variation in All Seasons of Air Temperature in New Valley.

C. Relative Humidity

The behavior of relative humidity can be discussed in Fig. 8, 9, and 10. It is obvious from the figures that for climatic conditions and relative humidity one can find that Frafra has the highest relative humidity 41.5%, the lowest relative humidity found in both Dakhla and Balat and equals 38.4%. The higher values of relative humidity were detected during winter months and lower in summer months as prospective.

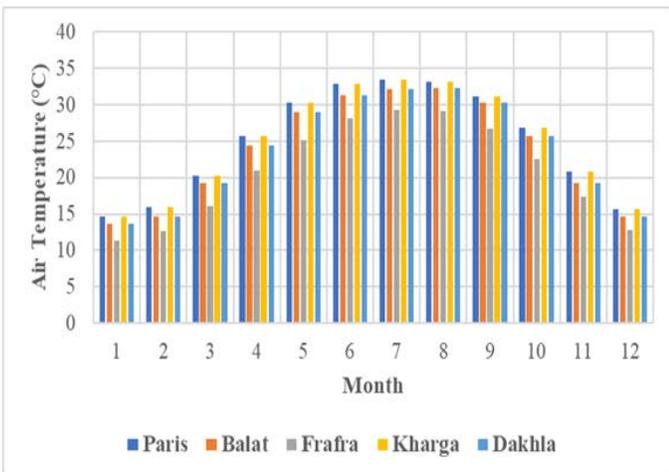


Fig. 5. The Total Monthly Variance of Air Temperature At Different Locations.

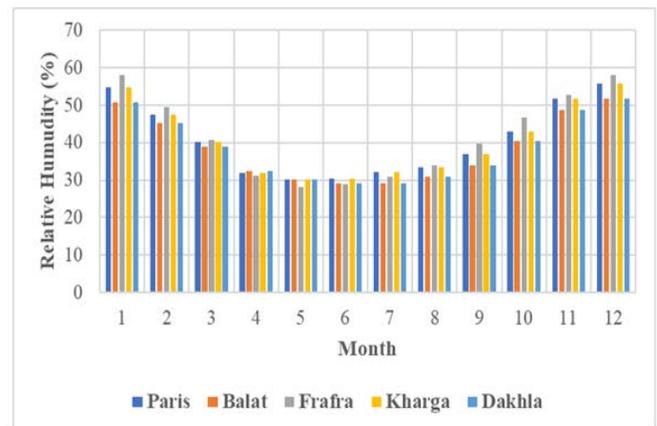


Fig. 8. The Monthly Values of Relative Humidity (%) in Different Locations.

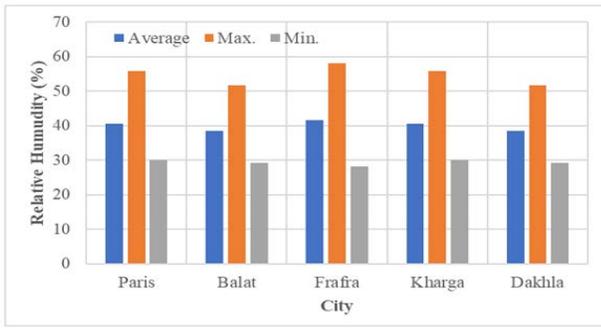


Fig. 9. Long-Term Averaged, Maximum, and Minimum Values of Relative Humidity (%).

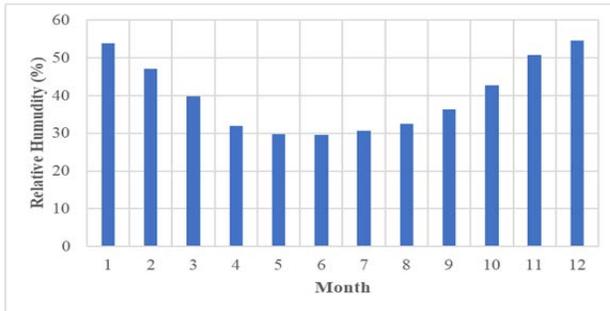


Fig. 10. The Total Variation in All Seasons of Relative Humidity (%) in New Vally.

III. PV MODULE SELECTION

The appropriate type of the photovoltaic module was chosen according to the proportion of the energy generated to the unit and its surface area (power/surface area 350w mono si-solaria power spandrel with a number of units equal 143 unit). The required PV module specifications are provided in Table I [21]. 143 units needed to establish the required 50 kW station. In our work, the fixed axis position of the modules will be chosen to reduce the initial cost 60kW inverters are selected with 95% efficiency.

TABLE I. SPECIFICATIONS OF THE CHOSEN PV MODULE

Model:	350w mono si-solaria power spandrel
Electrical Specifications	
Power P_{max}	350 W
Efficiency	17.3%
Open Circuit Voltage (V_{oc})	47.0V
Short Circuit Current (I_{sc})	9.6A
Max Power Voltage (V_{mp})	39.0V
Max Power Current (I_{mp})	9.0A
Temperature Performance	
Coefficient of P_{max}	-0.40% /C
Coefficient of V_{oc}	-0.30% /C
Coefficient of I_{sc}	+0.05% /C
Mechanical Specifications	
length	1495 mm [58.8in]
Width	1350 mm [53.1in]
thickness	11.5 mm [0.45in]
Weight	52 Kg [114lbs] approx

IV. RESULTS AND DISCUSSION

In this study, a 50kW PV power plant is proposed in the five locations and feasibility analysis of installing such a power plant in the New Vally region is assessed. The feasibility study is accomplished using RETScreen software. RETScreen Which means Clean Energy Management Software (usually abbreviated to RETScreen) is a package of programs for clean energy, improved and developed by the Canadian government that is an Excel-based software tool for making analyzes of clean energy projects, which helps decision-makers know both financial and technical feasibility in the field of energy efficiency, renewable energy and cogeneration (integrating heat and power generation) [8]. Conventional energy projects can also be designed and compared to more conservative alternatives. The program has the access to database of NASA and the values of the solar radiation on the horizontal surface, air temperature, relative humidity, and other meteorological data are applied to the RETScreen software [8, 22]. The program output is the annual total of energy produced from the proposed plant, the annual reduction of the GHG emissions thanks to the installed power plant in the selected locations.

A. Renewable Energy Production

The total annual power produced (MWh) is obtained from the RETScreen program as an output. Fig.11 shows the amount of energy produced in the five stations. The maximum energy amount is 100 MWh produced at Paris station. The minimum energy amount is 95 MWh produced at Balat station.

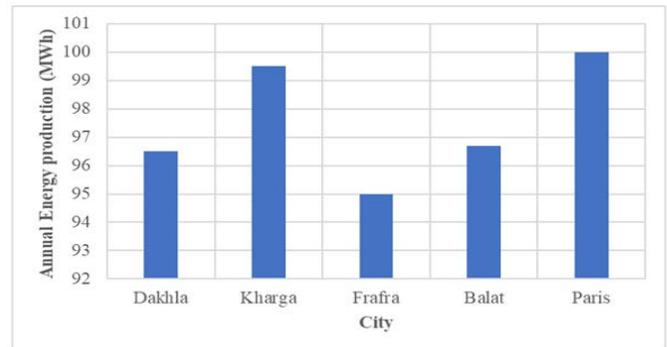


Fig. 11. Annual Energy Production of Different Stations.

B. Greenhouse Gases Reduction

One of the essential issues for validating the best-case study location is the Greenhouse Gases (GHG) Reduction. So, the model estimates the overall percentage and annual reduction in the amount of greenhouse gas emissions that would be incurred if the proposed case were implemented. The calculation is dependent on emissions of both the base case and the state systems selected on yearly basis. The units are given in CO2 equivalent annually per year (tCO2/yr). The annual GHG emission reduction thanks to the installation of the proposed 50kW PVPP in the five sites is presented in Fig. 12. According to this figure, it is noted that Kharga station has the highest rate of preventing greenhouse gases by 45 tons of CO2 can be prevented to reach the atmosphere. On the other hand, the Frafra station has the lowest rate of preventing greenhouse gases by 43.3 tons of CO2 can be prevented to reach the atmosphere.

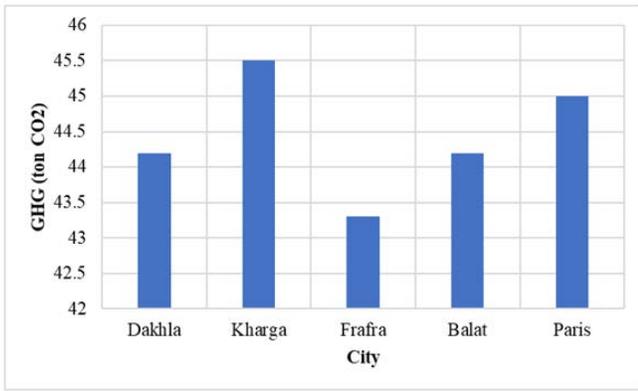


Fig. 12. Annual Reduction of GHG in the Selected Locations.

C. Economic Feasibility Analysis

RETScreen has been utilized to determine a series of economic indicators such as simple payback period (SPP), internal rate of return (IRR), benefit to cost ratio (BCR), net present value (NPV), annual life cycle saving (ALCS), years of positive cash flow (YPCF), and energy production cost (EPC) for each project (50kW PV power plant). These indicators are explained in detail in the following subsections. The initial economic parameters, like the growth rate of energy cost, the rate of inflation, plus the cost of the main components of the PV power plant, are taken following the data given in Table II [13, 14].

TABLE II. FINANCIAL PARAMETERS THAT ARE USED IN THE STUDY [22, 23]

Parameter	Value
Project life	25 years
Inflation rate	2.5%
Discount rate	5%
Reinvestment rate	9%
Fuel cost escalation rate	0.0%
Debt interest rate	7%
Debt term	20 years
Electricity export rate	0.1\$/kWh

a) *Internal Rate of Return (IRR)*: Through IRR, companies can learn the feasibility of investment of the projects in the long term as it is one of the methods of capital budgeting. It is the discount rate at which the present value result is equal to zero, and it is usually explained by the expected profits that an investment decision generates [5]. In general, If the rate exceeds the capital value for the project, this will represent an added value for the project. IRR is calculated as given in equation (1):

$$Initial\ Investment = \sum_{t=1}^N \frac{C_t}{(1 + IRR)^t} \quad (1)$$

where C_t denotes the Net of cash inflow during the period called t , IRR is defined as the internal rate of return, N represents the lifetime of the project, and t represents the number of time intervals.

b) *Simple Payback Period (SPP)*: A measure of the project's liquidity. It refers to the period of time during which the initial cost is recovered from the cash receipts, and this method is based on the fact that the more investment value is recovered in a shorter time, the more acceptable the investment or the paid-back capital period, which is one of the capital budgeting tools that the investor or financial manager relies on to implement an investment project. The higher the recovery period of the invested money in a shorter time, the better the investment in this project and vice versa.

c) *Years of Positive Cash Flow (YPCF)*: It refers to the movement of money in or out of a business, enterprise, or financial product. It is usually measured over a specific, limited period of time. It indicates that the company's core business activities are thriving. It provides an additional measure of an indication of the company's profitability potential.

d) *Annual Lifecycle Saving (ALCS)*: It is a nominal annual saving with an equation, with the exact value of the NPV discount rate and the project life, the value of ALCS is calculated.

e) *Net Present Value (NPV)*: It is defined as the present equivalent value of future payments. One of the tools that companies use to evaluate investment projects (long-term projects); The method of working the net present value depends on ensuring that the project being evaluated generates cash flows above the value invested in the project [5, 23]. The financial flows of the coming years can be transformed to the net present value by the next equation:

$$NPV = \sum_{t=0}^N \frac{C_t}{(1 + r)^t} \quad (2)$$

where C_t is annual cash flow, r represents the annual interest rate of annual interest, and t represents the current year.

f) *Benefit-Cost Ratio (BCR)*: An indicator used in the field of cost-benefit analysis that attempts to arrive at the total value of money needed by the project. The higher the ratio is more than 1, the better the project, and the safer investment in it:

$$BCR = \frac{Discounted\ value\ of\ fringe\ benefits}{The\ value\ of\ the\ additional\ costs} \quad (3)$$

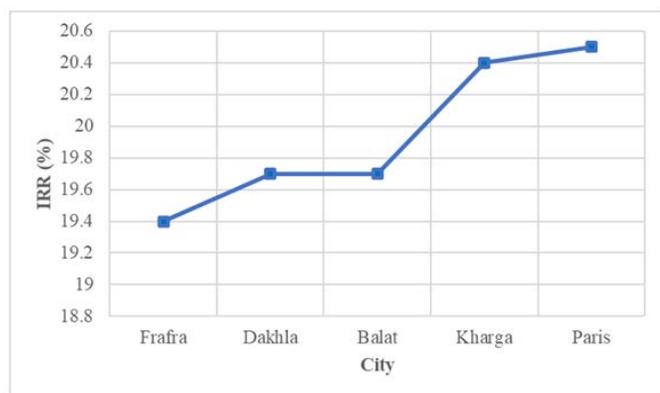
g) *Energy Production Cost (EPC)*: Energy production cost per kilowatt-hour for generated electricity. This value is equivalent to the electricity export rate required to make a net present value (NPV) equal to zero.

The variation of the aforementioned economic factors concerning the location of the PVPP is presented in Fig. 13. If the value of IRR equals to or greater than the required return rate value, the study will be passed. The IRR values of all sites are shown in Fig. 13(a). The greatest value of the internal rate of return (IRR) is 20.5% found at Paris station while the minimum IRR value is 19.4% found at Frafra station. For SPP which equal to the period of time during which the initial cost is recovered from the cash receipts. The values of SPP and

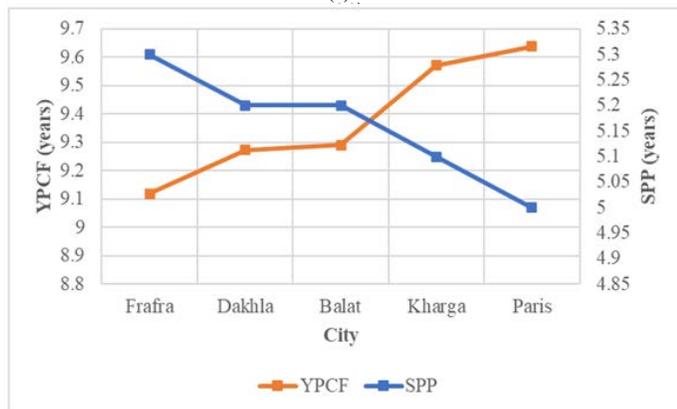
YPCF for all locations are shown in Fig. 13(b). The greatest value of the simple payback period is 5.3 years found at Frafra station while the minimum value is 5 years found at Paris station. For YPCF which indicates how long the project will take to recoup its initial investment. The maximum value of YPCF is 9.638 years found at Paris station. The minimum value of YPCF is 9.119 years found at Frafra station.

The values of NPV and ALCS of all sites are shown in Fig. 13(c). ALCS can be calculated using the discount rate, the NPV, and the life of the project. ALCS has a maximum value equal is 6.358 years found at Paris station. The minimum value of ALCS is 5.83 years found at Frafra station. The maximum value of net present value is 89.605 million USA\$ found at Paris station while the minimum value of NPV is 82.28 million USA\$ found at Frafra station. Positive NPV values ensure project feasibility. From the shown values in the figure, we can see that all the values are positive which ensures that all five stations are potentially feasible.

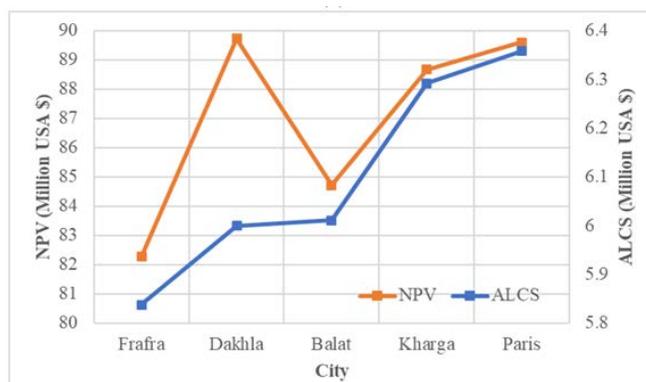
All the available values of BCR are greater than one so that it is an indicator for profitable and better projects. The values BCR for all locations are illustrated in Fig. 13(d). The greatest value of BCR is 2.9 at Kharga and Paris stations while the minimum value of BCR is 2.8 at Dakhla, Balat, and Frafra stations. The generating unit's production cost known as Energy Production Cost (EPC) is one of the outputs of the RETScreen program. The value EPC of all sites is shown in Fig. 13(e). The minimum value of EPC equals 3.7 cents/kilowatt hour in Kharga and Paris stations while the maximum value EPC equals 3.9 cents/ kilowatt hour in Frafra station.



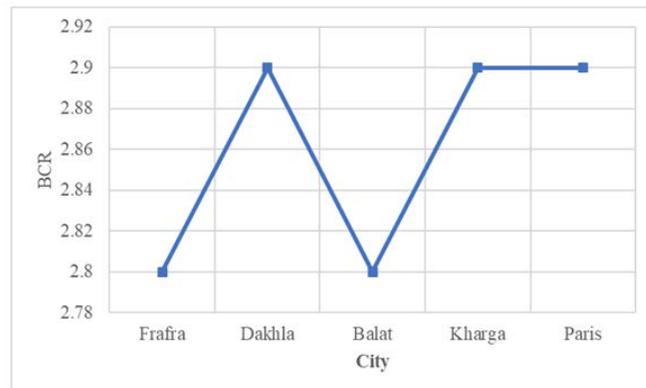
(a)



(b)



(c)



(d)



(e)

Fig. 13. Economic Indicators Varied According to the Location of the Photovoltaic Power Plants.

V. CONCLUSION

This paper presented a comprehensive study of the potential of a PVPPs installation in New Vally governorate, Egypt. For completing this study, five sites have been assessed for installing the same power plant with a capacity of 50 kW. This study has been accomplished to have a map for solar energy projects that will increase the international investment in the underutilized and superabundant environmentally friendly energy source in southwestern Egypt. After analyzing the meteorological conditions in the region it was obtained that the minimum solar radiation in the studied area is 5.8 kW/m²/day and obtained in Frafra station, while the maximum solar radiation is 6.04 kW/m²/day in the Kharga and Paris stations. The amount of annual energy produced from the

proposed PV power plants in the suggested locations ranges from 95 MWh at Frafra station to 100 MWh at Paris station, which ensures the small variation in the output of the PV power plant regarding the location in the New Valley governorate. From the environmental side, it has been found that the 50kW PV power plant located in Kharga oasis has the highest rate of reduction in the greenhouse gases emissions with 45 tons of CO₂ that can be prevented to reach the atmosphere annually, in addition, it has been found that the PV power plant located in Frafra oasis has the lowest annual greenhouse gases reduction by 43.3 tons of CO₂. From the financial point of view, it was concluded that the installation of a solar PV power plant in any of the proposed places is economically feasible, and the cost of energy produced by solar power plants ranges from 3.7 to 3.9 cents per kWh. All the proposed sites are fully recommended for the installation of PV projects from all the accessible indices, but the proposed PV power plant located in Paris is considered to be the best choice because of the high energy exported to the grid utility. The payback period of the PV project in Paris is also the shortest compared to other sites.

It is recommended that solar photovoltaic power plants may be installed in Paris to acquire new experiences and endorse the PV technology in the regional environment. Based on the results obtained in this study, the authors in their future work are looking forward to studying the impact of the integration of such PV power plants into the distribution network in Paris on the power quality indicators in the distribution network.

REFERENCES

- [1] H. M. Sultan, A. A. Z. Diab, N. K. Oleg, and S. Z. Irina, "Design and evaluation of PV-wind hybrid system with hydroelectric pumped storage on the National Power System of Egypt," *Global Energy Interconnection*, vol. 1, no. 3, pp. 301-311, 2018.
- [2] A. A. Z. Diab, H. M. Sultan, and O. N. Kuznetsov, "Optimal sizing of hybrid solar/wind/hydroelectric pumped storage energy system in Egypt based on different meta-heuristic techniques," *Environmental Science and Pollution Research*, pp. 1-23, 2019.
- [3] A. A. Z. Diab, H. M. Sultan, T. D. Do, O. M. Kamel, and M. A. Mossa, "Coyote optimization algorithm for parameters estimation of various models of solar cells and PV modules," *IEEE Access*, vol. 8, pp. 111102-111140, 2020.
- [4] K. D. Patlitzianas, "Solar energy in Egypt: Significant business opportunities," *Renewable energy*, vol. 36, no. 9, pp. 2305-2311, 2011.
- [5] H. M. Sultan, O. N. Kuznetsov, and A. A. Z. Diab, "Site selection of large-scale grid-connected solar PV system in Egypt," in *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2018, pp. 813-818: IEEE.
- [6] (2021, 20 January 2021). The official website of the Ministry of Electricity and Energy. Available: <http://www.moee.gov.eg>.
- [7] T. E. N. R. E. A. (NREA), "Public Disclosure for EIA for 20MW Photovoltaic Power Plant Project in Hurghada," 2012.
- [8] (2021, 20 January 2021). RETScreen International. Renewable energy project analysis software. Available: <http://www.etscreen.net/>
- [9] J. B. Sy, A. Haile, and W. Degife, "Feasibility Study of a 100MW Photovoltaic Power plant at Bati, Ethiopia Using RETScreen."
- [10] S. Bouchakour, A. H. Arab, F. Cherfa, S. Semaoui, and A. Razagui, "SOLAR ASSESSMENT IN ALGERIA USING RETSCREEN DATABASE."
- [11] M. I. Yuce and S. Yuce, "Pre-feasibility Assessment of Small Hydropower Projects in Turkey by RETScreen," *Journal-American Water Works Association*, vol. 108, no. 5, pp. E269-E275, 2016.
- [12] Y. İÇMEZ, C. EMEKSİZ, and Z. DOĞAN, "Wind Power Plant Feasibility Study in Tokat with RETScreen Analysis Program," *Journal of New Results in Science*, vol. 5, no. 11, pp. 56-63, 2016.
- [13] L. S. de Freitas, N. A. Sousa, A. N. Pinheiro, and M. L. M. de Oliveira, "Feasibility of An on-Grid Photovoltaic System: Case Study Using Retscreen1," *Revista Brasileira de Energias Renováveis*, vol. 6, no. 4, pp. 763-786, 2017.
- [14] A. B. Owolabi, B. E. K. Nsafon, J. W. Roh, D. Suh, and J.-S. Huh, "Measurement and verification analysis on the energy performance of a retrofit residential building after energy efficiency measures using RETScreen Expert," *Alexandria Engineering Journal*, 2020.
- [15] S. Sanni and K. Mohammed, "Residential Solar Photovoltaic System Vs Grid Supply: An Economic Analysis Using RETScreen™," *Journal of Solar Energy Research*, vol. 3, no. 2, pp. 107-114, 2018.
- [16] Y. Pan, L. Liu, T. Zhu, T. Zhang, and J. Zhang, "Feasibility analysis on distributed energy system of Chongming County based on RETScreen software," *Energy*, vol. 130, pp. 298-306, 2017.
- [17] A. B. Owolabi, B. E. K. Nsafon, J. W. Roh, D. Suh, and J.-S. Huh, "Validating the techno-economic and environmental sustainability of solar PV technology in Nigeria using RETScreen Experts to assess its viability," *Sustainable Energy Technologies and Assessments*, vol. 36, p. 100542, 2019.
- [18] K. Y. Kebede, "Viability study of grid-connected solar PV system in Ethiopia," *Sustainable Energy Technologies and Assessments*, vol. 10, pp. 63-70, 2015.
- [19] A. A. Z. Diab, S. I. El-ajmi, H. M. Sultan, and Y. B. Hassan, "Modified farmland fertility optimization algorithm for optimal design of a grid-connected hybrid renewable energy system with fuel cell storage: case study of Ataka, Egypt," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 8, 2019.
- [20] (2021, 20 January 2021). Surface meteorology and solar energy. NASA renewable energy resource website. Available: <http://eosweb.larc.nasa.gov/sse/>.
- [21] (2021, 20 January 2021). Power-Producing Spandrels for Buildings. Available: https://static1.squarespace.com/static/568f7df70e4c112f75e6c82b/t/59951dac1e5b6c2acef8d1d9/1502944685219/Datasheet_SolariaPowerSpandre1_350_R5_8-4-17.pdf.
- [22] S. Rehman, M. Ahmed, M. H. Mohamed, and F. A. Al-Sulaiman, "Feasibility study of the grid connected 10 MW installed capacity PV power plants in Saudi Arabia," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 319-329, 2017.
- [23] A. S. Al-akayshee, O. N. Kuznetsov, and H. M. Sultan, "Viability Analysis of Large Photovoltaic Power Plants as a Solution of Power Shortage in Iraq," in *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2020, pp. 1145-1150: IEEE.

Augmented Reality Adapted Book (AREmotion) Design as Emotional Expression Recognition Media for Children with Autistic Spectrum Disorders (ASD)

Tika Miningrum¹, Herman Tolle², Fitra A Bachtiar³
Faculty of Computer Science, Brawijaya University
Malang, Indonesia^{1,2,3}

Abstract—One of the Autism Spectrum Disorder (ASD) characteristics is their difficulty understanding other people's emotions. Their lack of skill of understanding emotion includes expression and appropriate emotional response for a certain situation. This paper proposes an adapted book that helps therapists and parents guide ASD children to learn facial emotional expression. The adapted book combined with video, animation, and Augmented Reality increases children with ASD at recognizing emotional expression. This research uses *User-Centered Design* (UCD) approach to design the AR Adapted Book application and design the social story to be observable in school and family environment. Based on usability testing tried on the application, the result shows that AREmotion has an average score of 82.73 percent, and the learnability aspect has the highest score of 86.7 percent. This preliminary usability testing proves that the design of the AREmotion application is ready to use in real implementation for children with ASD.

Keywords—Autism Spectrum Disorder (ASD); adapted book; Augmented Reality (AR); User-Centered Design (UCD)

I. INTRODUCTION

Recognizing facial emotional expression is an important part of daily life to build an excellent interpersonal relationship [1]. Yet, not all people, in general, can easily recognize other people's facial emotional expressions. People with ASD (Autism Spectrum Disorder) have difficulties recognizing facial expressions [2]. Autism is a developmental and neurobiological disorder that has powerful effects on children's growth and influences their communication skills and social interaction in daily life [3]. People with ASD spreads all over the world, and the people with ASD total number is not small. The World Health Organization (WHO) stated that one of 160 children is a child with Autism Spectrum Disorder [4]. This estimate is the average result of a population in a certain period in all studies that have been carried out. Since the time of their growth, children with ASD have a problem in some social interaction area, i.e. (1) The development of their interpersonal relationship is blocked. (2) The spontaneous willingness to share hobbies and interests is lacking. (3) A significant disruption in body language, eye gazing, and facial expression when they do social interaction [5]. There is bad impact if ASD children do not study about facial emotional expression. They will grow up with lack of empathy and they find difficulty to give appropriate reaction to other people emotional expression. To avoid those

problems, a therapy is needed for ASD children so they can studying recognize other people expression in early stage.

A therapy, treatment, and guidance performed by therapist, parents, shadow teacher is needed to train the children with ASD to be able to learn how to interact socially. Parents and professionals, especially in Indonesia, stated that they need more coordination to give education, therapy, and care to children with ASD [6]. Children spent most of their time with family. This situation puts parents and family in the crucial role to treat the children with ASD very carefully. The treatment should be done correctly as every growth of children with ASD could inhibit the children growth phase. As a result, the children are in a high risk to be isolated from society. The isolation is due to their studying ability, lack of communication skills, and social stigma of children with ASD [7]. To overcome the described problem, a medium is needed to help therapists and parents to be able to take care of ASD Children in correct manner, especially so that the therapy for ASD children becomes more easily yet effective in treating the children.

There are various methods can be used to conduct a therapy for children with ASD. One of the therapies is using the sketch from drawing or art. Sketch art is a psychotherapy medium for children in general [8]. In case either normal children or children with ASD have difficulties or disruption, sketch art can be used so they can communicate in a non-verbal manner [9]. In general, therapists for ASD use a flashcard with sketch art or an illustration book as a supporting medium in learning and recognizing facial emotional expressions.

Previously, various technology has been developed to increase the quality of education and learning process of children with ASD. The research entitled The Impact of Technology on People with ASD [10] stated that there are four categories in learning with technology for people with ASD. The categories are conceptual skill (cover the subtopics: language, mathematics, colors, money, science, and programming), practical skill (subtopics: health, transportation, and daily life), social skill (subtopics: interpersonal relationship, communication, emotion) and general skill. The researcher explained that the most recent examination focused on utilizing technology to increase children's social skills with ASD. The study also emphasized that user needs, experience, and comfort are very important

considerations in developing technology for children with ASD. The method for developing development designs based on user needs is User Centered Design (UCD).

One technology incorporated used in children ASD therapy is Augmented Reality (AR). Augmented Reality is a technology that visualizes virtual objects generated by a computer, and the virtual object resides in a real-world environment to give an interactive experience to the user. The virtual object consists of 2D or 3D objects, video, animation, and audio. Augmented Reality technology can be applied to many fields such as entertainment, advertising, education, and medicine [11]. Using AR in education, students will experience new things and give more exciting and interactive ways during the learning process [12]. However there are a lot of research about AR technology with positive findings that explained on systematic review that Augmented Reality Technology for ASD patients show that 56 % social skill learning using AR [13]. The study also state that social skills, cognitive skills, and communication skills can be improved by using AR-based learning. Most of AR learning implementation is carried out on children and adolescents with ASD on age range of 3-14 years old.

Previous research carried out innovation in interactive media by combining game books and Augmented Reality on tablet device to learn expression to children with ASD [14]. User with ASD confronted with choosing expression that match with scenario story that presented according the real environment, the game book was developed with mobile based technology. The author believed that the game book application would improve social skills of ASD children. The limitation of this research was they did not explain how to develop the application and social story on the game book, and it was not explained either how significant the used of application for the development of ASD children.

Previous research discussed ARSFM (Augmented Reality Self Facial Modeling) [15]. This research aimed the goal to help children with ASD to improve their social skills. The ARSFM used a mask with facial emotional expression to train their ability to recognize facial expressions. The result from the research, there was a significant difference between baseline and follow-up lines. The AR technology is implemented on a personal computer, not free to move while using the application. Furthermore the research using ARVMS (Augmented Reality Based Video Modeling Storybook), to teach social skills for children with ASD by understanding the expression and emotions of other people in the social environment [16]. ARVMS learning system has 2 layers : the first one is static image layer, and the second one contains videos. The system has 20 videos that focus on social interaction. These studies also confirmed that augmented reality can help children with ASD to improve their social skills through expression. Several limitations in this study undeveloped technology based on User-Centered Design principles in order to comply the requirement of ASD children learning process. The research also did not include how they started make the design from early stage and there is no assessment of the application using usability testing after the application was developed.

The current study proposed an Augmented Reality Adapted Book that combines static sketch in flashcards and adapted books as a tool for children with ASD. Each of these two combination tools is a common tool used by the therapist to help ASD children learn facial emotional expression. This research started by conducting an observation. The process of designing the adapted book's story and the application uses the User-Centered Design (UCD) approach. This technique will assist in the development process, to ensure the design meets the needs of children with ASD. After that, ASD's therapists will evaluate it by assessing the usability level of the developed application.

The hypothesis in this research is augmented Reality with virtual objects, i.e., animation, audio, and video, can help ASD children increase their ability to recognize facial emotional expressions. Comfortability and enjoyment experience in using the application will be aimed by using the User-Centered Design (UCD) principle to develop a technology on a handheld device. This research also hopes to help parents and therapists guide children with ASD in learning facial emotional expression. Therefore, people with ASD need special attention and special treatment to become people who can recognize and distinguish other people's facial emotional expressions.

II. MATERIALS AND METHODS

To comply with a request of ASD children's facial emotional recognition application developments, it should be designed as its requirement list. End user and stake-holder's experience are one of the important part so the product agrees with the requirement, the goals, and learning process for ASD children. Three special needs teachers was involved in the process of iteration to made AREmotion application. The iteration is based on User Centered Design, and from this step the research gets feedback from end user or stakeholder. The feedbacks influence the design so a user friendly application can be developed for ASD children [19]. Indirectly, UCD is a continual process a combination of direct identification, development, and brainstorming processes by involving several people or professionals from several disciplines to produce a product or media according to user needs [17].

To be more precise, this research used UCD for design and develop the AR application. There are four steps insist: 1) Specify the context of use to store user information, 2) Specify user requirements to gets information and start the product modification by inputting feedback that collected from previous step, 3) Design and development, 4) Evaluation of the design as an assessment in order to improve the design, so end user can comfortably interact with the application and the application can help them learning something. The methodology diagram is presented in Fig. 1.

In the development process of the AREmotion application for children with autism, 2 literacy sessions were carried out. This iteration process applies to the process of making stories in the adapted book (ARBook) and the development of the AREmotion application as a whole accompanied by a special needs teacher. In the evaluation phase, usability testing was carried out, which was carried out on the therapist for children with autism. Evaluation is an effort to find out whether the

overall design of the AR application developed in this study is easy enough before being used/implemented for special users, namely children with ASD.

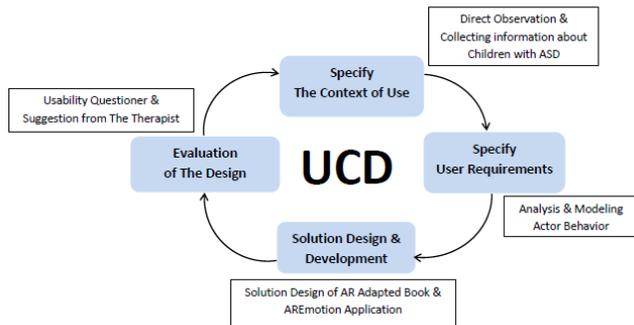


Fig. 1. UCD Process in AREmotion Application Development.

A. Understanding and Specify of Context of use

In the first phase, to understand and specify the context of use, direct observation is made on a Special Needs School. The teachers at the school were asked and interviewed to get some information about ASD children. Other information collected from the interview is: a general picture of ASD children's activity, their social interaction with school peers, and learning methods and therapy applied for them. The collected data will be used as a reference when designing the story and illustration on the adapted book. The information gathered from the interview, explained in Table I.

TABLE I. INFORMATION ABOUT ASD CHILDREN ACCORDING TO SPECIAL NEEDS TEACHER

No	Information
1.	In the studying process, ASD children often need a familiar guide (parents/teacher/therapist/shadow teacher)
2.	ASD children experience difficulties in understanding the abstract concept of social interaction, i.e., facial emotion recognition.
3.	It is often hard for ASD children to focus on their studies.
4.	Therapy and learning methods for ASD children rarely use an updated technology.
5.	ASD children are more interested in visual objects such as pictures, photos, and videos.

B. Understanding and Specify user Requirement

After getting information about therapy and the learning process, the next step is modeling the actor's behavior, as explained before using the case diagram. The diagram in Fig 2 shows that the main actor in this research is children with ASD. The children will interact with the application in four different use cases. The first use case is viewing the animation through Augmented Reality. The second is viewing the video footage menu for human emotion and expression. The third is finding the multiple-choice to answer the problem on ARBook through AR marker. And the last is checking the notification 'True' or 'False' on their answer through augmented Reality.

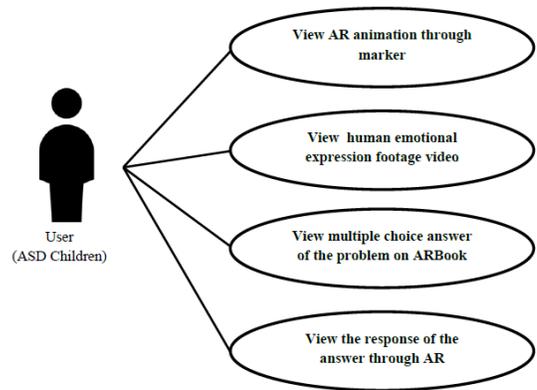


Fig. 2. Use Case Diagram for AREmotion Application.

III. DESIGN AND DEVELOPING OF AREMOTION APPLICATION

A. Design of Adapted Book (ARBook)

1) *Design of the story of adapted book:* The first design of storybooks that the researcher used is related to the book titled "A guide to producing written information in easy read" [18]. This book suggests using simple words, clear fonts, and sentences with less than 15 words, using pictures that are easy to understand and support the target sentences, and the images are made in the possible largest size.

After showing the first storybook's design to the ASD special needs teachers, they suggested dividing each expression into different categories. Every category contained some actions (in this field, the researcher made a limit in 3 actions in every category), and in 1 action would consists of 5 to 6 sentences.

In the second step, there are 18 accumulated stories divided into three parts. Each part contains six human expressions: happy, sad, angry, scary, disgust, and shocked. The therapist suggested a choice of word in every sentence the researcher made. Then, in the last step in developing the adapted book, the researcher formulated a flow to make a story for Autistic Spectrum Disorder (ASD). Here are the lists:

- Use time sequence story method
- Every story have to explain detailly
- Every sentence use 2 to 5 words only
- Give an illustration picture in every sentence
- The sentences must relate to the illustration picture
- Illustration picture must be in sequence
- Instead of abstract words, the explanation uses a concrete word. Take, for example, a dark sky, and there is a dark sky on the illustration.
- Avoid using the word like "want" and "will."
- Show familiar story for children

- Use variable character, such as curly hair, straight hair, bright-dark skin, to show there are many ethnics and race (SARA) in Indonesia.

After arranging the sentences, an illustration made from each of the sentences and dividing a page of the Augmented Reality adapted the book into three layout sections, as shown in Fig. 3.

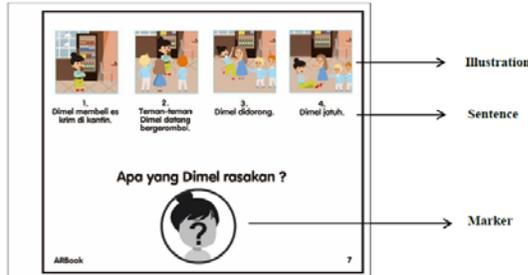


Fig. 3. Three-Section Layout on AR Adapted Book.

2) *Illustration design*: The adapted book illustration was made sequentially according to the formula list in Section III (A). Each page on the adapted book is divided into three layout sections, i.e., illustrative layout, sentence or narrative layout, marker layout. The illustrative layout contains pictures that visually explain the sentences beneath it on the narrative layout and marker layout allocated for the camera to detect the marker on that area.

3) *Marker ARBook design*: There are 18 different markers on each page designed on an Augmented Reality adapted book (ARBook). Every marker was made to suit facial expression on the adapted book pages. The marker can show the multiple-choice answer to the problem story in the adapted book.

B. Design of Augmented Reality Object on ARAnimasi

1) *ARAnimasi's character design*: A digital sketch was used in the process of designing and editing the animation character. The character was made and grouped into two categories, male character, and female character. When the animation showed, the upper body of each character will pop out. The illustration described in Fig. 4. While designing the sketch, it focused on facial features such as the shape of the eyes, eyebrow, and mouth which is described in Fig. 5. The shape of the facial feature is used to differentiate facial emotional character on each animation sketch.



Fig. 4. Sketch of Head, Upper Body, and Hand of Animation Character.



Fig. 5. Various Shape of Mouth, Eyes, and Eyebrow of Animation Character.

2) *ARAnimasi's marker design*: On the ARAnimasi menu, there is 12 flashcard-shaped marker choice with emotion animation object on augmented Reality. The ARAnimasi marker consists of two gender groups, male and female. Each group contains six emotional expression animations (angry, sad, happy, afraid, nauseous, and surprised). The design of the ARAnimasi's marker can be seen in Fig. 6.



Fig. 6. ARAnimasi Markers.

IV. SYSTEM IMPLEMENTATION

A. Tools and Technology

The development of the AREmotion application using Unity and Vuforia. Vuforia was chosen because this platform provides Augmented Reality SDK for image recognition ability to recognize the marker [19]. Every marker in the AREmotion application is saved in the Vuforia account. The engine that is used in the AREmotion application, Unity3D, all of the process is in PC. The result of the Unity3D engine is in the form of a smartphone application with Android systems.

Every animation and character in the ARAnimasi menu was drawn in Adobe Photoshop and Crazy Talk Animator. Wavepad was used for recording and balancing audio. The Adobe Premiere was used for the final rendering of the animation video. AREmotion application was used in the smartphone. The mobile device specification used was 2 GB RAM, 16GB internal memory, 16 MP camera to detect the markers, and the operating system Android 6.0 (Marshmallow).

B. System Overview

AREmotion is a combine application of adapted book and mobile Augmented Reality, which runs in smartphone with platform Android. AREmotion divided into three menus. They are video, ARAnimasi, and ARBook. The first menu is "video", consists of video footages with kinds of expression gesture like sad, happy, angry, afraid, disgust and shock. The second menu is "ARAnimasi", consists of a marker flashcard with six child expressions (for girls and boys) in AR is shown in Fig. 8(left). The third menu is "ARBook", which consists of questions in the form of a story/storybook, with the answers in AR is shown in. 8(right).



Fig. 7. Augmented Reality Adapted Book.

The AR adopted book's physical form is printed in the size 21cm x 16cm, as shown in Fig. 7. There are two menus in the book that is ARAnimasi and ARbook.



Fig. 8. Flashcard of ARAnimasi (Left) and Marker Sample of ARBook (Right).

C. Features Overview

In this study, there are two designs Augmented Reality, ARAnimasi dan ARBook. Fig. 8 shows an example of the ARAnimasi. Fig. 9 shows a menu containing 5 seconds animation video supported by clean audio to express human emotion. For example, there is the sound of someone's crying in the video. It is to show the children the expression of sadness.



Fig. 9. Marker Sample of ARAnimasi Menu.

Fig. 10 shows an example of the ARBook menu design when the smartphone camera is ready to detect the marker. While Fig. 11 shows, there is a multiple-choice answer organized as a UI AR button with slight animation from the detected marker. Animation of the answer pops out for 3 seconds, and it played automatically when the smartphone camera detects the marker. Fig. 12 shows the design of augmented Reality that pops out when the user chooses the correct answer and wrong answer.



Fig. 10. Example of ARBook Menu Design.



Fig. 11. Multiple Choice Interface.



Fig. 12. The Interface of Right Answer (Left) and Wrong Answer (Right).

V. RESULT AND ANALYSIS

The usability of AREmotion is tested on three therapists as the respondent. The therapist tested the usability by experiencing the AREmotion end fill out the questionnaire. Usability is a method that aims to measure whether the application media that has been built is in accordance with user needs. Usability testing can verify whether the applied design has worked well because there is an interaction how the user can cognitively assess the questions on the questionnaire, as the implementation of the interaction with the designed application [20].

There are 17 statements related to five categories of usability in the questionnaire, i.e., efficiency, learnability, errors, memorability, and satisfaction [21]. Each statement in the questionnaire is evaluated by the respondent using the Likert Scale. The respondent gives marks or rates from 1 to 5 to each statement [22]. Table II shows the Likert Scale which describes the symmetrical scale assessment of the respondent's level of agreement or disagreement [20], and Table III shows the questionnaire and all respondents's rates.

TABLE II. LIKERT SCALE DESCRIPTION

Score	Description
1	Very Disagree
2	Disagree
3	Neutral
4	Agree
5	Very Agree

TABLE III. USABILITY QUESTIONNAIRE STATEMENT AND RESULT

No	Statements	Score					Sum
		1	2	3	4	5	
Learnability							
1.	I think children with ASD can easily recognize the menu on AREmotion (Augmented Reality Adapted Book).	0	0	0	2	1	13
2.	From my point of view, AREmotion (Augmented Reality Adapted Book) is simple and easy to learn	0	0	0	3	0	12
3.	In my opinion, text on AREmotion (Augmented Reality Adapted Book) can be read easily	0	0	0	0	3	15
4.	It seems that the graphic display on AREmotion (Augmented Reality Adapted Book) screen is familiar to children with ASD.	0	0	0	2	1	13
5.	I think the icon on the AREmotion (Augmented Reality Adapted Book) media cognizable by ASD children.	0	0	0	3	0	12
Efficiency							
6.	In my opinion, AREmotion (Augmented Reality Adapted Book) can introduce human basic emotional expression to ASD children.	0	0	0	1	2	14
7.	I think the media has simple and easy use features.	0	0	0	3	0	12
8.	It seems that AREmotion (Augmented Reality Adapted Book) does not run at a slow speed.	0	0	3	0	0	9
Memorability							
9.	From my perspective, the animation object on the media is easily recalled by children with ASD.	0	0	0	3	0	12
10.	I think children with ASD can remember how to use AREmotion (Augmented Reality Adapted Book) media.	0	0	0	3	0	12
11.	In my opinion, ASD children recalling the variety of human facial emotional expressions more accurately when they learn it through AREmotion (Augmented Reality Adapted Book).	0	0	0	1	2	14
Error							
12.	In my opinion, children with ASD will not be making many mistakes when they use AREmotion (Augmented Reality Adapted Book)	0	0	2	1	0	10
13.	I think when children make some mistake in using the application, they will easily correct it.	0	0	2	1	0	10
Satisfaction							
14.	In my opinion, children with ASD use AREmotion (Augmented Reality Adapted Book) media comfortably.	0	0	1	1	1	12
15.	From my perspective, the animation on AREmotion (Augmented Reality Adapted Book) is running well.	0	0	0	2	1	13
16.	It seems to me that audio from AREmotion (Augmented Reality Adapted Book) heard by ASD children	0	0	1	2	0	11
17.	In my opinion, AREmotion (Augmented Reality Adapted Book) has an interesting design	0	0	0	3	0	12

Based on the result in Table III, the sum of each category will be calculated. Table IV and Table V show the calculation process.

TABLE IV. THE TOTTALLING SORE FOR EACH CATEGORY

No	Category	1	2	3	4	5	Sum
1.	Learnability	0 x 1	0 x 2	0 x 3	10 x 4	5 x 5	65
2.	Efficiency	0 x 1	0 x 2	3 x 3	4 x 4	2 x 5	35
3.	Memorability	0 x 1	0 x 2	0 x 3	7 x 4	2 x 5	38
4.	Errors	0 x 1	0 x 2	4 x 3	2 x 4	0 x 5	20
5.	Satisfaction	0 x 1	0 x 2	2 x 3	8 x 4	2 x 5	48
SUM		0 x 1	0 x 2	9 x 3	31 x 4	11 x 5	206

TABLE V. PERCENTAGE OF USABILITY SCORE

No	Category	Sum x Likert Score	Maximum Score	Result
1.	Learnability	65	75	$(65/75) \times 100\% = 86.7\%$
2.	Efficiency	35	45	$(35/45) \times 100\% = 77.8\%$
3.	Memorability	38	45	$(38/45) \times 100\% = 84.4\%$
4.	Errors	20	24	$(20/24) \times 100\% = 83.3\%$
5.	Satisfaction	48	60	$(48/60) \times 100\% = 80\%$
SUM		206	249	$(206/249) \times 100\% = 82.73\%$

AREmotion (AR Adapted Book) has a usability score in a total of 82.73 %.

VI. CONCLUSION

This paper focused on analyzing the design and developing the application augmented Reality adapted book (AREmotion) for children with Autism Spectrum Disorder (ASD). The primary participant in this research is therapist and ASD children. The therapist was involved in designing the application using the User-Centered Design (UCD) approach.

The adapted book's story design was based on a simple consecutively event concept explained in 2-5 sentences and illustration embedded in each sentence. There are three main menus on AREmotion, i.e., ARAnimasi consists of emotional expression animation video in augmented Reality, video emotion menu consists of a group of human emotional expression footage video, and the last menu is ARBook that consists of story problem series about emotional expression for ASD children includes the answer in augmented Reality.

In this initial research work, the application's design is tested on the therapist to en-sure that this application is efficient, easy, and understandable before children with ASD use it. Usability testing on AREmotion has 82.73%, and the highest score among all categories is learnability, with an 86.7% score. The result shows that the AREmotion application can be used as educational media for children with ASD when learning about facial emotional expression. The next step of this research is evaluating the usage of AREmotion to help children with ASD recognize emotional expression through this innovation.

REFERENCES

- [1] P. M. Niedenthal and M. Brauer, "Social functionality of human emotion," *Annu. Rev. Psychol.*, vol. 63, pp. 259–285, 2012, doi: 10.1146/annurev.psych.121208.131605.
- [2] J. Kätsyri, S. Saalasti, K. Tiippana, L. von Wendt, and M. Sams, "Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome," *Neuropsychologia*, vol. 46, no. 7, pp. 1888–1897, 2008, doi: 10.1016/j.neuropsychologia.2008.01.005.
- [3] M. Eroglu S., Toprak S., Urgan O, MD, Ozge E. Onur, MD, Arzu Denizbasi, MD, Haldun Akoglu, MD, Cigdem Ozpolat, MD, Ebru Akoglu, DSM-IV Diagnostic and Statistical Manual of Mental Disorder, vol. 33. 2012.
- [4] "Autism Spectrum Disorders," June.01.2021. Accessed on: 23 June 2021. [Online]. Available: <https://www.who.int/news-room/factsheets/detail/autism-spectrum-disorders>.
- [5] APA, "Supplement To Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition," *Diagnostic Stat. Man. Ment. Disord.*, no. August, p. 26, 2015.
- [6] A. C. Tucker, "Interpreting and Treating Autism in Javanese Indonesia," 2013.
- [7] Y. E. Riany, M. Cuskelly, and P. Meredith, "Cultural Beliefs about Autism in Indonesia," *Int. J. Disabil. Dev. Educ.*, vol. 63, no. 6, pp. 623–640, 2016, doi: 10.1080/1034912X.2016.1142069.
- [8] C. A. Malchiodi, *Handbook of Art Therapy*. The Guilford Press, 2004.
- [9] M. B. T. Sampurno, Y. S. Prabandari, and M. D. Marianto, "Theoretical Exploration of Art Therapy and Education," *Int. J. Indones. Educ. Teach.*, vol. 4, no. 2, pp. 260–276, 2020, doi: 10.24071/ijiet.2020.040209.
- [10] K. Valencia, C. Rusu, D. Quiñones, and E. Jamet, "The impact of technology on people with autism spectrum disorder: A systematic literature review," *Sensors (Switzerland)*, vol. 19, no. 20, pp. 1–22, 2019, doi: 10.3390/s19204485.
- [11] L. Daniela, *New perspectives on virtual and augmented reality: Finding new ways to teach in a transformed learning environment*. 2020.
- [12] D. N. Eh Phon, M. B. Ali, and N. D. A. Halim, "Collaborative augmented reality in education: A review," *Proc. - 2014 Int. Conf. Teach. Learn. Comput. Eng. LATICE 2014*, pp. 78–83, 2014, doi: 10.1109/LaTiCE.2014.23.
- [13] K. Khowaja et al., "Augmented reality for learning of children and adolescents with autism spectrum disorder (ASD): A systematic review," *IEEE Access*, vol. 8, pp. 78779–78807, 2020, doi: 10.1109/ACCESS.2020.2986608.
- [14] P. Cunha, J. Brandao, J. Vasconcelos, F. Soares, and V. Carvalho, "Augmented Reality for Cognitive and Social Skills Improvement in Children with ASD," no. February, pp. 328–329, 2016.
- [15] C. H. Chen, I. J. Lee, and L. Y. Lin, "Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders," *Res. Dev. Disabil.*, vol. 36, pp. 396–403, 2015, doi: 10.1016/j.ridd.2014.10.015.
- [16] C. H. Chen, I. J. Lee, and L. Y. Lin, "Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions," *Comput. Human Behav.*, vol. 55, pp. 477–485, 2016, doi: 10.1016/j.chb.2015.09.033.
- [17] G. S. Akbar, E. R. Kaburuan, and V. Effendy, "User interface (UI) design of scheduling activity apps for autistic children," *Proc. 2017 Int. Conf. Orange Technol. ICOT 2017*, vol. 2018-January, pp. 129–133, 2018, doi: 10.1109/ICOT.2017.8336105.
- [18] North Yorkshire Country Council - Health and Adult Services, "A guide to producing written information in easy read," no. June, 2018, [Online]. Available: <https://www.nypartnerships.org.uk/sites/default/files/Partnership%20files/Learning%20disabilities/Guide%20to%20easy%20read.pdf>.
- [19] Zeynep Tacgin, *Virtual and Augmented Reality: An Educational Handbook*. Cambridge Scholars Publishing, 2020.
- [20] E. Geisen and J. R. Bergstrom, *Usability Testing for Survey Research*. Todd Green, 2017.
- [21] S. B. C. Kate, *Fundamental of User Centered Design, a Practical Approach*. CRC Press, 2016.
- [22] R. Dermawi, H. Tolle, and I. Aknuranda, "Design and usability evaluation of communication board for deaf people with user-centered design approach," *Int. J. Interact. Mob. Technol.*, vol. 12, no. 2, pp. 197–206, 2018, doi: 10.3991/ijim.v12i2.8100.

Towards Evaluating Adversarial Attacks Robustness in Wireless Communication

Asmaa FTAIMI¹, Tomader MAZRI²
Laboratory of Advanced Systems Engineering
Ibn Tofail Science University
Kenitra, Morocco

Abstract—The emerging new technologies, such as autonomous vehicles, augmented reality, IoT, and other aspects that are revolutionising our world today, have highlighted new requirements that wireless communications must fulfil. Wireless communications are expected to have a high optimisation capability, efficient detection ability, and prediction flexibility to meet today's cutting-edge telecommunications technologies' challenges and constraints. In this regard, the integration of deep learning models in wireless communications appears to be extremely promising. However, the study of deep learning models has exhibited inherent vulnerabilities that attackers could harness to compromise wireless communication systems. The examination of these vulnerabilities and the evaluation of the attacks leveraging them remains essential. Therefore, this paper's main objective is to address the alignment of security studies of deep learning models with wireless communications' specific requirements, thereby proposing a pattern for assessing adversarial attacks targeting deep learning models embedded in wireless communications.

Keywords—Adversarial attacks; deep learning; wireless communication; security; robustness; vulnerability; threat

I. INTRODUCTION

Wireless communication represents an exciting and evolving field of study. It has significantly contributed to the development of telecommunications and had been at the source of LTE and 5G network implementations, and it continues today to lead advancement in 6G generation development. Nevertheless, this field of study has recently encountered several challenges to meet emerging telecom technologies' requirements. The necessity of optimisation and adaptability is crucial today to guarantee highly efficient wireless communications [1]. In this context, the researchers have immediately turned to the rapidly growing techniques of artificial intelligence, especially the deep learning models that have received considerable interest for their reliability in computer vision and object detection.

Deep Learning models have revolutionised many fields of study. Their expressivity and generalisation potentials have shown impressive outcomes in several areas, including wireless communications [2]. The latter have utilised deep learning models in the radio frequency spectrum and have harnessed their adaptability and flexibility to improve the wireless communication capacity. An obvious application of Deep Learning in Wireless Communications is spectrum estimation and detection and modulation classification. These two functionalities contribute significantly to enhancing the

transmission quality while handling the channel effects encountered at the receiver [3,4].

However, several research studies have revealed that Deep Learning models contain inherent vulnerabilities that an attacker could eventually harness to perform malicious actions compromising wireless communications systems. Nevertheless, researchers usually opt for different hypotheses and follow different methods when testing adversarial attacks, rendering confronting, and comparing the latter challenging and complicated since the platform used in literature while testing adversarial attacks are not standardised. Whereas some researchers hypothesise a scenario where the attacker possesses complete knowledge of the system, known as white-box attacks, others deal with attacks where the attacker's knowledge is constrained, known as grey-box attacks. This disparity in assumptions provides a complex platform to make an accurate comparison of attacks on various algorithms.

Since no unified scheme has been developed for assessing attacks robustness, we have proposed a framework to analyse and evaluate the robustness of adversarial attacks in the context of wireless communication [27-30]. This paper aims to provide a standardised and unified platform for comparing different adversarial attack strategies against wireless communication systems. Through our study of adversarial attacks in literature, we have derived a list of criteria that we have considered to evaluate the complexity of the attack and its impact on the target system to obtain a global view of its robustness.

In this paper, we will extensively review the work devoted to studying the particularities of adversarial attacks targeting wireless communications to propose a pattern designed to evaluate the robustness of these attacks while incorporating the specifications related to this field of study.

First, we will introduce the applications of deep learning models in wireless communications. Then we will discuss the theoretical aspect of these attacks as well as the different factors involved in their identification and classification. Afterwards, we provide a review of the work highlighting certain particularities of wireless communications that could impact adversary attack success. Finally, we will elaborate on a pattern being proposed to evaluate the adversarial attacks' robustness in the context of wireless communications.

II. RELATED WORKS

A new research field has been conducted to extensively study the security of deep learning models and adversarial attacks leveraging their flaws [5]. Adversarial learning addresses the empirical evaluation of adversarial attacks by testing them in the physical world to experience them in a realistic scenario [6]. This research field is similarly focused on studying the theoretical aspect of these attacks by proposing a taxonomy for their classification and a threat model to describe the different aspects of the attack. However, the multiple suggestions carried out towards assessing adversarial attacks have mainly focused on attacks targeting computer vision or object detection models. Few works have addressed adversarial learning in the wireless communications context [27-30]. Indeed, the latter presents certain specificities to be considered in the study of attacks targeting the models they employ. Hence, it is necessary to adapt threat models and methods for evaluating adversarial attacks' robustness to wireless communication's technical and functional specifications.

III. DEEP LEARNING APPLICATION IN WIRELESS COMMUNICATIONS

Wireless communications represent an essential field of study for developing networks and meeting innovative technologies' specific requirements. The evolution of telecommunications stems from this research field since it has brought LTE and 5G networks to the surface. It is also contributing significantly today to waveform design for 6G emerging networks. Wireless communications have been based on classical probabilistic and analytical methods. However, such an approach involves several limitations regarding channel modelling, interference handling, traffic management, error detection and correction, and security [7].

Wireless communications have evolved systems built on deep learning models to overcome the complexities encountered in earlier network generations. Wireless Communications leverage the expressiveness and capacity for generalisation of Deep Learning models toward addressing detection, classification, optimisation, and prediction problems and consequently guarantee quick, reliable, and secure communications.

A. Deep Learning for Communication Systems

The primary purpose of wireless communication is to ensure a message's reception in an optimal state by deploying resources efficiently. The transmission mechanisms deployed by wireless communications are handled through independent blocks, each dedicated to the specific functionality of the data transmission process, as shown in "Fig. 1". Conventional approaches have focused on enhancing each block's functionality separately, thereby failing to achieve a proper optimisation of the overall system. Nevertheless, deep learning models flexibility currently provides the ability to address the optimisation needed for different blocks in parallel [8].

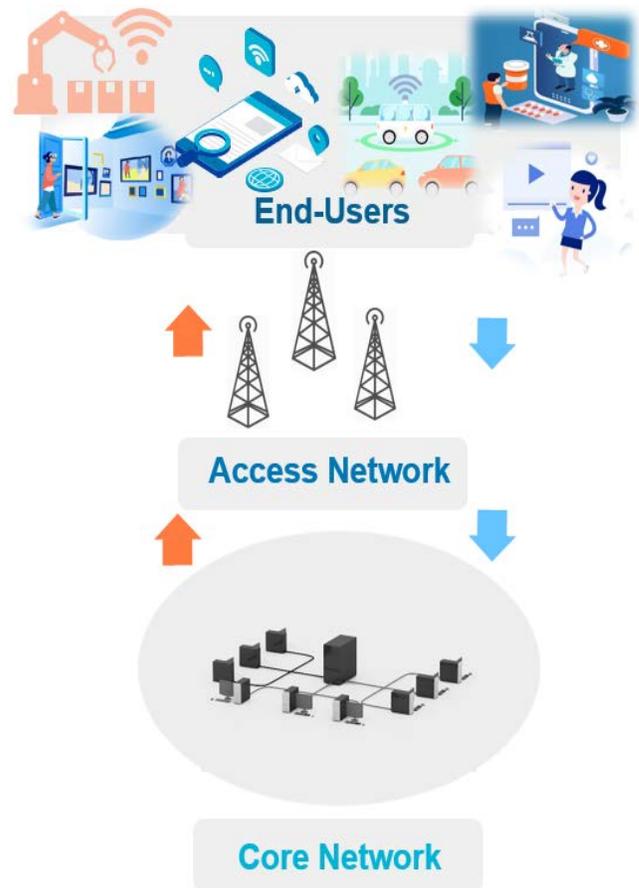


Fig. 1. Wireless Communication Architecture.

Moreover, Deep Learning in the end-to-end communication process has been employed significantly in implementing MIMO techniques. This technology involves integrating multiple antennas during the transmission and reception of signals to boost the spectrum's performance. Moreover, the implementation of multiple antennas appears to be computationally expensive and challenging in system optimisation. Thus, the use of Deep Learning in innovative studies [9,10] regarding the MIMO technique has overcome these challenges. The application of deep learning models has proven to be important when managing multiuser communication systems as well. These techniques could be applied to optimise the spectrum's exploitation for multiple users, yet they remained restrained by channel interferences.

Accordingly, an emerging technique known as Non-Orthogonal Multiple access NOMA has contributed to solving this trade-off by improving spectrum efficiency while minimising interference [11]. Currently, a new approach is also being considered for addressing these three issues in a synergistic scheme. It involves dealing with both emitter and receiver as one system designed as a single autoencoder [12], requiring a comprehensive optimisation.

B. Deep Learning for Spectrum Estimation, Detection, and Classification

Wireless communications have employed Deep Learning models' powerful capabilities in adaptability and flexibility to arm their systems with cognition. Cognitive radio could learn from collected data to adjust dynamically and rapidly the spectrum to performance and throughput demands. Among the crucial tasks that cognitive radios need to carry out, we cite signal detection and classification, although it is complex for classical feature-based algorithms that lack the flexibility to accommodate different types of signals. Deep learning model-based systems can overcome such problems since they have a high generalisation ability to classify and detect several types of signals, as shown in "Fig. 2".

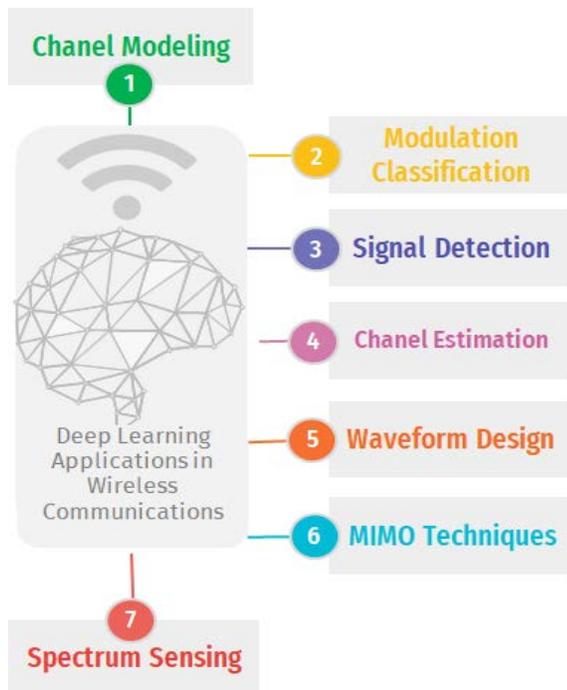


Fig. 2. Deep Learning Applications in Wireless Communications Systems.

1) *Channel modeling and estimation:* Channel modelling becomes essential to enhance the communication system's performance. For instance, autoencoders require a training phase approaching as closely as possible the real channel conditions. However, because of the channel effect, performing an efficient autoencoder training enfold several complexities. Hence reliance on GAN [13] to approximate interferences, noise, and multi-path effects to depict a channel model representing accurate and realistic behaviour [14].

2) *Signal detection:* Signal detection and classification is crucial functionality in wireless communications. It allows the control of system components and provides an up-to-date overview of the communications and events occurring in the system. Indeed, it ensures reliable detection of spectrum users and arising events, such as identifying interference sources for immediate response. Nevertheless, the spectrum is often shared for multiple simultaneous applications (TV, GSM, LTE, Radar, Etc.), which is challenging when identifying the wide variety

of waveforms used. The conventional general and specialised detection methods lack scalability and depend on the SNR for signal detection and classification. Therefore, detection using these methods is difficult to perform, especially when the SNR these to noise ratio is low [15].

3) *Consequently,* many studies consider employing CNN models that have proven their high performance in object detection and recognition, specifically in computer vision. O'Shea et al. [16] have examined the application of CNN models for signal detection and classification in the radio frequency spectrum. Their study utilised Gradient-Weighted Class Activation Mapping (Grad-Cam) for spectral event localisation and have achieved high performing results.

4) *Modulation classification:* O'Shea et al. [17] have evaluated the performance of CNN models in modulation classification by experimenting with channel effects such as multi-path fading to test the accuracy rate obtained under real-world conditions. Following the study carried out on a dataset of 11 types of modulations often used in wireless communication, the results obtained by CNN models vastly exceeded those produced by SVM or Naive Bayes, even for all used SNR ratios.

Although deep learning delivers considerable potential advantages, most recent studies have shown that they contain many vulnerabilities that attackers can harness to perform malicious manipulations [18]. Many researchers have recently examined the security of deep learning models. Some have considered the practical aspect by testing these attacks in the physical world, particularly in computer vision and object detection, while others have focused on studying the theoretical aspect of adversarial attacks. In the following section, we will present the taxonomy of these attacks and the different classifications proposed for their study.

IV. THREAT MODEL

Recently, several research studies have focused on scrutinising the security of Deep learning models. Experiments conducted in a variety of fields of study have shown that these models are vulnerable. Their high flexibility potential and their adaptability have contributed considerably to their weakness [19]. Today, several types of attacks have been tested to reveal the security flaws of deep learning models. Indeed, an attacker potentially poisons the model by introducing malicious data during the training phase to affect its behaviour to the new input data. The attacker could also severely compromise the model even in the prediction phase through carefully designed inputs to exploit its inherent flaws to mislead it into producing inaccurate results [20]. All these considerations have motivated researchers to explore deep learning models' security using conventional security approaches, especially the study of their confidentiality, integrity, and availability (CIA). In this regard, a group of researchers, namely Barreno et al. [21], have developed a taxonomy dedicated to the security of these models, providing a comprehensive classification of adversarial attacks by highlighting the opponent's goals, his knowledge of the targeted system, his capabilities and the strategy he may employ to carry out the attack as illustrated in "Fig. 3".

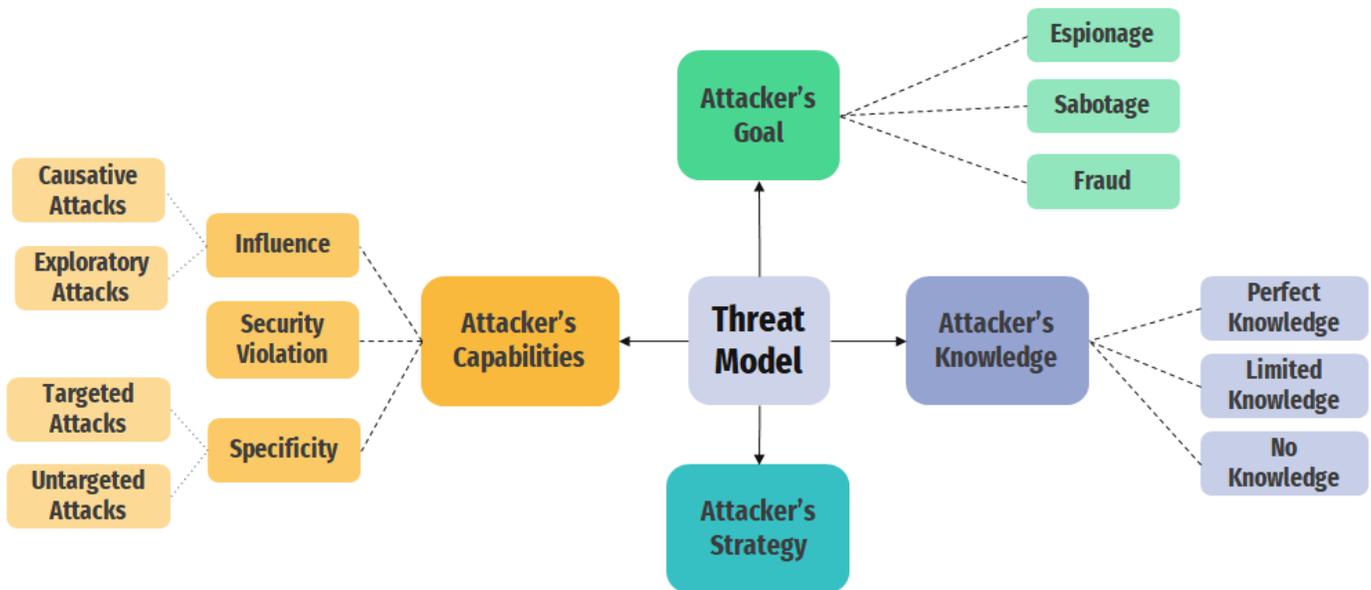


Fig. 3. Adversarial Learning Threat Model.

A. Attacker's Goal

While the same motivations might not necessarily guide attackers, their aims tend to converge around three main axes: espionage, sabotage, or fraud [22].

1) *Espionage*: In this context, the attacker seeks to derive sensitive information by exploiting vulnerabilities in the system. The leakage of sensitive information can compromise the confidentiality and the privacy of the system since the received information can be utilised to plan more advanced attacks and consequently cause very severe incidents.

2) *Sabotage*: The attacker can obstruct the system by either disabling important functionalities or by denying normal operations. Usually, this occurs when the adversary attempts to flood the model with incorrectly classified examples to increase the working time on false positive, or he can just as easily overload the system with a massive number of requests that require more computation time.

3) *Fraud*: Fraud in the Deep Learning system refers to the adversary's action of causing misclassifications or inaccurate predictions. In this case, the attacker takes advantage of existing vulnerabilities in the system to inject malicious inputs in the dataset or even modify the model's behaviour, and in this way, he causes severe damages.

B. Attacker's Knowledge

Regardless of the adversary's goals, the complexity of the attack he intends to carry out depends significantly on his knowledge of the targeted system. Papernot et al. [23] have elaborated a classification of the attacker's knowledge that can be represented in the following three levels:

1) *Perfect knowledge level*: in this scenario, the attacker knows everything about the model and the training dataset and can carry out white-box attacks. Nevertheless, this scenario

remains unrealistic because it is almost impossible to have perfect knowledge about the target system.

2) *Limited knowledge level*: This scenario is more realistic and practical; nevertheless, it presents a range of possibilities: 1) Limited Knowledge attacks with surrogate model: when the attacker has limited knowledge of the model, he can use an alternative model with features similar to the targeted system's model's features in order to craft effective attacks, 2) Limited knowledge attacks with the surrogate dataset: when the adversary has no access to the training data he can use substitute dataset with similar characteristics to carry out efficiently the attacks.

3) *Oracle or no knowledge*: In this case, the attacker has no prior information about the training set, the model, or its features; he can perform black-box attacks.

C. Attacker's Capabilities

Besides factors seen in the previous sections, the opponent's means and potentials contribute largely to determining the attack's success [24]. The adversary's potentials could be categorised according to the following three main axes:

1) *Influence*: The attacker can compromise the targeted system using either a causative attack to introduce malicious data into the algorithm's training set or by exploiting the model's weaknesses by introducing specific inputs, often called adversary examples. The causative attack influences the model's behaviour contrary to the exploratory attack, where the adversary does not affect the model's behaviour [25].

2) *Security violation*: The security infraction committed by the adversary relies on the actions taken to compromise the targeted system. The attacker can cause integrity violations by crafting false-negative inputs that bypass the model without altering the usual tasks. However, the adversary causes an availability violation when he conducts false positives, leading

to a denial of service. Moreover, the adversary can also perform privacy violation attacks that aim to derive sensitive information about the users of the targeted system, the training dataset, or the features of the model [26].

3) *Specificity*: This characteristic determines how much the adversary is specific while performing the attack. Indeed, when the attacker has intentions to mislead the model for specific instances, he can perform targeted attacks, whereas, if he chooses to compromise the predictions or the classifications carried out by the model for a broad range of inputs, he must implement indiscriminate attacks [22].

D. Attacker's Strategy

An attack strategy can be elaborated by leveraging the model's vulnerabilities and flaws by considering the attacker's goals, knowledge of the targeted system, capabilities, and potentials. Therefore, the attack strategy in question is nothing more than an optimisation problem aiming to minimise the model performance by carefully crafting efficient and imperceptible perturbations to achieve its malicious objectives successfully [21].

V. EVALUATING ADVERSARIAL ATTACKS IN WIRELESS COMMUNICATION

The study of adversarial learning in Wireless Communications presents new aspects beyond the scope of the examination of attacks targeting computer vision. Indeed, the adversarial attacks must be meticulously studied in this context. In the following paragraphs, we will examine the types of attacks targeting wireless communication systems. Then we will highlight important metrics that must be considered while assessing their robustness [27].

A. Type of Adversarial Attacks

Adversarial attacks in Wireless communications could be classified into two categories:

1) *Direct Access Attacks (DAC)*: This category of attacks exploits the direct access to the classifier's input dataset to carry out malicious actions. The results obtained in [27] for this type of attack have shown that for symbol energy and jamming signal ratio E_s/E_j of 30db, the FGSM attack produces a higher degradation than the one caused by the addition of AWGN Gaussian noise.

2) *Over the Air Attacks (OTA)*: In computer vision, the attack reliability depends on selecting the perturbations that lead the model into misclassifications while remaining imperceptible to human eyes. Similarly to computer vision, Self-protect attacks are also interested in misleading the classifier while guaranteeing information transmission to the receiver with a defined modulation.

B. Metrics to Evaluate Adversarial Attacks in Wireless Communications

Besides, the evaluation of attacks and their success rates in Wireless Communications must adopt additional metrics aligned with signal transmission performance measures. Therefore, it is essential to consider the Bit Error Rate (BER) computation and the ratio of perturbing noise and modulated

signal to estimate the opponent's attack's success rate, among other metrics.

1) *Frequency offset*: Before classifying the wideband signal, the systems initially identify the frequency of the signals and the time of transmission to convert these signals back to the baseband. Nevertheless, such operations may induce errors, as shown in [13], especially in centre frequency estimation, leading to frequency offset. The authors in [13] have noticed that raw-IQ-based AMC model accuracy dramatically decreases after adding frequency offsets. This encouraged Flowers et al. to examine frequency offsets' impact on adversarial attacks' success rate. The obtained results for 10 and 20 dB SNRs showed that even the most minor errors in frequency offset estimation could reduce the effect of adversary examples by increasing the model's accuracy by approximately 10%.

2) *Time offset*: To estimate transmission start and end times, the system employs an energy detection pattern that utilises a specific threshold of frequency power to determine the signal's existence in each instant. Incorrect evaluation of this threshold can result in false alarms or delays in the estimated transmission start time. In [28], the authors studied the impact of these parameters on the model accuracy rate. Indeed, in the absence of adversarial examples, the time offset does not significantly affect the model's accuracy rate. However, under adversarial conditions, they have noticed that translating the time-offset by four samples enhances the model accuracy rate by 20% for an E_s/E_j ratio of 12 dB. Thus, they have concluded that the time offset can considerably reduce adversarial perturbations' impact on modulation classifier.

3) *Multiple antenna usage*: In [29], the authors have examined a wireless communication system in which the transmitter emits signals to receivers using several different modulation types. The receiver identifies the modulation types using a deep learning model classifier. In this context, an opponent can potentially introduce adversarial noise by employing multiple antennas to mislead the classifier and decrease its accuracy. They have demonstrated that using multiple antennas could enhance the opponent's attack robustness using a technique used in previous work [30] known as the maximum received perturbation power MRPP. They have evaluated this attack by emulating two different scenarios. In the first one, they have attacked while using adversaries operating in separate locations with only a single antenna. In the second scenario, they have performed the Elementwise Maximum Chanel Gain EMCG attack involving a single opponent yet with multiple antennas. These two scenarios have applied different techniques for power allocation. Kim et al. [29] also considered the opponent's attack on modulation classifiers while maintaining two essential requirements: 1) the perturbations introduced to the signal must be conceived to drive the targeted model to misclassify the modulations; 2) the power of the perturbations must not exceed maximum permissible levels so that they remain imperceptible to the receiver. The experimental findings indicated that the use of

multiple opponents would not degrade the classifier's performance significantly as a single adversary had employed the same antenna power [31]. The reason behind the obtained results is the lack of coordination and collaboration between multiple opponents in the second case since they do not focus on the same adversarial goal. The authors have demonstrated in their study the efficiency of EMCG attack. The EMCG attack associated with Gaussian noise gives the weakest results proving that Gaussian Noise's presence makes the adversarial perturbations detectable by the signal receiver.

VI. THE PROPOSITION OF ASSESSMENT PATTERN OF ADVERSARIAL ATTACKS ROBUSTNESS FOR WIRELESS COMMUNICATION

Considering the results obtained in the study of adversarial attacks in the context of wireless communications, we will propose in this section a pattern to evaluate attack robustness, as shown in "Fig. 4". This pattern is inspired by our previous

works [32] related to adversarial examples robustness assessment in computer vision. It is more adapted to the specificities of wireless communications.

The process is initiated by evaluating the attacker's knowledge of the targeted system. Attacks requiring complete knowledge of the targeted system are the least robust, while those designed with limited knowledge of the target system are proven to be the most robust. Afterwards, the attacker's potential and capabilities are analysed. Indeed, we have extensively detailed through Kim et al.'s [30] experience the impact of this factor in enhancing the robustness of the opponent's attack. In the previous section, we have already demonstrated that the adversary can improve the attack robustness by maximising the used antenna's power or increasing the number of antennas employed to carry out the adversarial attack.

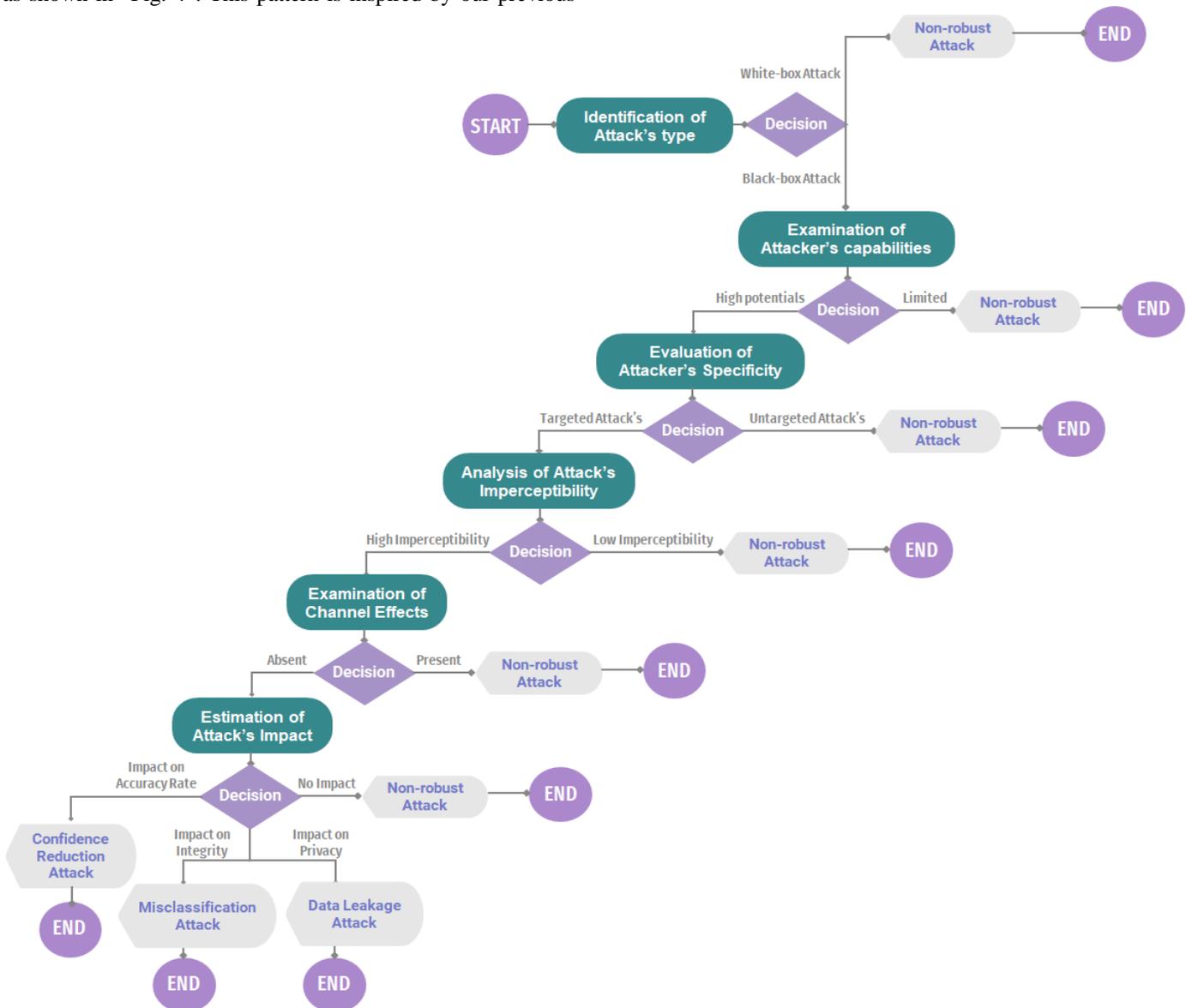


Fig. 4. Adversarial Attacks Robustness Evaluation Pattern.

Subsequently, the focus is on investigating the specificity of the attack. Attacks aiming at specific targets appear to be more robust than those seeking to introduce errors into the model without focusing on a specific target. Further, we analyse the imperceptibility of the perturbations introduced into the targeted system. Indeed, this parameter remains highly relevant to the attack's success rate. Further, we analyse the imperceptibility of the perturbations introduced into the targeted system. Indeed, this parameter remains highly relevant to the attack's success rate. It has been demonstrated in the studies conducted by [33] that as long as the crafted perturbations are perceptible, it is highly probable that they will be detected. This could be achieved in the context of wireless communications by keeping the power of the perturbations below the maximum permissible thresholds and thereby fooling the model without disrupting the signal transmission to the receiver.

In addition, the proposed pattern also involves examining the channel effects of Additive White Gaussian Noise, sample time offsets, and centre frequency offsets on the receiver. As in [34], researchers have shown that adding AWGN would significantly impact the adversarial examples compared to the model's accurate input data. In the presence of AWGN, the success rate of the adversarial attack is significantly reduced due to the identified sensitivity of adversarial examples to additive white Gaussian noise. Therefore, it is essential to include channel effects in the assessment process of adversarial attacks to match conditions that reflect a realistic scenario perfectly.

Finally, the pattern concludes with an estimation step of the attack's impact. To this end, we have proposed an analysis, including the adversary's objective and damaged components in the targeted system. Indeed, if the attacker fails to carry out malicious actions, we consider the attack with no impact on the target system. However, if it allows unauthorised access to the system by the adversary, then the attack has a significant impact on the confidentiality of the target system, yet its impact remains limited. Nevertheless, when the adversary carefully designs malicious perturbations to reduce model confidence in its predictions by diminishing its accuracy rate, then the impact is higher than the previous case. On the other hand, if crafted perturbations alter the model's output or influence its behaviour regarding the input dataset, then the impact is considered high since the attack affects target system integrity [35].

VII. DISCUSSION

The approach we proposed in the previous section is designed to guarantee several advantages, including evaluating adversarial attack robustness according to several metrics, such as the type of attack, its specificity, its imperceptibility, and its impact. Indeed, we have considered analysing the adversary's knowledge and capabilities since they directly affect the attack's success rate. As far as the attacker is familiar with the components of the system, he can perform attacks causing tremendous damage. This feature allows the attacker to design imperceptible adversarial examples to carry out the attack and increase its specificity by targeting precise targets. In addition, We have included the attack's adaptability as an essential

criterion for estimating its robustness. Indeed, the latter considerably affects the adopted attack strategy flexibility and therefore increases the challenge of the adopted defensive mechanism to mitigate the vulnerabilities of the targeted system. Moreover, we have conducted a comprehensive study of the attack's impact through an in-depth examination of the opponent's goal. Accordingly, we have developed a framework standardising the study and the assessment of different types of adversarial attacks. Our process provides the significant advantages of adaptability and generality since it can be applied to different models and can be tailored to different attacks and strategies that attackers may adopt in wireless communication systems.

VIII. CONCLUSION

Recently, growing attention in the scientific community has been dedicated to adversarial attacks. Throughout their studies, researchers have adopted several methods and established different hypotheses. This has made evaluating these attacks challenging since the platforms used in the experimentations are not standardised. Therefore, the proposed work suggests a unified method for evaluating adversarial attacks targeting deep learning models in wireless communications. This work has highlighted essential security aspects of the deep learning model used in wireless communications. In the different sections, we have explored the theory behind adversarial attacks as well as their practical application in the physical world. We have also examined the different studies carried out in this direction to draw a set of characteristics specific to wireless communications that greatly influence the success rate of the opponent's attack. At the end of this article, we have proposed a pattern devoted to evaluating adversarial attacks' robustness. Finally, the proposed model is designed to highlight the different characteristics of the attacks to provide an exhaustive evaluation that approximates the scenarios that can be encountered. Our future work will focus on implementing this framework using various Deep Learning models and different attacks to test its reliability in assessing adversarial attacks robustness in wireless communication systems.

REFERENCES

- [1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, 23(2), 201–220, 2005, doi:10.1109/JSAC.2004.839380.
- [2] C. Clancy, J. Hecker, E. Stuntebeck, T. O'Shea, "Applications of Machine Learning to Cognitive Radio Networks," *IEEE Wireless Communications*, 14(4), 47–52, 2007, doi:10.1109/MWC.2007.4300983.
- [3] S.M. Dudley, W.C. Headley, M. Lichtman, E.Y. Imana, X. Ma, M. Abdelbar, A. Padaki, A. Ullah, M.M. Sohil, T. Yang, J.H. Reed, "Practical Issues for Spectrum Management With Cognitive Radios," *Proceedings of the IEEE*, 102(3), 242–264, 2014, doi:10.1109/JPROC.2014.2298437.
- [4] L.J. Wong, W.C. Headley, S. Andrews, R.M. Gerdes, A.J. Michaels, "Clustering Learned CNN Features from Raw I/Q Data for Emitter Identification," in *MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM)*, IEEE, Los Angeles, CA: 26–33, 2018, doi:10.1109/MILCOM.2018.8599847.
- [5] D. Tsipras, S. Santurkar, L. Engstrom, A. Turner, A. Madry, "Robustness May Be at Odds with Accuracy," *ArXiv:1805.12152 [Cs, Stat]*, 2019.

- [6] S. Qiu, Q. Liu, S. Zhou, C. Wu, "Review of Artificial Intelligence Adversarial Attack and Defense Technologies," *Applied Sciences*, 9(5), 909, 2019, doi:10.3390/app9050909.
- [7] T. Erpek, T.J. O'Shea, Y.E. Sagduyu, Y. Shi, T.C. Clancy, "Deep Learning for Wireless Communications," *ArXiv:2005.06068 [Cs]*, 2020.
- [8] N. Rahimi, J. Maynor, B. Gupta, "Adversarial Machine Learning: Difficulties in Applying Machine Learning Existing Cybersecurity Systems," 8.
- [9] N. Samuel, T. Diskin, A. Wiesel, "Deep MIMO Detection," *ArXiv:1706.01151 [Cs, Math, Stat]*, 2017.
- [10] H. He, C.-K. Wen, S. Jin, G.Y. Li, "A Model-Driven Deep Learning Network for MIMO Detection," *ArXiv:1809.09336 [Cs, Math]*, 2018.
- [11] T. Erpek, S. Ulukus, Y.E. Sagduyu, "Interference Regime Enforcing Rate Maximization for Non-Orthogonal Multiple Access (NOMA)," in *2019 International Conference on Computing, Networking and Communications (ICNC)*, IEEE, Honolulu, HI, USA: 950–994, 2019, doi:10.1109/ICCNC.2019.8685624.
- [12] I.J. Goodfellow, J. Shlens, C. Szegedy, "Explaining and Harnessing Adversarial Examples," *ArXiv:1412.6572 [Cs, Stat]*, 2015.
- [13] T.J. O'Shea, T. Roy, N. West, B.C. Hilburn, "Physical Layer Communications System Design Over-the-Air Using Adversarial Networks," *ArXiv:1803.03145 [Cs, Eess]*, 2018.
- [14] H. Ye, G.Y. Li, B.-H.F. Juang, K. Sivanesan, "Channel Agnostic End-to-End Learning based Communication Systems with Conditional GAN," *ArXiv:1807.00447 [Cs, Math]*, 2018.
- [15] T. O'Shea, T. Roy, T.C. Clancy, "Learning robust general radio signal detection using computer vision methods," in *2017 51st Asilomar Conference on Signals, Systems, and Computers*, IEEE, Pacific Grove, CA, USA: 829–832, 2017, doi:10.1109/ACSSC.2017.8335463.
- [16] T.J. O'Shea, T. Roy, T. Erpek, "Spectral detection and localisation of radio events with learned convolutional neural features," in *2017 25th European Signal Processing Conference (EUSIPCO)*, IEEE, Kos, Greece: 331–335, 2017, doi:10.23919/EUSIPCO.2017.8081223.
- [17] T.J. O'Shea, J. Corgan, T.C. Clancy, "Convolutional Radio Modulation Recognition Networks," *ArXiv:1602.04105 [Cs]*, 2016.
- [18] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, R. Fergus, "Intriguing properties of neural networks," *ArXiv:1312.6199 [Cs]*, 2014.
- [19] M. Jagielski, A. Oprea, B. Biggio, C. Liu, C. Nita-Rotaru, B. Li, "Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning," *ArXiv:1804.00308 [Cs]*, 2018.
- [20] B. Biggio, I. Corona, D. Maiorca, B. Nelson, N. Srndic, P. Laskov, G. Giacinto, F. Roli, "Evasion Attacks against Machine Learning at Test Time," *ArXiv:1708.06131 [Cs]*, 7908, 387–402, 2013, doi:10.1007/978-3-642-40994-3_25.
- [21] M. Barreno, B. Nelson, A.D. Joseph, J.D. Tygar, "The security of machine learning," *Machine Learning*, 81(2), 121–148, 2010, doi:10.1007/s10994-010-5188-5.
- [22] L. Muñoz-González, E.C. Lupu, *The Security of Machine Learning Systems*, Springer International Publishing, Cham: 47–79, 2019, doi:10.1007/978-3-319-98842-9_3.
- [23] N. Papernot, P. McDaniel, A. Sinha, M. Wellman, "Towards the Science of Security and Privacy in Machine Learning," *ArXiv:1611.03814 [Cs]*, 2016.
- [24] Z. Abaid, M.A. Kaafar, S. Jha, "Quantifying the impact of adversarial evasion attacks on machine learning based android malware classifiers," in *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*, IEEE, Cambridge, MA: 1–10, 2017, doi:10.1109/NCA.2017.8171381.
- [25] L. Muñoz-González, E.C. Lupu, *The Security of Machine Learning Systems*, Springer International Publishing, Cham: 47–79, 2019, doi:10.1007/978-3-319-98842-9_3.
- [26] L. Huang, A.D. Joseph, B. Nelson, B.I.P. Rubinstein, J.D. Tygar, "Adversarial Machine Learning," 15.
- [27] B. Flowers, R.M. Buehrer, W.C. Headley, "Evaluating Adversarial Evasion Attacks in the Context of Wireless Communications," *ArXiv:1903.01563 [Cs, Eess, Stat]*, 2019.
- [28] S.-M. Moosavi-Dezfooli, A. Fawzi, O. Fawzi, P. Frossard, "Universal Adversarial Perturbations," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Honolulu, HI: 86–94, 2017, doi:10.1109/CVPR.2017.17.
- [29] B. Kim, Y.E. Sagduyu, T. Erpek, K. Davaslioglu, S. Ulukus, "Adversarial Attacks with Multiple Antennas Against Deep Learning-Based Modulation Classifiers," *ArXiv:2007.16204 [Cs, Eess, Stat]*, 2020.
- [30] B. Kim, Y.E. Sagduyu, K. Davaslioglu, T. Erpek, S. Ulukus, "Over-the-Air Adversarial Attacks on Deep Learning Based Modulation Classifier over Wireless Channels," in *2020 54th Annual Conference on Information Sciences and Systems (CISS)*, IEEE, Princeton, NJ, USA: 1–6, 2020, doi:10.1109/CISS48834.2020.1570617416.
- [31] B. Kim, Y.E. Sagduyu, K. Davaslioglu, T. Erpek, S. Ulukus, "Channel-Aware Adversarial Attacks Against Deep Learning-Based Wireless Signal Classifiers," *ArXiv:2005.05321 [Cs, Eess, Stat]*, 2020.
- [32] A. Ftaimi, T. Mazri, "Evaluation and Analysis of Robustness of Adversarial Examples Attacks in Deep Neural Networks," 6 in press.
- [33] M. Sadeghi, E.G. Larsson, "Adversarial Attacks on Deep-Learning Based Radio Signal Classification," *ArXiv:1808.07713 [Cs, Eess, Math, Stat]*, 2018.
- [34] J. Wang, J. Sun, P. Zhang, X. Wang, "Detecting Adversarial Samples for Deep Neural Networks through Mutation Testing," *ArXiv:1805.05010 [Cs, Stat]*, 2018.
- [35] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, D. Mukhopadhyay, "Adversarial Attacks and Defences: A Survey," *ArXiv:1810.00069 [Cs, Stat]*, 2018.

An Approach for Optimal Feature Selection in Machine Learning using Global Sensitivity Analysis

G.Saranya¹

Research Scholar
Dept. of Computer Science and Engineering
Sathyabama Institute of Science and Technology
Tamil Nadu, India

Dr.A.Pravin²

Dept. of Computer Science and Engineering
Sathyabama Institute of Science and Technology
Tamil Nadu, India

Abstract—The classification application is an important procedure for selecting the feature. The classification is mainly based on the features extracted from the object. You can select the best feature using the following three methods: wrapper selection, filter and embedded procedure. All three practices have been implemented by single or combined two approaches. As a result, there is no important feature in the classification process. This problem is solved by the proposed integrated global analysis of sensitivity. Each feature is selected in a classification based on the sensitivity of the feature and the correlation from the target vector in this integrated sensitivity and correlation approach. Likewise, the GSA approach uses a variety of filtering techniques for ranking attributes and optimization using particle swarm technique. Then, the optimum attributes are trained and tested using the Random Forest Classifier grid search via MATLAB software. In comparison to the existing method, wrapper-based selection, the performance of our integrated model is measured using sensitivity, specificity and accuracy. The experimental results of our proposed approach outweigh the sensitivities by 93.72%, 94.74% and the accuracy of 89.921% and 90% where, wrapper selection approach as sensitivity by 89.83% and the accuracy of 93%.

Keywords—Feature selection; feature sensitivity; feature correlation; global sensitivity analysis; classification

I. INTRODUCTION

In a period of large amounts of large amounts of data, social media, health care, bioinformatics, online training and other media are omnipresent. The machine learning methodology depends greatly on useful basic data from a large data pool. Redundant information degrades learning process performance. It is a critical task to determine useful data and remove redundant information. The process for discovering knowledge is used to gather promising information in large data pools. Pre-processing data steps are sub-processed, such as data cleaning, data inclusion, processing and information reduction. A feature is a quantifiable characteristic of a particular process. Feature selection (FS) is the process of selection from a certain dataset for the most important features. In many cases, FS can further enhance the performance of a learning model [1]. A number of features can be used to classify many machines. Real world data has numerous unimportant, superfluous and noisy features. Deleting these characteristics by FS decreases storing and computer costs while preventing important information loss or trying to learn performance degradation. The general approach to FS is shown in Fig. 1.

FS brings numerous benefits as the system's classification increases predictability, knowledge, usability and broad capacity. It also reduces computer systems' complexity and storage, offers a rapid and effective method for knowledge discovery, and plays a critical role. [2] In literature, the majority of FS methods can be classified as wrappers, filters, embedded or hybrid as illustrated in Fig. 2.

Wrapper methods making wise the quality of the selected features by predicting a predefined, learning algorithm. A traditional method takes two different steps, as shown in Fig. 3, according to a particular learning algorithm. (1) a sub-set of features will be searched and (2) the selected feature will be evaluated. [3] Wrapper repeats (1) and until certain stop criteria have been met (2). The search feature set first generates an inferior set of features, but the algorithms studied are then used as black boxes for quality evaluation. For example, the required number of features or the enhanced learning efficiency is achieved iteratively. When a selected function is returned the sub-set of functions provides maximum performance.

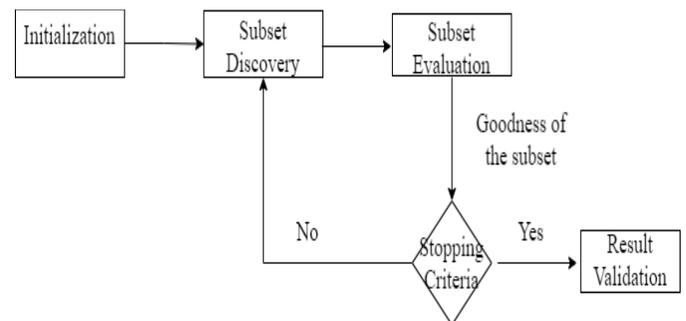


Fig. 1. Basic Feature Selection Process.

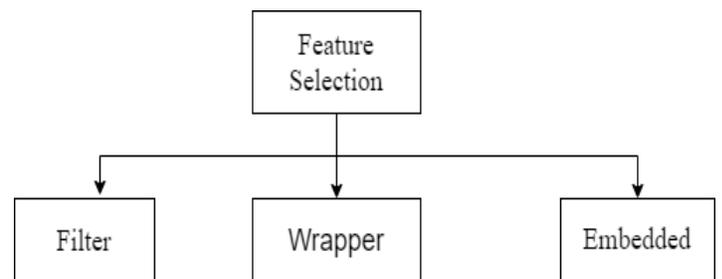


Fig. 2. Classification of FS Approaches.

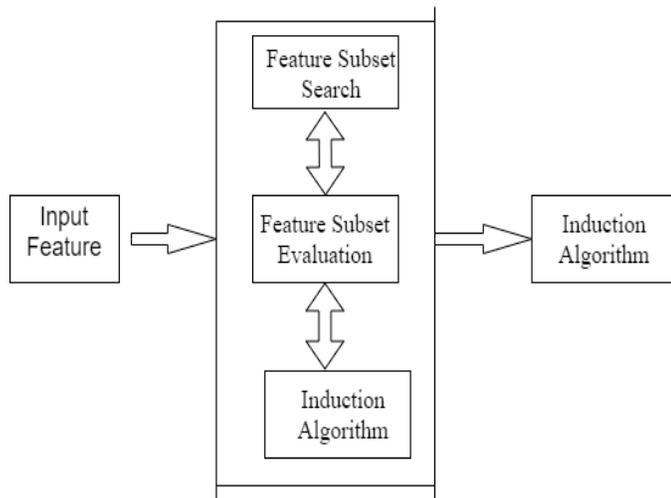


Fig. 3. Wrapper Feature Selection Working Principle.

In the majority of investigators, the selection of supervised features uses the filter evaluation framework [4]. No learning algorithms are available for the filters. You rely on data properties to measure the importance of its features. In general, filter methods are computationally effective rather than wrapping. However, because the feature selection phase is not governed by any certain algorithm, the selected features cannot be optimal for algorithms for the study. There are two steps to a typical filter method. In the first step, the importance of features is classified by certain criteria for feature assessment. The evaluation process may be one-size-fits-all or multivariate. Low quality features are filtered out in the second phase of the typical filter method. Examples are filter methods [5-9]. A typical FS filter technology diagram is shown in Fig. 4.

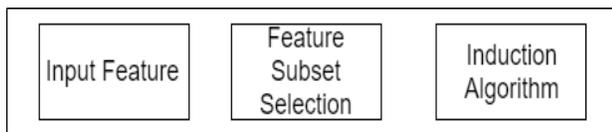


Fig. 4. Filter based Feature Selection Method.

The last type of feature selection is called embedded methods in Fig 4, a compromise that includes the selection of functions for model learning between the filtering and the wrapping methods [10]. Therefore, these approaches are worthy of wrappers and filter (1) because they need to interact with the classification algorithm; and (2) because they do not require an iterative evaluation of the functional sets. The major integrated approaches are the regulation model that can be adapted to the study model by minimizing fitting mistakes and

forcing less coefficients (or exact zero). The official outcome will then be returned, both the regularization model and the chosen functions. In this article we focus on new FS technologies based on a global sensitivity approach.

The paper is organized accordingly: Section II covers the existing related works and its shortcomings. Section III details the work of the proposed methodology. Section IV describes implementation and the results are discussed. Section V summarizes its performance to conclude this document. The future improvement work was suggested in Section VI.

II. RELATED WORKS

A mixture of several methodologies can be considered hybrid techniques to select features (e.g., filters, wrapper, and embedded). The main goal of many traditional approaches for selection is to solve problems of instability and interruption. Hybrid methods are examples of [11-12]. A small data disturbance, for example, can lead to completely different selection results in small, high-dimensional data. The results are consistent and therefore the integrity of the selected functions is managed to improve by combining several selected subassemblies from different methods.

Table I shows the comparative analysis of various existing feature selection techniques in the classification cardiovascular diseases. The support vector machine was preferred as classifier by most algorithms. It is due to the classification of the binary class. Due to the hierarchical arrangement of data only few works were suggested by the random forest classifier. The wrapper selection technique allows the three filter techniques to be highly precisely designed. This is due to the attribute's selection process. The attributes are selected according to the classifier's performance. On this basis, the current Cleveland data set classification technology is selected for the wrapper feature. You can use the grey-wolf optimization algorithm [13] to choose a feature from the dataset here. The selected characteristics were then given training and variable classified. It can achieve 89.83% higher grading rates. However, the wrapper selection algorithms have the following problems.

The wrapper method depends entirely on the grading option. With a different classification, the result can vary.

The classification efficiency determines the selection of fitness and variables.

These disadvantages have been overcome through the proposed global classification based on sensitivity analysis. The following Section III provides a brief explanation of the proposed approach.

TABLE I. COMPARATIVE STUDY OF DIFFERENT METHODS OF EXISTING FEATURE SELECTION

Author	Type	Technique	observations
Subanya and Rajalakshmi [13]	Wrapper selection	Artificial bee colony algorithm on heart diseases dataset SVM classification is performed Only 7 features were used.	<ul style="list-style-type: none"> • Fitness is not defined. • The results do not verify their identification nature • accuracy is the only measure.
Shah et al. [14]	Wrapper selection	Selection of features: mean fishing score and accuracy Reduction of feature: principal analysis component. Radial base kernel-based classification Vector support machine support	<ul style="list-style-type: none"> • Computational time high due to multiple feature extraction. • Cleveland datasets have minimal accuracy when compared to other datasets such as Hungary and Switzerland.
Chiroma et al. [15]	Wrapper selections	Rule based algorithm like Prism is used for feature selection and classification Wrapper algorithms like decision tree, LibSVM, K-nearest neighbor were used. Wrapper algorithm utilized forward selection and backward elimination algorithm	<ul style="list-style-type: none"> • Cleveland's f-score is a minimum of the other dataset. • As with prism, the algorithm of the Cleveland wrapper can also produce better results.
Moorthy and Gandhi [16]	Wrapper selection	Combination of ANOVA and whale optimization is used. Classification: naïve bayes, SVM and K-NN	The whale optimization algorithm for Cleveland data sets in any classification type can achieve greater precision. Compared to others, SVM produced Cleveland's best results in classification. However, it used all the classification attributes.
Wang and Li [17]	Wrapper selection	Immune system artificial	The best sonar and cardiac statolog result are produced. It can be as accurate as the Cleveland data set approach.
Saqalin et al. [18]	Filter selection	Selection of the feature: fisher score, selection forward and selection backward. Reduction feature: coefficient of mathematical correlation Classification: SVM based on radial basis	Cleveland's 81 percent data set is less accurate. The Switzerland heart disease dataset produced the best result. Less feature reduces the performance of the classification.
Garate Escamila et al. [19]	Filter selection	Chi-square and Principal Component analysis is used for feature selection. Classifier: six machine learning algorithms is used	<ul style="list-style-type: none"> • Random forest produced the best classification result. • Processing time is high for more attributes.
Gupta et al. [20]	Filter selection	Correlation and Squared correlation are used for feature selection. Random forest classifier is best among many classifiers	Able to achieve high classification rate. The number of features is high.
Ayon et al. [21]	Filter selection	Feature selection: Correlation Classifier: logistic regression, and many machine learning algorithms	Five- fold SVM based classification produced the best result. Real time features require modification
Muhammad et al. [22]	Ensemble Selection	Feature selection: LASSO, Relief, mRMR Different classifiers.	Proposal is made based on the analysis. Optimization techniques can improve the results further.

III. PROPOSED METHODOLOGY

This work introduces efficient approaches to select features such as integrated sensitivity and correlation and a global hybrid feature sensitivity analysis. We're using the data set for cardiac disease in Cleveland [23] here. It consists of 76 variables, 14 of which were only selected for use. Of the 14

attributes, 13 are predictors, and the final attributes are the target. In studies, 270 cases, 120 of which were categorized as CHD patients and 150 cases as CHD-free patients, were considered for the elimination of missing-value cases [14]. The characteristics and range of values are explained in Table II.

TABLE II. CLEVELAND HEART DISEASE DATASET [23]

Features	Description	Ranges
age	Age (in years)	29-77
sex	Gender	1: male; 0: female
Cp	chest pain type	0-3
Trestbps	resting blood pressure	94-200
Chol	serum cholesterol in mg/dl	126-564
Fbs	Fasting blood sugar	0-1
Restecg	resting electrocardiographic results	0-2
thalach	maximum heart rate achieved	71-202
Exang	exercise induced angina	0-1
Oldpeak	ST depression induced by exercise relative to rest	0-6.2
Slope	the slope of the peak exercise ST segment	0-2
ca	number of major vessels (0-3) colored by flourosopy	0-4
thal	3 = normal; 6 = fixed defect; 7 = reversable defect	0-3

A. Feature Sensitivity

The first stage uses the sensitiveness approach to identify the sensitivity to output variations in a particular model for each input factor. The results from the feature sensitivity analysis are highly dependent on the factors to be carefully selected. We have here classified their significance in determining the CHD risk. Ranks have been determined using functional sensitivity in a learned classification algorithm. After removing the least preferred characteristics, design was gradually trained according to these rankings. This phase continued until compared with the previous one the model performance deteriorated. In this approach, we examine the model, in order to analyze the differences between the characteristics of the development of the learning model.

The sensitivity of the i^{th} feature $Sen(M, m_i)$ is determined by a condition that differs between the original and the deformed data set by adding very little noise (known as μ) in the developed model.

$$Sen(M, m_i) = \frac{1}{N} \sum_{v_k} | RFO_{\text{Output}_k}(M(m_i + \delta)) - RFO_{\text{Output}_k}(M) | \quad (1)$$

In the $RFO_{\text{Output}_k}(M)$ and $RFO_{\text{Output}_k}(M(m_i + \delta))$, the inputs k are the output, with the original input data set M , and then the result with a noisy input ($X(x_i + \delta)$) is the very small noise δ to i^{th} . All sensitivities were measured individually with a single sensitivity. μ was randomly selected in the range [a1, 0.0010].:

B. Feature Correlation

We analyze the characteristics of model prediction outcomes and evaluate them. When changes were affected by features in the input for the preview performance, features were deemed correlated. This means that the value of the property is increased when training the model if a feature improves its severity. In addition, the relating characteristics can be compared if the size of the increase greatly affects the other features. Selection of features Correlation sees the class and value-based correlation of the subfunctions as an ideal set of characteristics:

$$CFS = \max_{s_i} \frac{kr_{cf}^-}{\sqrt{k+k(k-1)r_f}} \quad (2)$$

Where $(r_{cf})^-$ the average correlation of all the features is equal to r_f^- the average correlation of all features-class.

In this study, we analyzed features for category relationships and evaluated if they were correlated with the results of classification predictions mainly during the feature correlation analysis stage. If any of the features affects the possibility that the correlation will contribute to the output, the features were considered to be correlated.

Fig. 5 shows the flow diagram for our proposed integrated sensitivity model and correlation. The flowchart begins with the data pre-processing where missing data is processed using a data imputation method followed by min-max standardization. Any negligible input features can reduce the output of the classifier. It is therefore very difficult to select from a collection of features for the prediction mission an exact and rigorous set of attributes. Feature selection is made through the combination of feature sensitivity and functional correlation in the presented design. Each approach will evaluate the rating of features and then measure the value of the response variable by using the amount from both approaches. Functions are selected first in order to increase 1 to 13, and then in the second scenario, in order to reduce 13 to 1. This can be used to calculate and check the optimal feature subset. Random forest models are generated after the input rating with different number of features in order to estimate heart disease. A novel integrated feature selection result is linked in comparison with existing classification models like naive Bayes, decision tree, regression analysis and support vector machine.

The following is a pseudocode for feature selection based on sensitivity analysis:

Input: X = x1, x2, x3, ...xi.... xn /* features of Cleveland dataset
#Choose the feature subset based on sensitivities
#Assess all feature sensitivities
Rank the characteristics according to their sensitivities
#Add the features according to their rank to the feature subset
Output: Xs /* Chosen feature sub-set */

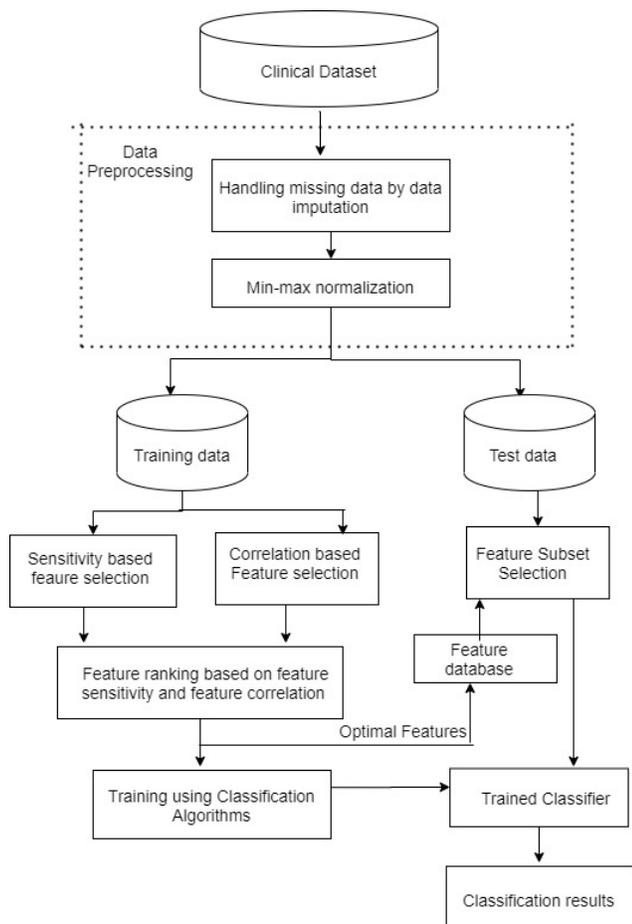


Fig. 5. Flowchart of Integrated Feature Sensitivity and Correlation Approach.

C. Global Sensitivity analysis - Particle Swarm optimization Feature Selection

This proposes an optimized filter technique for the classification process feature selection as shown in Fig. 6. There have been two phases in which to select classification attributes. The very first stage is based on a global analysis of feature sensitivity. In the second stage, a wrapper optimization determines the predominate attribute from the first phase. Due to this multi-stage functions selection, the framework developed can be used to apply every type of machine learning application. The purpose of the approach proposed is to attain the following objectives.

- 1) The method proposed does not depend on the first phase ranking of the selection of features.
- 2) Larger data sets can offer greater precision.

The importance of the attribute is determined by three individual filtering approaches:

- 1) The coefficient of correlation of the input vector to the objective output is calculated by Pearson. The correlation values are -1 and 1. The classification depends on the negative correlation approach. The attribute is listed below and the negatively correlated attribute is higher. Each ranking of attributes is performed on this basis.

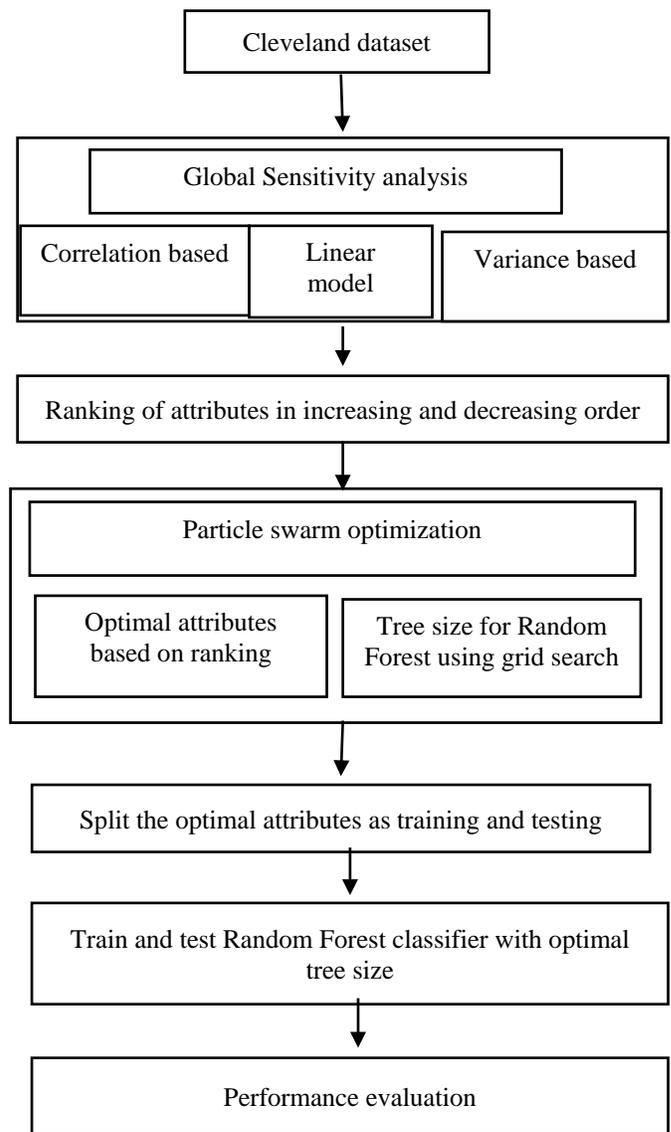


Fig. 6. Flowchart of GSA-PSO Feature Selection.

2) *Linear model fitting*- The value of the attribute is determined by the minimum average vector input error. Input attributes X1 to Xn and destination Y are included in this data set.

3) *Variance based*- The variance between the input and the target vector is calculated in order to determine its importance. Two steps for determining the difference. First, it calculates the medium of the attribute. The significant difference is also used to calculate the difference between the single attribute and its mean value.

D. Classification Methods

1) *Random forest*: The Random Forest (RF) as a supervised method of learning has been introduced recently to engineering practice [23–25]. This RF procedure combines two powerful ML techniques, bootstraps [26] and a random subspace [27]. It is therefore excellent to generalize this technique, as it adds the results of several decision-making

bodies, but computer cost is substantial. N arbitrary samples and trainings are derived using the bagging technique in this algorithm. Bootstrap sets are used to decide trees. Each node tests a feature and the leaf nodes are the output labels. The solution is achieved by combining all of the outputs [28-29], as follows: The solution is:

$$m = \frac{1}{n_{tree}} \sum_{i=1}^{n_{tree}} m_i(x) \quad (3)$$

Where y = the mean output of the total ntree amount;
 $m_i(x)$ = the prediction of the individual tree for the vector input x .

Fig. 10 illustrates the RF model comparison with other learning models. Fig. 11 shows the RF model comparison to the Proposed RF FSFC.

Pseudocode for feature selection based on GSA-PSO feature selection is given below:

Inputs: Cleveland dataset with thirteen attributes (Mn) and two classes (Pn)
Begin GSA
CR=corr(Mn,Pn) /* GSA based on Correlation*/
LM=fitlm(Mn,Pn) /*GSA based on Linear fit Model*/
VM=Var (Mn,Pn) /* GSA based on Variance-based*/
End GSA
#Ranking of attributes in increasing and decreasing order
Begin PSO /*Particle Swarm optimization*/
#Optimal attributes based on ranking
#Find Tree size for random forest using Grid search
End PSO
#Simulate the model with train and test/* Split the optimal attributes as training and testing*/
#Evaluate the model Performance
Output: Performance evaluation, Per. /* Performance evaluation of the classifier with optimal tree size*/

E. Optimization of the Feature and Size of the Tree by Particle Swarm

The optimal attribute and tree size is determined for the random forest classification via an optimization approach. The best solution for the problem is to solve the fitness function. The fitness focus is to prevent the rate of errors in the random forest classification. The fitness function is given with the following equation 5.

$$FitnessFunc_{GSA-PSO} = \text{minimum}(\text{error rate of RandomForest}) \quad (4)$$

This is achieved by finding the optimum attribute through the grid search algorithm based on a global sensitivity analysis and the best tree size between 10 and 130. By minimizing the classificatory error rate on the search algorithm, the optimum tree can also be determined. The common fitness function for optimizing the algorithm is therefore used.

IV. IMPLEMENTATION, RESULTS AND DISCUSSIONS

With the MATLAB Software R2018a, the proposed method is presented in Windows 10. The selection of characteristics based on integrated sensitivity and correlation is computed and the selection of characteristics calculated using the Global Particle Swarm Optimization Sensitivity analysis. And the approaches are compared on the basis of accuracy, sensitivity, specificity and time. The Table III shows an integrated selection of sensitivities and correlation functions. Therefore, the sensitivity of properties is determined by Equation 1 and the correlation between properties is calculated by Equation 2. The higher the sensitive value and the lower the sensitive one. And the less correlated feature is higher. In the case of the calculation of a total function range, features of the same rank will be given priority.

By combining the ranking of these techniques and identifying the importance of each attribute in classifying the data set, the integrated feature selection procedure by feature sensitivity and correlation is performed. Table III shows the general classification of these two methods for each attribute. Likewise, in order for the importance of attributes to be determined during the grading process, there was a combination of rankings with Pearson's correlation and variance-based tests as shown in Table IV. There have been sensitivity analyses.

The classified attributes were arranged in two ways in an increasing and decreasing order to determine the optimized attributes for the classification process. The ranking of lower to higher attributes is indicated by the increasing order. Lower order indicates the ranking of the higher to lower attributes.

The Cleveland dataset is used for random classification based on the above optimal attributes. The classification system is trained to increase order formats with the optimal attributes of 1 to 13 percent and 70 percent of the data. The remaining 30 percent data is then used to test and evaluate trained classifiers.

The greater ranking of an integrated method that allows patients with 93.72% sensitivities effectively to identify cardiac diseases compared with the control patient with 83.28% specificity. The overall precision of the increased order rating is 89.921%.

TABLE III. RANKING OF ATTRIBUTES WITH INTEGRATED FEATURE SENSITIVITY AND CORRELATION

S.no	Attribute_name	Integrated Sensitivity and Correlation		Ranking			Overall Ranking Score	
		Sensitivity	Correlation	S	C	Total Rank	Increasing order	Decreasing order
1	age	0.024	0.2254	6	8	14	6	8
2	sex	0.048	0.2809	5	9	14	7	7
3	Cp	0.003	-0.4338	13	1	14	4	10
4	Trestbps	0.011	0.1449	9	7	16	10	4
5	Chol	0.005	0.0852	11	6	17	11	3
6	Fbs	0.1	0.028	1	5	6	1	13
7	Restecg	0.049	-0.1372	4	4	8	2	12
8	thalach	0.004	-0.4217	12	2	14	5	9
9	Exang	0.082	0.4368	2	13	15	9	5
10	Oldpeak	0.073	0.4307	3	12	15	8	6
11	Slope	0.01	-0.3459	10	3	13	3	11
12	ca	0.012	0.3917	8	11	19	13	1
13	thal	0.013	0.344	7	10	17	12	2

TABLE IV. RANKING OF ATTRIBUTES USING GSA APPROACH

S.no	Attribute_name	Sensitivity factors			Ranking				Overall Ranking Score	
		Variance	Pearson	Significance	V	P	L	Total Rank	Increasing Order	Decreasing Order
1	age	82.21233	0.2254	0.76113	4	8	12	24	8	6
2	sex	0.216449	0.2809	4.2449e-05	12	9	3	24	9	5
3	Cp	1.061617	-0.4338	8.4015e-07	6	1	1	8	1	13
4	Trestbps	306.5713	0.1449	0.11441	3	7	9	19	5	9
5	Chol	2677.560	0.0852	0.40255	1	6	11	18	3	11
6	Fbs	0.126458	0.0280	0.77112	13	5	13	31	13	1
7	Restecg	0.275615	-0.1372	0.21282	10	4	10	24	10	4
8	thalach	522.9148	-0.4217	0.0079882	2	2	5	9	2	12
9	Exang	0.219978	0.4368	0.0053868	11	13	6	30	12	2
10	Oldpeak	1.343646	0.4307	0.010847	5	12	8	25	11	3
11	Slope	0.378481	-0.3459	0.063453	8	3	7	18	4	10
12	ca	1.042272	0.3917	6.2486e-06	7	11	2	20	6	8
13	thal	0.373645	0.3440	0.00095231	9	10	4	23	7	7

The reduction of ranking classifications is also trained, and all its attributes are tested using the random classification using 70% of data and 30% of data. In comparison with normal patients in the lower classifications, it can also efficiently detect cardiac patients. Its total precision, however, is 78.32%.

Table V and Fig. 7 shows the comparison between the two approaches proposed by the use of performance assessment in increased and decreasing order.

TABLE V. COMPARISON OF THE INTEGRATED FSFC APPROACH TO PERFORMANCE ASSESSMENT

Method	Increased order ranking (%)	Decreased order ranking (%)
Accuracy	89.921	78.32
Sensitivity	93.72	80.23
Specificity	83.28	78.47

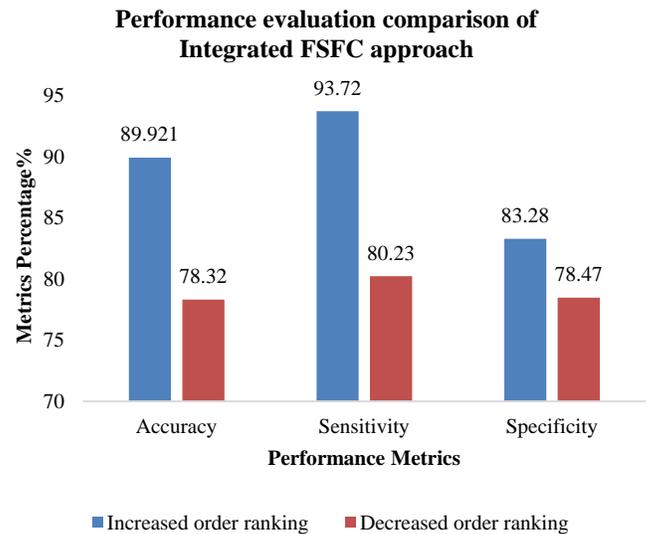


Fig. 7. Comparison of the Integrated FSFC Approach to Performance Assessment.

TABLE VI. COMPARISON OF GSA APPROACH PERFORMANCE ASSESSMENT

Method	Increased order ranking (%)	Decreased order ranking (%)
Accuracy	90.00	73.33
Sensitivity	94.74	69.23
Specificity	81.82	76.47

Performance evaluation comparison of GSA approach

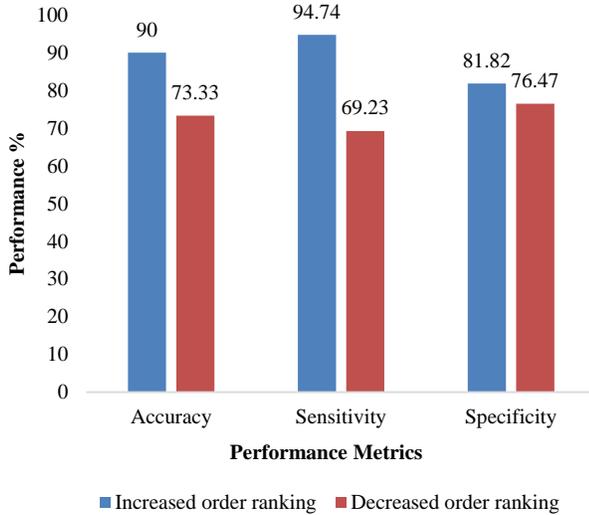


Fig. 8. Performance Evaluation Comparison of GSA Approach.

Similarly, with 94.74% sensitivity, 81.82% specialty and the overall accuracy for the increased ranking classifier is 90.00%, the GSA method. And with a sensitivity of 69.23 percent, 76.47 percent are declining and 73.33% are indicated in Table VI and Fig. 8 for the total accuracy of the classifier.

The performance evaluation shows that the higher rankings can effectively classify heart disease by having higher precision compared to decreasing order precision.

This means that in the Cleveland dataset, the classification of patients with normal and heart disease is best determined by the increasing order.

Then, the proposed integrated approach and global sensitivity analysis approach is compared to the existing wrapper selection method with the same Cleveland data set, using a grey wolf optimization support vector machine classifier.

In addition, our approaches, when compared with the existing wrapper selection performance, are shown in Table VII and Fig. 9 compared with the current wrapper 93% above the wrapper selection by precise determination of cardiac diseases with high sensitivity at 93.72% and 94%.

TABLE VII. PERFORMANCE COMPARISON BETWEEN GSA AND EXISTING WRAPPER SELECTION

Method	GSA analysis (%)	Integrated sensitivity and correlation analysis (%)	Existing wrapper selection GWO-SVM
Accuracy	90.00	89.921	89.83
Sensitivity	94.74	93.72	93
Specificity	81.82	83.28	91

Performance comparison between Proposed approaches and Existing Wrapper selection

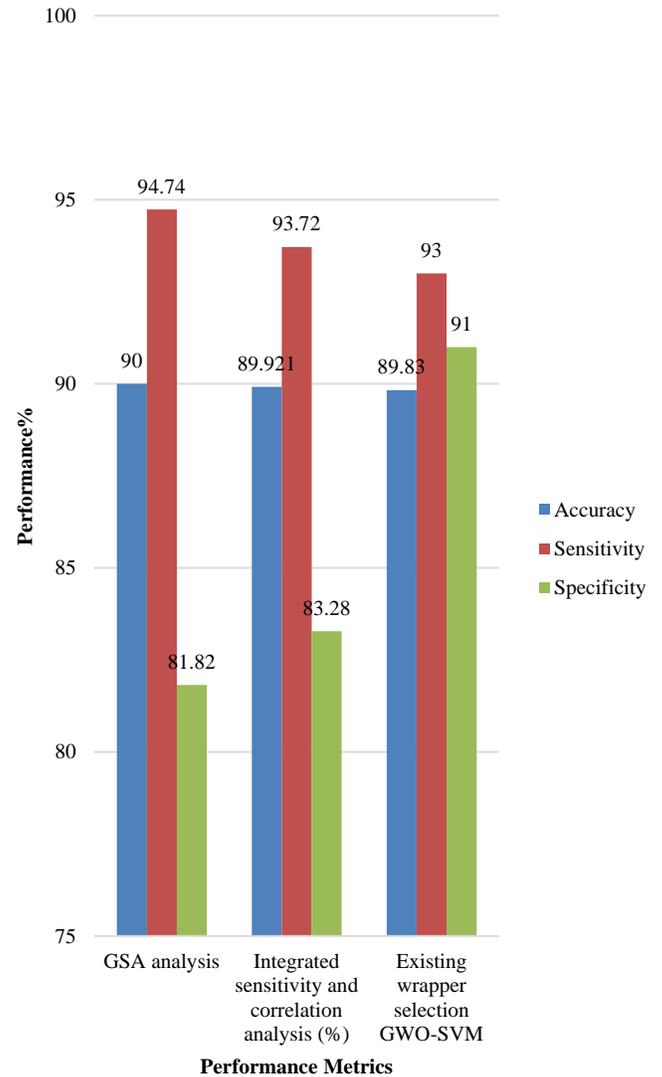


Fig. 9. Performance Comparison between Proposed Approaches and Existing Wrapper Selection.

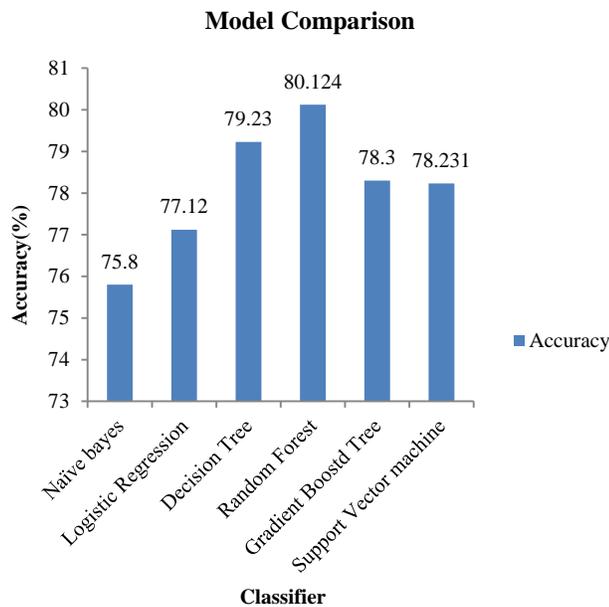


Fig. 10. Comparison of RF Model with the Other Learning Models.

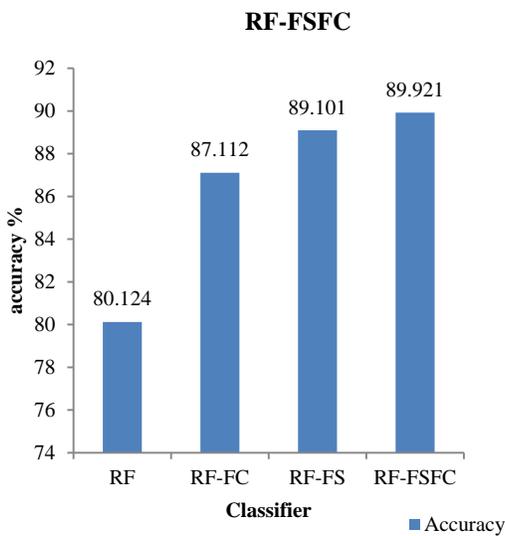


Fig. 11. Comparison of RF Accuracy based on Sensitivity and Correlation.

V. CONCLUSION

The purpose of the features is assessed based on an integrated approach to sensitivity and correlation and a global sensitivity analysis based on optimization. The integrated sensitivity and correlation of properties are implemented in two phases as follows:

- 1) The first stage is to rank each attribute based on feature sensitivity analysis between the vector and the target.
- 2) The second phase consists of classifying each attribute by a correlation between the vector function and the variable objective.

The global sensitivity analysis is similarly conducted in two phases:

- 1) The first step is a classification based on different sensitivity analyses of each attribute.
- 2) The second Phase defines a more sensitive grading attribute based on the optimization of particle swarm.

The proposed approaches assess the ranking of the random forest classifier's trees and the reduced order ranking with optimum dimensions. It then uses its accuracy, sensitivity and specificity to assess its performance. The lower to higher classification produced the best results for heart disease patients. This classification helps to avoid the least classified features.

By comparing the performance with the wrapper selection-based classification, our proposed approaches integrated feature sensitivity and feature correlation outperforms it with the accuracy 89.921% and the sensitivity with 93.72% and also the global sensitivity analysis outperforms the wrapper selection by finding heart diseases accurately with the accuracy 90% and high sensitivity of above 94% as compared to existing of 93%.

The integrated approach to feature sensitivity and correlation and the global approach to sensitivity play a significant role in choosing the best classification feature especially in comparison to each feature selection procedure.

VI. FUTURE ENHANCEMENT

With the aim of reducing calculation times for the selection of attributes and improving the performance of the learning model, the results from our suggested approaches can be further strengthened.

REFERENCES

- [1] A Shroff, K. P., & Maheta, H. H. (2015). A comparative study of various feature selection techniques in high-dimensional data set to improve classification accuracy. 2015 International Conference on Computer Communication and Informatics (ICCCI). <https://doi.org/10.1109/iccci.2015.7218098>.
- [2] Kishore, R., & Tripathi, S. (2016). A comparative analysis of enzyme classification approaches using hybrid feature selection technique. 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT).
- [3] El Aboudi, N., & Benhlima, L. (2016). Review on wrapper feature selection approaches. 2016 International Conference on Engineering & MIS (ICEMIS). <https://doi.org/10.1109/icemis.2016.7745366>.
- [4] Yin, P., Mao, N., Zhao, C., Wu, J., Sun, C., Chen, L., & Hong, N. (2018). Comparison of radiomics machine-learning classifiers and feature selection for differentiation of sacral chordoma and sacral giant cell tumour based on 3D computed tomography features. *European Radiology*, 29(4), 1841-1847.
- [5] Saranya, G., & Pravin, A. (2020). Feature selection techniques for disease diagnosis system: A survey. *Artificial Intelligence Techniques for Advanced Computing Applications*, 249-258. https://doi.org/10.1007/978-981-15-5329-5_24.
- [6] M Robnik-Šikonja, I Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF" - Machine learning, 2003 – Springer.
- [7] Feng Yang and K. Z. Mao. 2011. Robust feature selection for microarray data based on multicriterion fusion. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 8, 4 (2011), 1080–1092.
- [8] Lei Shi, Liang Du, and Yi-Dong Shen. 2014. Robust spectral learning for unsupervised feature selection. In *ICDM*. 977–982.

- [9] Hiromasa Arai, Crystal Maung, Ke Xu, and Haim Schweitzer. 2016. Unsupervised feature selection by heuristic search with provable bounds on suboptimality. In AAAI. 666–672.
- [10] Y Guo, J Ji, H Huo, T Fang, D Li, “SIP-FS: a novel feature selection for data representation on Image” 2018 - jivp-urasipjournals.springeropen.
- [11] X Liu, J Tang, “Mass classification in mammograms using selected geometry and texture features, and a new SVM-based feature selection method” - IEEE Systems Journal, 2014 - ieeexplore.ieee.org.
- [12] D Singh, AA Gnana, “A novel feature selection method for image classification” 2015.
- [13] Subanya, B., & Rajalaxmi, R. R. (2014, February). Feature selection using Artificial Bee Colony for cardiovascular disease classification. In 2014 International Conference on Electronics and Communication Systems (ICECS) (pp. 1-6). IEEE.
- [14] Shah, S. M. S., Shah, F. A., Hussain, S. A., & Batool, S. (2020). Support Vector Machines-based Heart Disease Diagnosis using Feature Subset, Wrapping Selection and Extraction Methods. Computers & Electrical Engineering, 84, 106628.
- [15] Chiroma, F., Cocea, M., & Liu, H. (2019). Evaluation of rule-based learning and feature selection approaches for classification. OASICS.
- [16] Moorthy, U., & Gandhi, U. D. (2020). A novel optimal feature selection technique for medical data classification using ANOVA based whale optimization. Journal of Ambient Intelligence and Humanized Computing, 1-12.
- [17] Wang, Y., & Li, T. (2020). Local feature selection based on artificial immune system for classification. Applied Soft Computing, 87, 105989.
- [18] Saqlain, S. M., Sher, M., Shah, F. A., Khan, I., Ashraf, M. U., Awais, M., & Ghani, A. (2019). Fisher score and Matthew’s correlation coefficient-based feature subset selection for heart disease diagnosis using support vector machines. Knowledge and Information Systems, 58(1), 139-167.
- [19] Gárate-Escamila, A. K., El Hassani, A. H., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. Informatics in Medicine Unlocked, 19, 100330.
- [20] Gupta, A., Kumar, R., Arora, H. S., & Raman, B. (2019). MIFH: A machine intelligence framework for heart disease diagnosis. IEEE Access, 8, 14659-14674.
- [21] Ayon, S. I., Islam, M. M., & Hossain, M. R. (2020). Coronary artery heart disease prediction: a comparative study of computational intelligence techniques. IETE Journal of Research, 1-20.
- [22] Muhammad, Y., Tahir, M., Hayat, M., & Chong, K. T. (2020). Early and accurate detection and diagnosis of heart disease using intelligent computational model. Scientific reports, 10(1), 1-17.
- [23] Cleveland Clinic Foundation, "Heart Disease Data Set ", Available at: <http://archive.ics.uci.edu/ml/datasets/Heart+Disease>.
- [24] J. Zhou, X. Shi, K. Du, X.Y. Qiu, X.B. Li, H.S. Mitri, Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel, Int. J. Geomech. 17 (2016) 04016129.
- [25] Y. Zhou, S.Q. Li, C. Zhou, H.B. Luo, Intelligent approach based on random forest for safety risk prediction of deep foundation pit in subway stations, J. Comput. Civ. Eng. 33 (2019) 05018004.
- [26] P. Zhang, R.-P. Chen, H.-N. Wu, Real-time analysis and regulation of EPB shield steering using random forest, Automat. Constr. 106 (2019) 102860.
- [27] Saranya, G., & Pravin, A. (2020). A comprehensive study on disease risk predictions in machine learning. International Journal of Electrical and Computer Engineering (IJECE), 10(4), 4217. <https://doi.org/10.11591/ijece.v10i4.pp4217-4225>.
- [28] Hemanth Reddy, K., & Saranya, G. (2020). Prediction of cardiovascular diseases in diabetic patients using machine learning techniques. Artificial Intelligence Techniques for Advanced Computing Applications, 299-305. https://doi.org/10.1007/978-981-15-5329-5_28.
- [29] A. Liaw, M. Wiener, Classification and regression by random forest, R News 23 (2002) 18–21.

Energy Material Network Data Hubs

Software Platforms for Advancing Collaborative Energy Materials Research

Robert R. White¹

Materials, Chemical, and
Computational Science
(MCCS) Research Operations
National Renewable Energy Laboratory
Golden, CO USA

Kristin Munch²

Materials, Chemical, and
Computational Science
National Renewable Energy Laboratory
Golden, CO USA

Nicholas Wunder³, Nalinrat Guba⁴
MCCS Data Analysis and Visualization
National Renewable Energy Laboratory
Golden
CO USA

Chitra Sivaraman⁵

Computational and Data Engineering
Group
Pacific Northwest National Laboratory
Richland, WA
USA

Kurt M. Van Allsburg⁶

Catalytic Carbon Transformation and
Scale-Up Center, National Renewable
Energy Laboratory, Golden, CO USA

Huyen Dinh⁷

MCCS Chemistry and Nanoscience
National Renewable Energy
Laboratory, Golden, CO USA

Courtney Pailing⁸

Maxar Technologies
Westminster, CO USA

Abstract—In early 2015 the United States Department of Energy conceived of a consortium of collaborative bodies based on shared expertise, data, and resources that could be targeted towards the more difficult problems in energy materials research. The concept of virtual laboratories had been envisioned and discussed earlier in the decade in response to the advent of the Materials Genome Initiative and similar scientific thrusts. To be effective, any virtual laboratory needed a robust method for data management, communication, security, data sharing, dissemination, and demonstration to work efficiently and effectively for groups of remote researchers. With the accessibility of new, easily deployed cloud technology and software frameworks, such individual elements could be integrated, and the required collaboration architecture is now possible. The developers have leveraged open-source software frameworks, customized them, and merged them into a platform to enable collaborative energy materials science, regardless of the geographic dispersal of the people and resources. After five years in operations, the systems are demonstratively an effective platform for enabling research within the Energy Material Networks (EMN). This paper will show the design and development of a secured scientific data sharing platform, the ability to customize the system to support diverse workflows, and examples of the enabled research and results connected with some of the Energy Material Networks.

Keywords—Energy materials research; cloud computing; virtual laboratories; data management; consortium; network

I. INTRODUCTION

The Energy Materials Network (EMN) is a United States Department of Energy (DOE) and Office of Energy Efficiency and Renewable Energy (EERE) network of consortia formed to accelerate the process of materials discovery, characterization, scale-up, and commercial deployment focused on solving the nation's toughest materials challenges in the energy sector. Building on the working concepts of High-Throughput

Experimentation [1] and the Materials Genome Initiative [2], the EMN was envisioned to be a coordinated resource network for advanced materials R&D, enabling industry and university access to world class materials capabilities at the National Labs. The concept of the EMN can be defined broadly as a virtual laboratory. The concept of virtual laboratories has been around for over a decade but has primarily been centered on the concept of supporting educational activities for remote students [3]. Unlike these virtual laboratory ideas, this concept is not a reference to a simulated environment but is centered around extending research accessibility in the real world. However, the key points that make it ideal for educating remote students, can make it equally beneficial for bringing together geographically remote resources of expertise and equipment to advance research. Bridging that expertise and equipment allows specific capabilities in each remote lab to be leveraged efficiently and provide cost saving measures in reduced equipment purchases and travel. Additionally, a virtual laboratory can still operate effectively in a changing world where environmental conditions demand social distancing and limited travel.

To facilitate this kind of research structure requires a common communication and data sharing platform for the collaborations to function efficiently. A major underpinning of the EMN strategy is the data and tool collaboration framework referred to as the Data Hub. The Data Hub is the means for capturing data, tools, and expertise developed at each of the EMN consortia facilities (called nodes) such that they can be shared and leveraged throughout the EMN and in future programs. The EMN Data Hub infrastructure facilitates accelerated learning and materials development through the establishment of data repositories, distribution of data to the greater scientific community, and development of data informatics tools. The Data Hub must also address certain challenges with creating any viable repository; data quality,

result duplication, provenance, relevance, data standards, security, and access [20].

Each Data Hub supports collaborative materials science through the establishment of accessible, searchable data resources for its EMN consortium. The Data Hubs host materials data associated with each consortia's technical portfolio, and integrate data of heterogeneous data types, sizes, and sources, including materials experimental results, theoretical and simulation data from computation, remote automated data acquisition from multiple sources, and performance and characterization benchmarking data. The Data Hubs enable robust data workflows from consortia resources and define metadata and standards for each technical area. Perhaps most importantly, the EMN Data Hubs are collectively managed by each consortiums Data Team, a group of data experts representing each member lab that focuses on coordinated development efforts, thereby increasing opportunities for data analytics and data outreach (see Fig. 1). It cannot be understated that establishing useful data hubs for materials science takes a team effort and lots of coordination and communication to facilitate development that crosses so many technical, data, and data source boundaries.

Developing a data archiving and collaboration platform that can consolidate experimental and theoretical data along with analytics tools presents many challenges. Due to the limited

development resources available, EMN Data Hubs needed a software framework that is not only customizable within the needs of each technological domain of a particular EMN, but one that can be utilized across all the EMNs. The Data Hub must house a variety of heterogenous instrumentation data, which would need to be not only archived, but contextually searchable requiring developing both standards of metadata elements along with means to facilitate metadata capture. To allow for programmatic access by analytic tools and to facilitate automated data uploads, an Application Programming Interface (API) would either need to be available or constructed for any Data Hub. Since the system would be supporting active projects and not just public release of data, any Data Hub would need project level security to restrict access to protected datasets.

This article discusses the architecture of the EMN Data Hubs with respect to both implementation, operations, and collaboration details, with an emphasis on showcasing capabilities and data available now on several of the EMN Data Hubs. The article will discuss the data governance concerns and the EMN Data Hub approach to data security and data accessibility, including efforts for standardized metadata for materials science datasets. Finally, there is a discussion of the ongoing virtual laboratory development efforts across the EMN's and the important research being conducted within them.

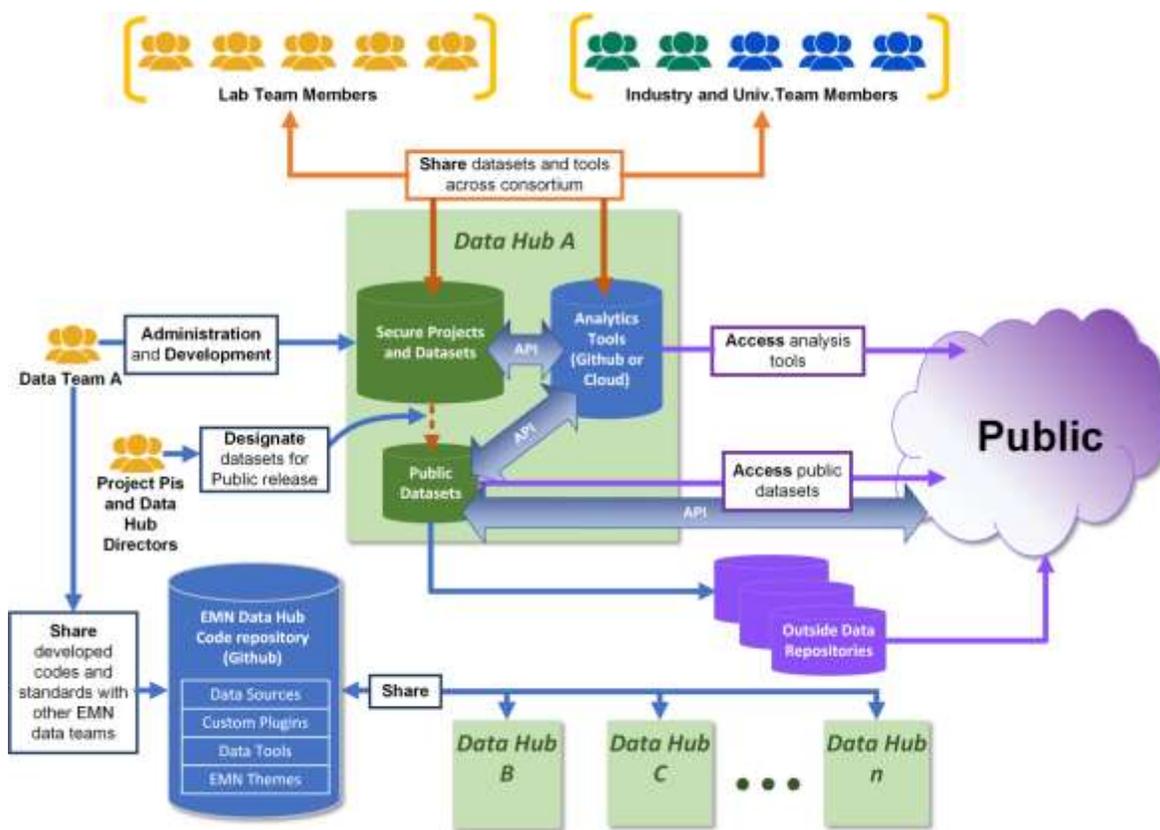


Fig. 1. EMN Data Hubs Collaboration Workflow. Not only do the Researchers Share Data and Resources Across the Platforms, but the Developers can Share Code and Administrative Resources Allowing for a More Efficient Development to Support Multiple Data Hubs. EMN Researchers from Across the National Labs, Industry and Academia Leverage the Collaborative, Virtual Lab Capabilities of the EMN Data Hubs. Each EMN Supports a Data Team, Made up of Data Experts Responsible for the Development and Operations of the Data Hub and Developed Resources are Shared among Data Team Members. Data is Released to the Public through a Managed Data Release Process.

II. COLLABORATIVE SCIENCE

From writing and publishing papers to experimental research, a vast majority of science is done as a collaborative effort. Some work is done in isolation, but often that piece is only part of a larger collaborative effort. While many ubiquitous modern technological elements can improve remote collaborative efforts (e.g. Dropbox, Google Drive, Teams, Slack, Figshare, etc.), there are often limits or pitfalls with many of those that had to be addressed for the EMN Data Hubs so that it could facilitate more efficient remote collaboration. The EMN Data Hubs were also able to facilitate the varied workflows and processes that researchers would normally use in more traditional collaborations. Many of these aspects were a result of the creation of EMN data teams, comprised of senior scientists, data architects and software engineers, to customize and refactor the basic software platform to meet each EMN's needs.

One of the basic research workflows needed by all the EMNs is round-robin experimentation (see Fig. 2), and a good example of this workflow can be found in one of the initial DuraMAT projects studying of hydrophilic and hydrophobic coatings.

Sharing data while the research under way is only one aspect of the collaborations that must be addressed. In any project there are often many ancillary documents and information that the working group will need to access. Constantly emailing documents and information, when project personnel changes, can be an onerous process, often requiring additional time by project leads to make sure the correct and complete information is distributed appropriately. In many cases, there have been established repositories within the Data Hub for both project level and EMN level information covering topics such as benchmarks, data standards and formats, tutorials, other data sources, and supporting research papers.

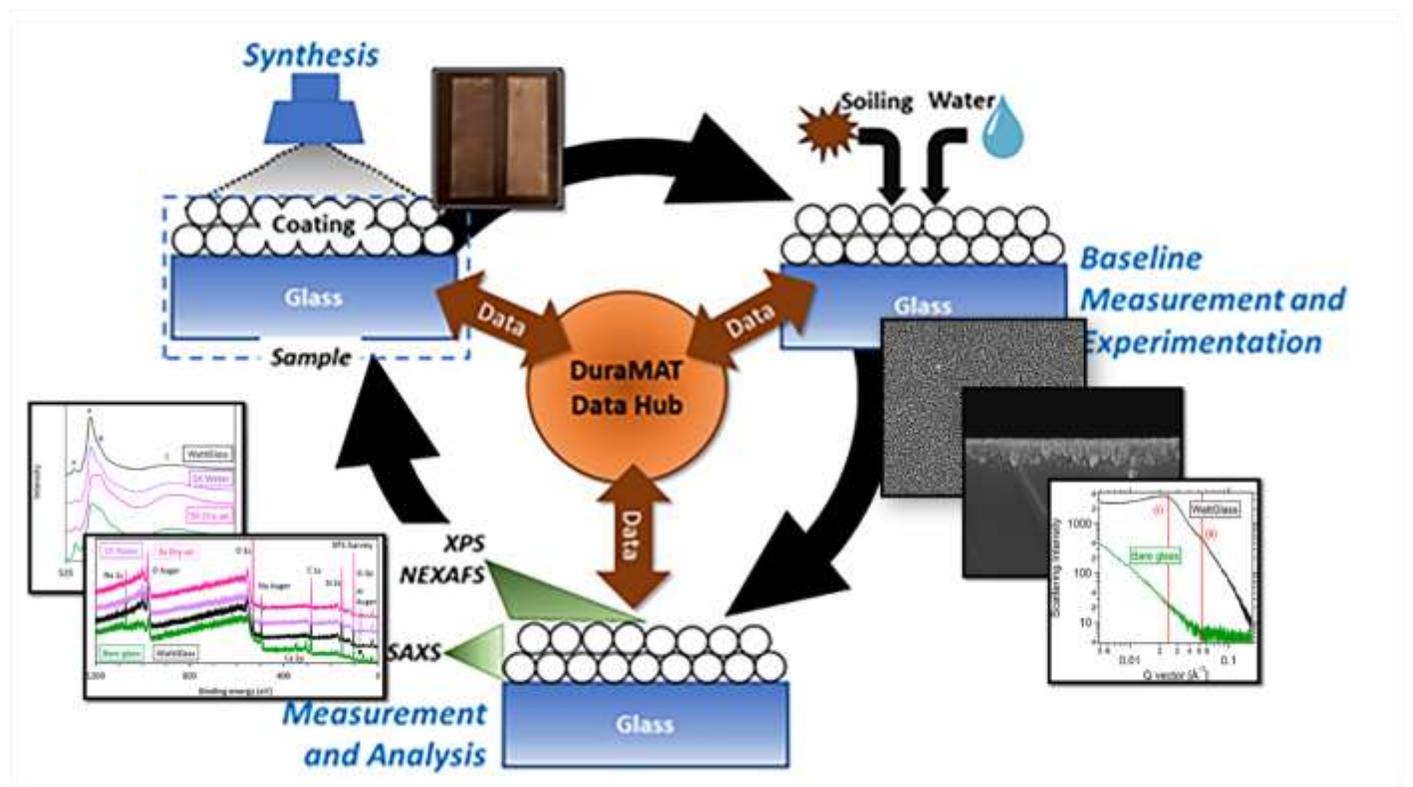


Fig. 2. Round-Robin Experimentation Workflow to Support the Anti-Reflection Anti-Soiling Project on DuraMAT. This Demonstrates How The Dispersed Resources Can Work In Concert Through The Collaborative Platform: Group 1 Performs Synthesis of Coating Material On glass. Group 2 Creates the Initial Baseline Measurements along with Soiling and Wash Cycles through in-Situ or Interim Processes. Group 3 Performs Any Additional Measurements Plus Analysis of Data. The Process Repeats Itself with Subsequent Samples. Data is managed through the Data Hub, Providing each Link in the Chain Access to the Project's Data. All Plots and Images; S.L. Moffitt, R.A. Flemming, et. al. [8].

III. THE EMN DATA HUB INFRASTRUCTURE

The purpose of the EMN Data Hubs is to build up a meaningful data resource from the research being done across each of the EMN nodes. The Data Hubs are designed to enable multiple levels of data governance, so that data can move from private, team-based sharing to public availability within the same infrastructure. Each Data Hub also enables the design of consortium-specific customizations for managing, analyzing, and visualizing the data types common to that consortium. The Data Hubs are hosted in the Amazon Web Services (AWS) cloud and are based on the open-source Comprehensive Knowledge Archive Network (CKAN) platform, which has a plug-in architecture for customizations.

The CKAN platform provides a project and dataset architecture, where resources within datasets can be either files or links to other internet resources. In addition, data teams have implemented several cloud resources as part of the Data Hub environment, to further the availability and manageability of materials data beyond being just a data repository.

There are currently seven EMN Data Hubs:

- LightMat (data.lightmat.org)
- DuraMAT (datahub.duramat.org)

- HydroGEN (datahub.h2aws.org)
- ChemCatBio (datahub.chemcatbio.org)
- ElectroCat (datahub.electrocat.org)
- H-Mat (data.h-mat.org)
- HyMARC (datahub.hymarc.org)

Fig. 3 shows a diagram of the Data Hubs and the cloud resources used in the Data Hub environment. At the Application level, it illustrates the Data Hub web front-end applications, which make use of a centralized authentication service. This allows researchers working within multiple EMN consortia to have a single account across all the Data Hubs they work in. CKAN also enables API accessibility to the data within the Data Hubs, so that researchers and developers can create applications, such as Python analysis and data harvesting methods that work programmatically with the Data Hub repository. The CKAN level is where the CKAN server is located, along with the CKAN database and file system. The cloud architecture enables the incorporation of other custom applications, such as a centralized sample management system, various EMN-specific materials databases, and distributed databases and big data storage.

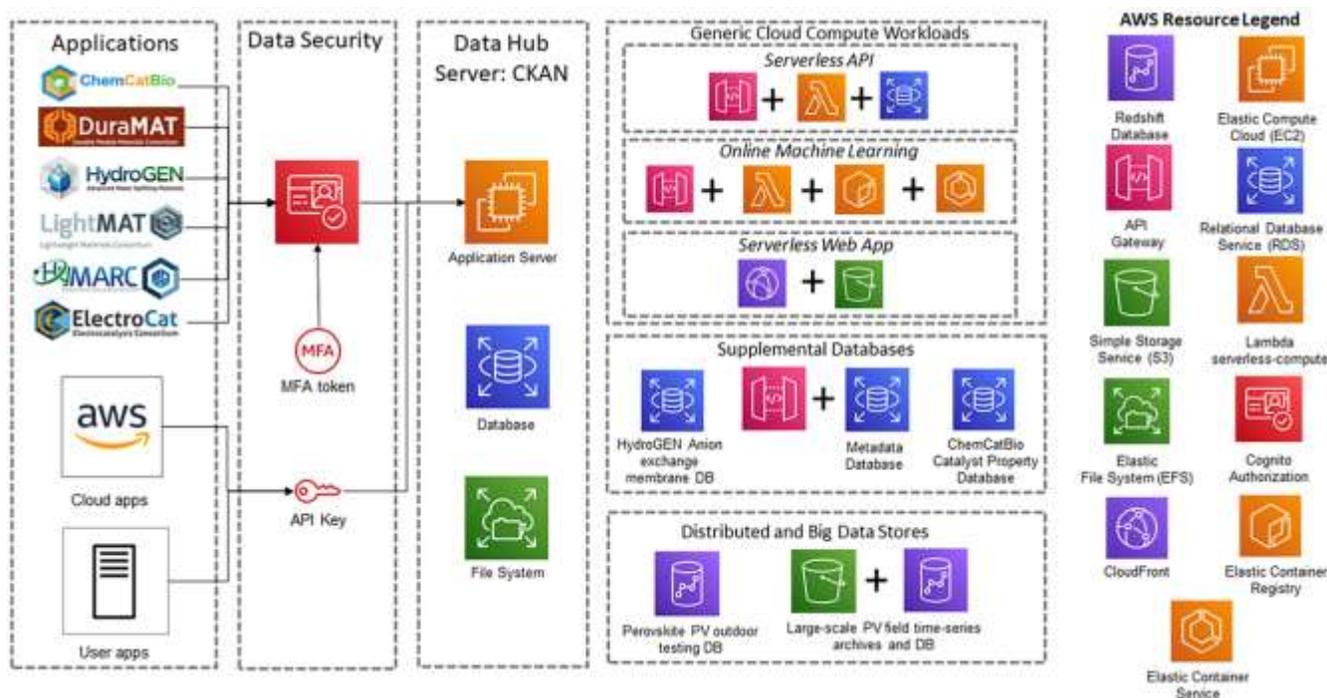


Fig. 3. EMN Data Hub Architecture. The Deployment Location of these CKAN Platform (Columns 2 and 3 on Left) was a Critical Decision that allowed for the Expansion of the Capability through Access to Easily Available, Utility Scale (Cloud) based Storage and Analysis Resources. This Extension of the basic Deployment by Integrating Automated Serverless Processing and Databases Provided the Capability for Sample Tagging and Tracking. Integrated Custom Designed Material Databases and Administration Tools were deployed as Separate Apps from the Main CKAN Deployment. Leveraging the Available Big Data Applications Allowed for Deployment of Cloud-based Time-Series Databases, Large Scale Archives, and Machine Learning and Deep Learning Resources as Ancillary Elements to the Data Hubs.

IV. DATA GOVERNANCE

As part of data governance, the seven Data Hubs managed by multiple institutions desired to adopt a common set of policies and terms. A governance team comprising of the Data Hub representatives, the legal team of the institutions and the DOE's Assistant Chief Counsel for Intellectual Property jointly worked on the language for the terms and agreement document and privacy policy. A common set of terms and agreement (<https://data.lightmat.org/terms>) and privacy policy (<https://data.lightmat.org/privacy>) was adopted by all the Data Hubs, and is available on each Data Hub or from a data team member. This governance team determined that three tiers of data could be hosted on the Data Hub. The Public-Unlimited Rights data is the most open data where data and metadata can be viewed by the public without restriction. This tier of data could be downloaded, but within some of the datahubs, this would require user authentication, so that essential metrics could be gathered on who is downloading data thereby allowing notification when a new version of the data became available. The Embargoed-protected data tier consists of data and metadata that are government sponsored and limited to authorized Non-Disclosure Agreement (NDA) or Cooperative Research and Development Agreement (CRADA) members until the embargo period ends; not to exceed five years. The embargoed-protected data will automatically become public-unlimited after the period. Finally, the Restricted-Proprietary level consists of government or non-government sponsored data that is limited to authorized NDA or CRADA members for a standard maximum of 5 years, unless terms of the agreement specify a longer period. At the end of the period, the data must be removed or released to the public.

All registered users are required to accept the terms and agreement before logging into the Data Hub. Users must also register with the Data Hub to submit data, but most Data Hubs do not accept submissions from outside of the consortium. It is the data producer's responsibility to provide metadata, upload data that accurately represents their work, and ensure that the data can be reproduced. The data producers are responsible to ensure that no personally identifiable information or classified information is uploaded. It is also the responsibility of the users to acknowledge the author, institution and DOI's on any publication related to data downloaded from the Data Hub.

Some of the Data Hubs also require an administrative review of datasets prior to the dataset being made public, to ensure accuracy and completeness. Each Data Hub has its own streamlined process for making a dataset public, which typically includes a committee review and approval. Additional supporting documentation is often required before the dataset is made public, to help understand and utilize the data. Data can be made public on a Data Hub with or without a publication; each Data Hub has an account on OSTI.gov to request a Data Object Identifier (DOI) for a new public dataset. A DOI is an example of a persistent unique identifier (PUD) and is important especially for scientific data because the unique persistent identifier is utilized and referenced in publications; therefore, as a publication will persist, so should the data that supports the research findings.

Today, there are many general-purpose data hubs pertaining to different areas of research, which elevates the importance of data governance for the EMN consortia and others [20]. Considering the growing number of such data repositories and in support of scientific data discovery and re-usability, instead of attempting to host all data, one goal of the EMN Data Hubs is to ensure there exists integrations and linkages to other external databases, datasets, and related resources that pertain to a particular public EMN dataset [21]. The Data Hubs allow external links to reference each dataset; the dataset DOI helps ensure these linkages persist for publications and other references. In some cases, the developers have consolidated dataset hosting on the EMN Data Hubs, as well as provided multiple ways to access and analyze data from the Data Hubs.

A. Data Security

The EMN Data Hubs are hosted on the Federal Risk and Authorization Management Program (FedRAMP) certified AWS cloud platform. The servers running the Data Hub application are secured via 2-factor authentication, follow network security standards and best practices, and have been vetted against DOE Cyber Security compliance rules. Data within the Data Hub is stored on secured AWS file storage, that is only accessible by the Data Hub servers. Users must be pre-approved by the Project PI for access to a private project and must use their authentication credentials to access project-specific data. Only datasets that have been marked public are available for view and download by others outside of the consortium. The security for the platform is scalable and capable of handling everything from low level to proprietary or embargoed data.

B. Data Accessibility

A backbone of modern scientific discovery is sound data management practices with a chief goal of free and easy access to data, following the FAIR (Findability, Accessibility, Interoperability, and Reliability) guiding principles for scientific data stewardship [4][21]. The EMN Data Hubs help users adhere to the best practices for data science. When possible, data should be non-proprietary, unencrypted, uncompressed, and easily interoperable by machines and humans. Typically, this means simple ASCII or UTF-8 encoding and CSV format for datasets, or simple text (.txt) format for all other relevant information. The Data Hubs provides additional support to the users with resources for applying best data management practices including:

- Documentation of metadata standards.
- Data source generation information for the discovery, reuse, and citation of scientific data.
- Consistent and reusable procedures for data release.
- User tutorials and workflow documentation.

These resources are available within the public domain on the Data Hubs. The platform enables programmatic accessibility to the data via an API and each registered user is provided an API key. This API key can be used within scripts to programmatically query and download any data to which the user has access.

V. METADATA FOR MATERIAL SCIENCE DATASETS

Metadata is the basic information that defines the criteria by which data was acquired or computed (e.g., the 2 θ setting of an X-Ray Diffraction instrument), thereby providing context to the underlying data. Experimental metadata can be elusive to capture, most often residing in researcher notebooks. Experience has demonstrated that many instrumentation manufacturers either do not capture important metadata criteria in a data file, or they bury it in proprietary binary formats, accessible only through their analysis platforms. Metadata is an important element in the Data Hubs since it enables the means to verify the accuracy of the data and provide the information needed for reproducibility, discovery, and reusability. Datasets must include all applicable metadata, data dictionaries, schemas, and technical specifications as appropriate. Discipline-specific metadata standards should be used whenever possible [5][6][7].

The EMN Data Teams view metadata as a central component of the Data Hubs, and early in the development process they focused on defining and refining community metadata definitions. This process was done iteratively, in close collaboration with the materials researchers within the EMN and at other institutions [21]. The Data Hubs have been designed to support multiple metadata types and methods, including both pre-defined metadata choices and user-defined tagging methods. This metadata can then be used for querying through both the web interface as a navigation tool, and via the API.

The Data Hub developers enabled collection of metadata in the Data Hubs using multiple methods to increase flexibility in acquiring this important information. One method utilizes predefined interactive CKAN pages with dynamic forms which are filled out by researchers when they create a new dataset or resource. This same process can be performed through the API, where a programmer can provide the predefined metadata as part of an API call. In another method, metadata templates are defined for specific materials data types, and users upload a completed template file when they add new data. This method has its advantages in that new metadata for a given dataset is within a single file and is quickly accessible. Additionally, a metadata plugin for CKAN was developed, which defines both dataset-level and resource-level metadata in JSON formatted files and used as part of the deployment configuration of CKAN. These JSON files are loaded when CKAN is started, and the web front-end uses these configuration files to create a dynamic user-input form. This method is useful in that metadata is archived as part of the internal key-value store of CKAN and can be referenced and searched directly in API calls.

VI. THE EMN DATA HUBS

The following section highlights examples of the EMN consortia and the materials research they focus on. Each EMN has produced unique and valuable datasets, and most have also developed custom data tools and models to explore and understand the underlying materials phenomena within their datasets. As part of the overarching design of the EMN Data Hubs, these tools have been engineered to be reusable across

most, if not all, the Data Hubs, enabling each of the consortia to leverage this tool development work for their own data.

A. The DuraMAT Consortium

DuraMAT is an EMN virtual laboratory focused on studying and improving the materials, reliability, and manufacturing of PV modules. This includes everything from the PV module framework to solder to encapsulants. The types of projects can include lab experimental studies, long term time-series studies of modules in the field and under accelerated testing and designing new data tools to explore and improve the understanding of the PV degradation modes. The projects in DuraMAT follow several different workflows and utilize both automated data harvesting and upload through the Hub's native API and web-based manual processes. The Data Hub is very adept at supporting a variety of experimental methods and workflows including key data management solutions that provide the dissemination of critical raw and processed time-series data to support internal collaborations and public research.

A critical DuraMAT project involves looking at degradation effects seen in PV modules, but these processes can take months or years in the field to see the cumulative damage (see Fig. 4). It was important for the Data Hub to be able to archive large, accelerated testing, time-series datasets for internal project members and to easily disseminate the information to outside researchers, stakeholders, and the public. Not only was the gathered data important, but the accelerated testing process and instrumentation would provide significant advancements in the field. With increased manufacturing levels of PV, combined with new technologies, packaging, and production methods; understanding long term degradation effects can be critical in determining PV and system component reliability. This understanding is pivotal to all stakeholders in the PV supply chain and a small savings at the PV system level could mean billions of dollars in savings overall [9][10]. By combining many of the stress factors as they occur in the natural environment, rather than sequentially, Combined-Accelerated System Testing (C-AST) can accurately provoke failure modes seen in the natural environment, reduce test time, and number of samples in parallel test. It can also help avoid costly over design by minimizing test failures not seen in the field [8]. Experiments in the C-AST system can accurately reproduce stress failures seen in the field over a longer time. Notable failure effects seen during the process were solder bonds, light-induced degradation, backsheet cracking, corrosion, and cracking of photovoltaic cells [9].



Fig. 4. Example Images [13]. A Few of the known Degradation Modes Found in PV Modules in the Field. These Modes and others are Targeted for Reproduction by C-AST Protocols. From Left to Right: Corrosion, Polyamide Backsheet Cracking, Edge Seal Failure, Delamination, Snail Trails.

During operation, the C-AST system produces a great deal of time-series data containing environmental chamber recipe data, chamber monitoring data, module monitoring data, current-voltage (IV) measurements, and electroluminescence (EL) imaging. Both the raw and processed data is extracted from the C-AST databases using automated API processes and is hosted within the DuraMAT Data Hub.

Most analysis of environmental effects on PV modules and many other physical processes have relied on the Koppen-Geiger (KG) classification of climates zones. The KG classifications are based on seasonal variations in precipitation and temperature and do not target all the climate factors that can possibly affect PV degradation. This makes using the KG model directly more difficult to ascertain true impacts of climate on PV modules. A key DuraMAT study was to re-examine the environmental climate factors and tune the models to be looking at those various climate factors known as contributors of PV degradation and rebuild the model and map to help advice improvements in PV array development, deployment, and costs (see Fig. 5).

The specific stressors used in this new climate model are based on mean module temperature, mean module temperature rate of change, extreme low ambient temperature, specific humidity (relative to damp heat), UV exposure, and wind loading [10]. The model and associated maps now provide a better insight into expected performance characteristics and expected failure and degradation modes in geographic zones. All the data and maps are available via the Data Hub to the public. Researchers can use an interactive web app, supplied through the DuraMAT Data Hub, to enter latitude and longitude coordinates to see PV climate stressors for any location in the US.

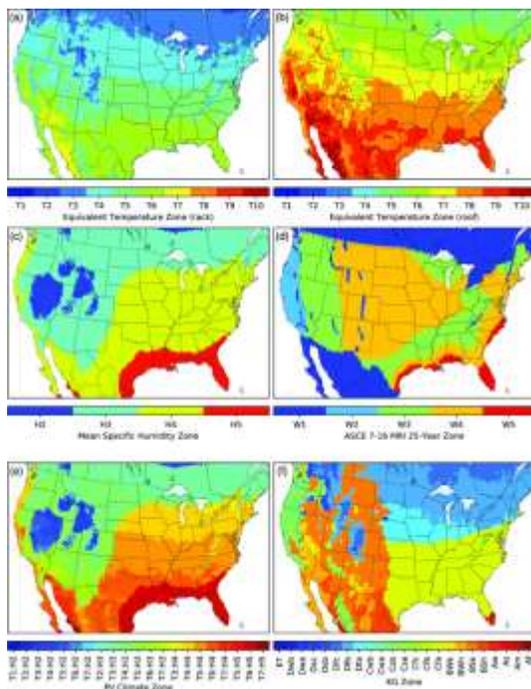


Fig. 5. Photovoltaic Stressor Climate Zones. (a) Temperature Zones for Continental United States (b) Specific Humidity Zones. (c) Wind Zones. (d) Combined Temperature and Humidity Zones. (e) Comparative Koppen-Geiger Temperature Zone [12].

B. The HydroGEN Consortium

The Hydrogen GENeration (HydroGEN) consortium is focused on accelerating the foundational R&D of innovative materials for multiple advanced water-splitting technologies to enable clean, sustainable, and low-cost hydrogen production. The consortium uses a collaborative, multi-lab, and theory-guided R&D approach to address the critical research gaps identified for the following technology readiness level (TRL) for water splitting pathways:

- low temperature electrolysis (LTE), including polymer electrolyte membrane (PEM) and anion-exchange membrane (AEM) electrolysis,
- high temperature electrolysis (HTE), including oxygen-conducting (o--SOEC) and proton-conducting solid oxide (p-SOEC) water electrolysis,
- solar thermochemical (STCH).
- photoelectrochemical (PEC) water splitting hydrogen production.

Within HydroGEN, a host of collaborative experimental, benchmarking, and round-robin data is being produced and archived within the Data Hub. Experimental data comprise materials characterization (e.g., x-ray diffraction, x-ray photoelectron spectroscopy, Raman spectroscopy, microscopy, stagnation flow reactor), device performance (e.g., voltage-current electrolysis, PEC photocurrent-voltage, solar-to-hydrogen efficiency), and materials durability (e.g., thermal gravimetric analysis, membrane conductivity, current or voltage vs. time).

In addition to the experimental data, there is structural modelling data that provides critical insights to a STCH water splitting material. $BaCe_{0.25}Mn_{0.75}O_3$ (BCM) material is of interest for solar thermochemical Hydrogen (STCH) generation. The BCM polytype structures from density functional theory (DFT) data is hosted on the Data Hub (<https://datahub.h2awsm.org/dataset/metadata/bcm-structures>) and provides explicit structure models for both the 12R (ground state) and 10H (metastable at ambient temperature) polytypes of BCM. These explicit structure models can be used for further electronic structure calculations and comparisons with experimental data [11].

The first principles materials theory (FPMT) for advanced water splitting pathways, a computational modeling capability (hosted by NREL), which produced these BCM data, supports the mission of the consortium by providing computational materials data for STCH water splitting. Access to the data and this modeling capability on the Data Hub, will allow researchers to accelerate the selection of suitable materials for synthesis, characterization, and optimization for water splitting.

The HydroGEN Data Hub also has public plenary presentations and breakout summaries from the three annual Water Splitting Technologies Benchmarking and Protocols Workshops. This is a national community effort with international engagement and participation in all four advanced water-splitting technologies. Presentations, discussion summaries, and action items obtained from the breakout groups

(e.g., high-level roadmaps for each advanced water-splitting technology, technoeconomic analysis, and protocols) are currently publicly available on the Data Hub (<https://datahub.h2awsm.org/dataset/2021-water-splitting-technologies-benchmarking-and-protocols-workshop>) (<https://datahub.h2awsm.org/dataset/2019-water-splitting-technologies-benchmarking-and-protocols-workshop>), (<https://datahub.h2awsm.org/dataset/2018-water-splitting-technologies-benchmarking-and-protocols-workshop>).

The 36 developed test protocols are still being finalized and will be published on the Data Hub soon. The development of standard test protocols and benchmarking of advanced water-splitting materials are one of HydroGEN's most important consortium cross-cutting activities and is critical to accelerate materials discovery and development.

As more experimental, computational, and benchmarking data are added to the data hub, along with the metadata that are being developed for the different water splitting technologies and the tools that help with the batch uploading of datafiles and data visualization, the HydroGEN Data Hub is an invaluable, central, searchable advanced water-splitting materials data source for the entire advanced water-splitting hydrogen production community.

C. The HyMARC Consortium

The Hydrogen Materials Advanced Research Consortium (HyMARC) addresses the scientific gaps to advancing solid-

state hydrogen storage materials. HyMARC focuses on the basic understanding, synthesis development, protocols, characterization tools, and validated computational models to accelerate exploration and discovery of materials that meet industry requirements for hydrogen storage, carriers for distribution of hydrogen from production to storage, or user facilities.

The HyMARC Sorbent machine learning (ML) application is a great example of a standalone data tool developed for the consortium and accessible through the Data Hub (see Fig. 6). In some cases, a standalone tool is the better solution than embedding the tool within the Data Hub framework. Internal linkages to the Data Hub data make it work seamlessly and make it easily accessible to consortium researchers and the public. The Sorbent web application interacts with a ML model and is centrally hosted with the model, code, and data storage, to help accelerate discovery of optimal hydride design (<https://sorbent-ml.hymarc.org/>). The public sorbent ML application can be extended to host high throughput runs of the model and or interactive dataset models; to provide a pre-computed visualization with varying input parameters.

Additional details concerning the input values to these ML codes could be needed by the researchers and public and can be explored within the HyMARC Data Hub projects. The work (<https://datahub.hymarc.org/dataset/machine-learning-ready-metal-hydrides>) provides any needed information on the seeding data [17] supporting the interactive tool.

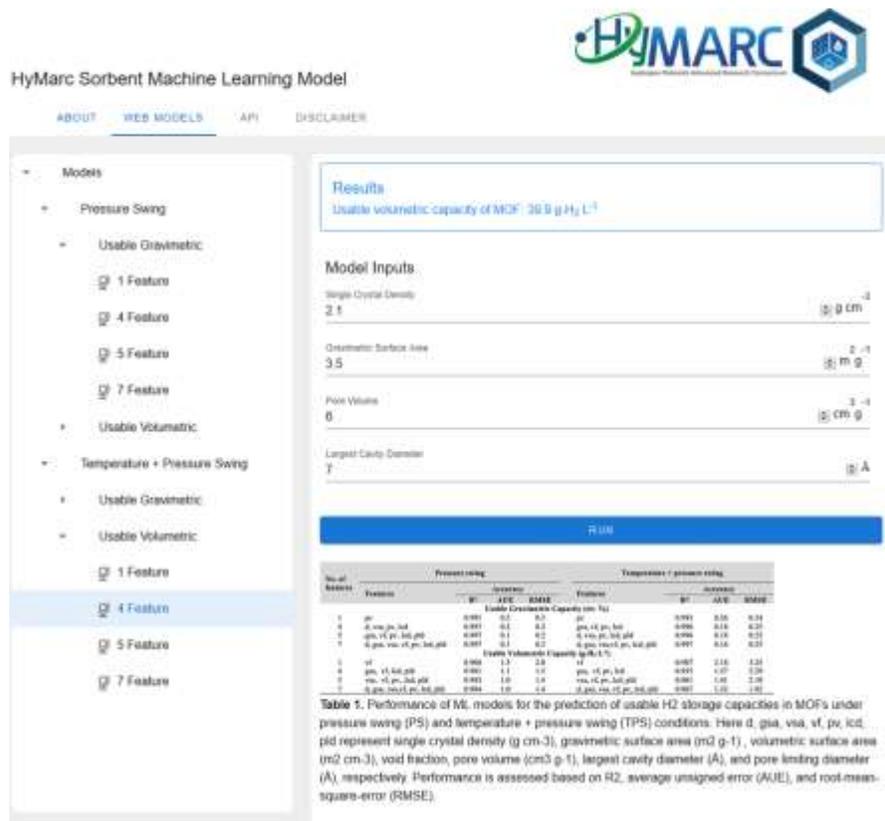


Fig. 6. An Example of the HyMARC Data Hub's Ancillary Application Capability. Built using Modern Web Technologies and Hosted on a Serverless Cloud Platform, a Public Machine Learning Application for Prediction of Usable H₂ Storage Capacities of MOFs Varying Input Parameters.. Interactive Dataset Models could Provide a Pre-Computed Visualization with

D. The ElectroCat Consortium

The ElectroCat (Electrocatalysis) consortium focuses on fuel cell energy conversion devices by accelerating the development and deployment of non-platinum group metal (PGM-free) catalysts. With a systematic approach using high-throughput, combinatorial methods, potential catalysts can be synthesized and analyzed rapidly and comprehensively. This results in an accelerated development of PGM-free catalysts, reducing fuel cell costs, by utilizing more abundant materials, thereby increasing U.S. competitiveness in manufacturing fuel cell electric vehicles (FCEVs) and other applications. The consortium has pooled electrocatalysis knowledge from several national laboratories, with an overarching goal of refining and streamlining the hardware and software tools necessary to model, analyze, and optimize PGM-free catalyst and electrode structure and performance. These tools have become enduring capabilities and will grow the publicly available dataset as a resource on the ElectroCat the Data Hub.

Of particular importance is the Unmix X-Ray Diffraction (XRD) data tool, which is a prime example of a custom CKAN embedded tool. This tool gives researchers the ability to perform component analysis on experimental x-ray diffraction spectra to estimate the contributions of each of a set of reference spectra, thereby determining a percentage distribution of each reference material within the experimental sample. Unmix XRD provides a form of automated Rietveld analysis for thin-film and powder x-ray diffraction spectroscopy. By providing a series of reference patterns in addition to an experimental spectrum, Unmix XRD determines the percentage of each reference pattern present within the spectrum. Reference patterns may be provided by the user or determined automatically from a diverse selection of the National Institute of Standards and Technology Inorganic crystal structure database of powder diffraction spectra (see Fig. 7).

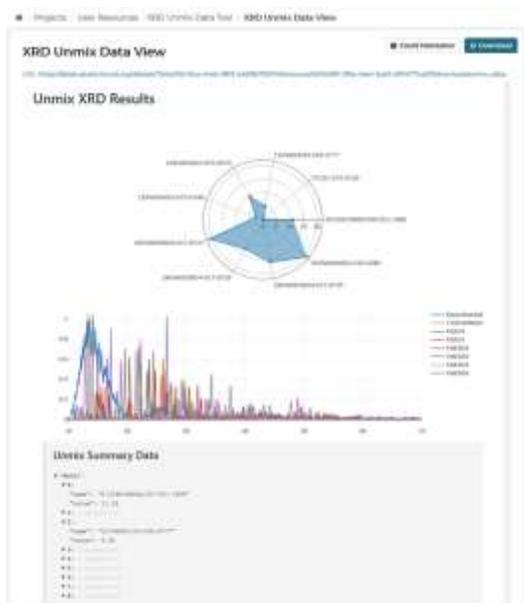


Fig. 7. The Unmix XRD Tool, which Enables the Researcher to Analyze XRD Data and Estimate the Relative Contributions of a Set of Reference Material Spectra. The Result is an Estimate of the Percentage of each Reference Material in the Experimental Sample.

This tool takes input from a reference spectra set, and any number of XRD experiment result files, performs non-negative least squares fit, and outputs a heatmap showing each "sample" (experimental file) and how much of each of the references were present within each sample. The second output file is a csv file of the same data. The Unmix XRD data tool is designed and built with modern front-end web application technologies and may be decoupled from the CKAN framework. In addition to the web application, Unmix XRD runs a lightweight compute process on the Data Hub servers, leveraging AWS Cloud resources whenever possible.

E. The ChemCatBio Consortium

The Chemical Catalysis for Bioenergy (ChemCatBio) Consortium aims to accelerate the development of catalysts and processes for next-generation biomass conversion processes. Through faster maturation of catalytic technologies, ChemCatBio ultimately intends to speed a transition to a circular economy for fuels and chemical products. It accomplishes this mission through a range of core catalytic technology projects including catalytic upgrading of biochemical process intermediates, catalytic fast pyrolysis, upgrading of one- and two-carbon intermediates (e.g., methanol and ethanol), and CO₂ utilization. These projects are accompanied by a set of enabling capabilities such as synthesis and characterization, cost analysis, deactivation studies, and theoretical modeling. ChemCatBio includes researchers from six DOE national labs, along with a network of advisors and research partners from industry and academia. In order to accomplish its mission with a geographically distributed team and large datasets, structured data management is critical. The ChemCatBio Data Hub offers this structured data-sharing resource, as well as a platform for specific data and tools developed by ChemCatBio.

Two ChemCatBio tools that use the Data Hub platform include CatCost™ and the Catalyst Property Database. CatCost is a free, public estimation tool designed to help catalyst researchers understand the cost of producing their materials at industrial scale. It was developed by ChemCatBio and released in 2018 (<https://catcost.chemcatbio.org>), in both spreadsheet and web app versions. Both versions of CatCost use the Data Hub extensively, including to store the version history for the spreadsheet downloads and the web app's libraries of materials, equipment, and spent catalyst information. CatCost development has continued since its initial release, using the resources of the Data Hub to provide continuity and support. The Catalyst Property Database (CPD) is a structured, searchable database of catalyst property information, designed to accelerate literature searching, collaboration, benchmarking, and data science applications in catalyst research, and ultimately to make the process of catalyst discovery faster. It currently contains density functional theory-computed adsorption energies for intermediates on catalytic surfaces. An initial release of the CPD was published in 2020 (<https://cpd.chemcatbio.org>) and it is under active development to allow external user uploads, establish curation procedures, offer user training, and add user-requested features. As CPD development progresses, data structures offered by the Data Hub guides development and helps ensure reliability.

VII. CONCLUSIONS

The need to meet the rapid changing landscape of scientific research and the ability to handle data efficiently and effectively is becoming more critical each year and has created a need for developers to create these collaborative data hubs to support more than just the searching and sharing of publications [21]. A simple internet search will reveal most data hubs are designed to support searchable publication repositories. Some focus on data archives, modeling and simulation, or are focused on education [17][19]. Other hub platforms are engineered to provide informatics, Machine Learning (ML) and Artificial Intelligence (AI) analysis resources, and reporting. Research teams working on pathogen surveillance are designing and deploying similar systems to the EMN data hubs, sharing the concepts of open data, structured sharing, extensibility, discovery, and data standards [18] all designed to meet research needs and follow FAIR principles of data management and discovery [4]. But what makes the EMN data hubs unique is that they can provide those same capabilities, but with a strategic focus to support ongoing experimental data sharing, secured experimental projects, analysis, and research team communication making them more analogous to a virtual laboratory.

The EMN Data Hubs have been in operation for nearly five years. Each has successfully been able to provide their community with a platform to support the acquisition and dissemination of project data, securely within project teams, and to the public. There are more users, data, visualization tools, and data analytic tools coming online in the Data Hubs every month. The EMN Data Hubs have demonstrated that the cloud is the most suitable platform to provide needed accessibility, disseminate the unique capabilities of a virtual laboratory, and make available a myriad of cloud services and workloads to the community. From web applications and API hosted on the Data Hub infrastructure, to cloud native workloads tailored to specific analytics, the platform on which these Data Hubs are built ultimately enables their EMN to host robust tools for open, reproducible, and on-demand compute workloads for data science, machine learning, and visualization.

With resources and time often at a premium for research organizations, a centralized virtual laboratory like an EMN demonstrates how to maximize the potential of available instrumentation and expertise across a dispersed set of resources. The Data Hub is the binding element of the EMNs, facilitating easy, yet secure communications and data sharing for all the projects and consortium members. While each of the Data Hubs contain elements that are unique to them and the supported EMN, the baseline architecture of the platform enables software re-use and efficient development processes across all the consortia. With a careful eye to modularity and templated deployment, a Data Hub built on these ideas can be quickly deployed and be highly scalable and agile to meet the needs of a distributed research team. As recent world events have shown, there is sometimes a need to avoid travel or direct interaction, and a research consortium built around a virtual laboratory Data Hub can still allow the researchers to continue to perform well; saving time, money, and possibly lives.

VIII. FUTURE WORK

The Data Hubs are always in active development and continue to grow in capability and data resources. Several development thrusts are underway to improve the Data Hub operability and improve the overall user experience. The importance of metadata capture cannot be understated but neither of the current two methods to upload metadata, web user interface (UI) and API, provide a robust validation process (see Fig. 8). Development of a metadata service layer is in progress that will enable more robust metadata curation and validation. This metadata-as-a-service will integrate with the existing CKAN architecture. This service will provide agility to modify existing metadata, reduce developer workload, and easily add new metadata as capabilities change and grow in the future, while still meeting guidelines for community standards [21].

As the Data Hub data repository features have matured, how data tools are developed and hosted is becoming more of a focus of the EMN data teams. Throughout the development of the data repository, several visualization and data analysis tools were developed to showcase common needs as defined within the EMN scientific domain (see above). Some of these existing tools, by design, are integrated within the CKAN framework, however, all of them could be decoupled from CKAN and hosted in the cloud, as independently accessible, interoperable services. Additionally, while open-source software development is a viable method for community sharing and development, it does not guarantee that the code will remain practical or useable (i.e. easy to install and run) for researchers. A more appropriate solution would be to not only develop a tool as open-source research software, but to host it as a service in the cloud and enable other researchers to run the software on-demand with their choice of dataset.

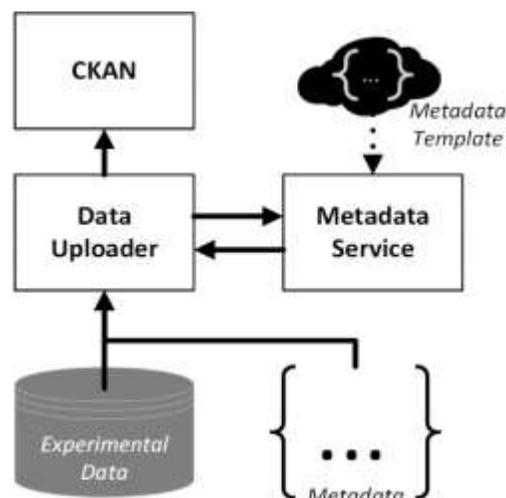


Fig. 8. PI/Leads May Define Custom Templates Per Experiment to Ensure Consistency and Provide Flexibility, Enabling the Curated Metadata Upload for each Project or Capability/Node.

The work being done, and the data being generated in many of the consortia is helping create extensive libraries of materials characteristics. In some cases, this data is being passed into materials science databases to allow for accelerated

discovery through machine learning and deep learning techniques [14], but there is still more to be done. It is hoped that future work can begin to either build additional large scale material databases for experimental data or merge these libraries into existing systems such as the Materials Project [16]. On a limited scale, some of the Data Hubs are currently leveraging a custom Laboratory Information Management Systems (LIMS) to help automatically assure the gathering and consolidation of all the metadata and data products and it is expected that to be extended to additional consortia partners and institutions soon [15].

The developers are constantly working with the consortium researchers and data team members to look at improvements of the general CKAN UI that can provide a better work environment and experience. There are several issues with project hierarchy and data association that while functional, are not as intuitive to use as researchers would like, and the developers are examining ways to improve them during the next rounds of Data Hub development.

ACKNOWLEDGMENT

The Energy Material Network Data Hubs are a large-scale project across several national laboratories and institutions. Much of the work in design, development, deployment, and operations is handled by the members of the EMN data teams, who we wish to thank for their hard work, insight, and dedication.

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Hydrogen and Fuel Cell Technologies Office. Funding provided as part of the Durable Modules Materials Consortium (DuraMAT), an Energy Materials Network Consortium funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, and Solar Energy Technologies Office agreement number 302509. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

- [1] J. R. Hattrick-Simpers, A. Zakutayev, S. C. Barron, Z. T. Trautt, N. Nguyen, K. Choudhary, et al., "An inter-laboratory study of Zn-Sn-Ti-O thin films using high-throughput experimental methods," *ACS Comb. Sci.*, 21(5), pp. 350-361, March 2019. doi: 10.1021/acombisci.8b00158.
- [2] M. L. Green, C. L. Choi, J. R. Hattrick-Simpers, A. M. Joshi, I. Takeuchi, S. C. Barron, et al., "Fulfilling the promise of the materials genome initiative with high-throughput experimental methodologies," *Applied Physics Reviews* vol. 4, 011105, pp. 1-18, March 2017. doi: 10.1063/1.4977487.
- [3] R. P. Vargas, A. C. Caminero, D. Sanchez, R. Hernandez, S. Ros, A. Robles-Gomez, L. Tobarra, "Laboratories as a service: using cloud technologies in the field of education," *Journal of Universal Computer Science*, vol. 19, no. 14, pp. 2112-2126, 2013. doi: 10.3217/jucs-019-14-2112.
- [4] M. Wilkinson, M. Dumontier, I. Aalbersberg, G. Appleton, M. Axon, A. Baak, et al., "The FAIR guiding principles for scientific data management and stewardship," *Sci Data* 3,160018(2016). doi: 10.1038/sdata.2016.18.
- [5] Library of Congress, "Recommended formats: datasets/databases," *Library of Congress, Preservation* Feb 21, 2020. http://www.loc.gov/preservation/resources/rfs/data.html#datasets_
- [6] Duke University Libraries, "Research data management," *Research Data Management*, Feb 21, 2020. https://guides.library.duke.edu/c.php?g=633433&p=4429351_
- [7] Stanford University Data Management Services, "Best practices for file formats," Feb 21, 2020. https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-formats_
- [8] S. L. Moffitt, R. A. Fleming, C. S. Thompson, C. J. Titus, E. Kim, L. Leu, et al., "Advanced X-ray scattering and spectroscopy characterization of an antisoiling coating for Solar module glass," *ACS Appl. Energy Mater.*, vol. 2, no. 11, pp. 7870-7878, October 2019. doi: 10.1021/acsaem.9b01316.
- [9] P. Hacke, M. Owen-Bellini, M. Kempe, D. C. Miller, T. Tanahashi, K. Sakurai, et al., "Combined and sequential accelerated stress testing for derisking photovoltaic modules," *Advanced Micro- and Nanomaterials for Photovoltaics*, 2019, pp. 279-313. doi: 10.1016/B978-0-12-814501-2.00011-6.
- [10] S. V. Spataru, P. Hacke, M. Owen-Bellini, "Combined-accelerated stress testing systems for photovoltaic modules," *IEEE 7th World Conference on Photovoltaic Energy Conversion*, June 2018. doi: 10.1109/PVSC.2018.8547335.
- [11] S. Lany, "BCM polytype structures from DFT – data and resources," *Energy Materials Network HydroGEN*, June 2019. doi:10.17025/1532370.
- [12] C. B. Jones, T. Karin, A. Jain, W. B. Hobbs, C. Libby, "Geographic assessment of photovoltaic module environmental degradation stressors," *IEEE 46th Photovoltaic Specialists Conference*, 2019. doi: 10.1109/PVSC40753.2019.8980741.
- [13] P. Hacke, "Development of combined and sequential accelerated stress testing for derisking photovoltaic modules," *DuraMAT Webinar presentation*, May 2019. <https://www.nrel.gov/docs/fy19osti/73984.pdf>.
- [14] A. Zakutayev, N. Wunder, M. Schwarting, J. D. Perkins, R. R. White, K. Munch, et al., "An open experimental database for exploring inorganic materials," *Sci. Data* 5, 180053, April 2018. doi: 10.1038/sdata.2018.53.
- [15] R. R. White, K. Munch, "Handling large and complex data in a photovoltaic research institution using a custom laboratory information management system," *MRS Proceedings*, 1654, 1104, March 2014. doi: 10.1557/opl.2014.31.
- [16] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, "The materials project: a materials genome approach to accelerating materials innovation," *APL Materials* 1, 011002, July 2013. doi:10.1063/1.4812323.
- [17] M. Witman, S. Ling, D.M. Grant, G. S. Walker, S. Agarwal, V. Stavila, M. D. Allendorf, "Extracting an empirical intermetallic hydride design principle from limited data via interpretable machine learning," *J. Phys. Chem. Lett.* 11, pp. 40-47, December 2019. doi: 10.1021/acs.jpcclett.9b02971.
- [18] C. Amid, N. Pakseresht, N. Silvester, S. Jayathilaka, O. Lund, L. Dynovski, et al., "The COMPARE data hubs", *Database*, vol. 2019, pp. 1-14, December 2019. doi: 10.1093/database/baz136.
- [19] T. A. Faultens, P. A. Bermel, A. Buckles, K. A. Douglas, and A. H. Strachan, "nanoHub.org: A gateway to undergraduate simulation-based research in materials science," *MRS Proceedings*, 1762, pp. 7-14, February 2015. doi: 10.1557/opl.2015.80.
- [20] D. Brickley, M. Burgess, and N. Noy, "Google dataset search: building a search engine for datasets in an open web ecosystem," in *The World Wide Web Conference*, 2019, pp. 1365-1375. doi: 10.1145/3308558.3313685.
- [21] A. Dima, S. Bhaskarla, C. Becker, M. Brady, C. Campbell, P. Dessauw, R. Hanish, et al., "Informatics infrastructure for the materials genome initiative," *JOM*, vol. 68, pp 2053-2064. doi: 10.1007/s11837-016-2000-4.

A New Approach based on Fuzzy MADM for Enhancing Infrastructure Selection in the Case of VANET Network

Abdeslam Houari¹, Tomader Mazri²

Laboratory of Advanced Systems Engineering, Ibn Tofail Science University
Kenitra, 14020, Morocco

Abstract—The emergence of the 5G mobile network has a huge impact on the evolution of services and functionalities offered to its customers; this latest version of mobile networks will allow the simultaneous connection of a significant number of people and IoT devices, in addition to the improvement of several other features. 5G will serve in a large part of smart cities and especially in the field of intelligent transportation systems (ITS). Vehicular Ad-Hoc Network (VANET) is one of the promising projects on which the ITS is relying on. Its main purpose is to provide communication and information-sharing support for the vehicles in its network. VANET is based on a heterogeneous network architecture composed mainly of two infrastructures, the first one is the cellular infrastructure, and the second is the road infrastructure. This paper proposes a new approach based on Fuzzy multiple attribute decision making (MADM) methods for the selection of the most appropriate infrastructure in the VANET network and consequently enhance the number of executed vertical handover to move from one infrastructure to another without loss of connection.

Keywords—5G; Vehicular adhoc network; internet of things; intelligent transportation system; cellular infrastructure; road infrastructure; fuzzy; vertical handover; MADM

I. INTRODUCTION

Nowadays, improving the vehicle's autonomy level has become one of the major concerns of the automotive sector. This new orientation is dictated by the ever-increasing number of accidents and traffic jams, which have irreversible social, ecological, and economic consequences. For this reason, the automotive sector is gradually turning to intelligent transportation systems (ITS).

VANET (vehicular ad-hoc network) is an important project related to ITS, a specific case of MANET (Mobile Ad-hoc Networks) network in which nodes are smart vehicles. The idea behind it is to connect the vehicles and share data and resources between them to understand their surroundings better and deal with the unexpected events that may occur on the road, such as traffic jams, which are one of the biggest issues for road users. Two main communication modes are provided; vehicles can communicate with each other via the V2V (vehicle-to-vehicle) mode using devices such as AU (application unit) or OBUs (On-Board Units), as they can communicate with the road infrastructure using the V2I (vehicle-to-infrastructure) mode which allows them to communicate with RSUs and Base station installed along the

road to benefit from several services such as internet access. The high node mobility is one of the main characteristics of VANET networks caused by the random movement and high speed of the vehicles, which directly impacts the response time and the topology of the network.

The emergence of 5G mobile networks is highly expected by smart cities and especially intelligent transportation. This new generation of mobile networks is equipped with a set of advanced features, such as higher capacity in comparison to 4G, allowing device-to-device communication and providing a better density of mobile broadband users [1]. As a result, vehicles moving in such a heterogeneous network will benefit from various services offered by both infrastructures. Therefore, the network's right choice is an important step to improve the quality of services (QoS) and the vehicles' efficiency. The heterogeneous nature of the VANET network raises the challenge of service continuity; as a result, the performance of this type of network requires the effective use of a vertical handover process to know the most appropriate infrastructure according to predefined criteria. Generally, the selection process can be composed of three phases: data collection, decision transfer, and transfer execution after conducting a comparative study in the previous work presented at the ISAECT conference [2]. In this paper, we focus on the decision-making step. The decision-making in the VHO process depends on several QoS network parameters and criteria such as network state, capacity, reliability, and others. In order to solve the problem of the multitude of criteria that increases with each network upgrade, the multiple attribute decision-making (MADM) approach is proposed in this study. MADM is considered a promising solution to the network and infrastructure selection problem. It is simple to implement and does not require any particular physical resources. The best-known algorithms of the MADM are the hierarchical analysis algorithms such as AHP (analytic hierarchy process) or the analytic network process (ANP) used in the criteria evaluation and weighting step, as well as VIKOR Saw and TOPSIS for ranking the available choices (networks or infrastructures). In the first part of this article, a presentation of the VANET network is performed, as well as its different characteristics, the modes of communication, the main components as well as the main challenges it faces. The next section describes the cellular infrastructure in the 5G era while detailing the main technologies that characterize it compared to previous generations of mobile networks. The following section is

dedicated to the proposed approach, which explains the implemented system model and the applied methods for ranking the two infrastructures. The penultimate part concerns the simulations performed as well as the discussion of the results obtained. Finally, the article ends with a conclusion.

II. VANET NETWORK

As explained earlier, VANET is a form of MANET network, enabling data exchange communication between vehicles in a specific geographical area and between vehicles and equipment installed at the roadside, known as roadside equipment. Several industries are interested in this project, especially the transport sector, mainly for the safety aspect that it promises to ensure by reducing the number of accidents and traffic jams, positively impacting the ecology by reducing harmful gas emissions from road traffic. In this section, a focus on the different modes of communication supported by the VANET network is highlighted. After that, the main equipment on which it is based are listed; later, we explain the data processing cycle. We end with the main features and challenges it faces.

A. Communications Mode

In a VANET network (Fig. 1), vehicles communicate with each other through the vehicle-to-vehicle mode and with road equipment through the V2I (vehicle-to-infrastructure) mode; this will allow drivers and authorities to benefit instantaneously from different information on road conditions and to take advantage of several other services offered by the network (entertainment applications, games, localization...) allowing drivers and passengers to travel safely and comfortably. There is also another type of communication, which not only interconnects vehicles and infrastructure but also includes cellular networks and other devices, this new mode called Vehicle-to-everything (V2X) is composed of V2V, V2I modes in addition to V2P (Vehicle-to-Pedestrian), and Vehicle-to-Device (V2D).

Furthermore, VANET networks are essentially based on two types of applications, one constituting the core of the ITS, which ensures the improvement of road safety, and the second type of application for the comfort of the driver and his passengers[3].

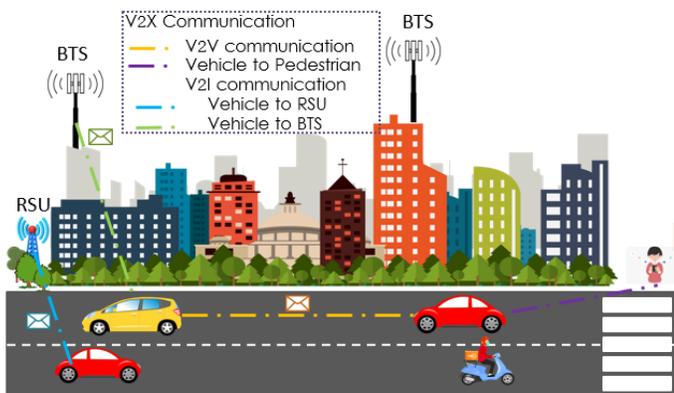


Fig. 1. Communication Modes.

B. Network Architecture Component

The data exchange at the network layer is handled by DSRC/WAVE wireless medium, whose role is to secure data transfer and reduce latency. The resulting communication leads to more reliable information and allows the deployment of safety applications that enhance travelers' safety and make driving more pleasant. Among the main components of the VANET network, there are:

- OBU: Commonly used with the RSU or other OBUs, this device is placed on the vehicle's board. Its role is to share information with the other vehicles. It is also used to retrieve information from the sensors installed in the vehicle. The connection with the RSU and other OBU is established through an IEEE 802.11p wireless link[4].
- AU: The Application unit exploits the communication means provided by the OBU in order to communicate with the network[4]. It is used in various applications; it can be used for security applications and internet access devices such as PDA (personal digital assistants).
- RSU: the road side-unit is a wave device used to extend the range of communication. It is also used to send warning messages on the road to prevent accidents[4].
- Base Station: It is equipment whose role is to connect mobile devices to access a telecommunication network; it comprises a transmission antenna and a reception antenna.

The VANET network vehicles are equipped with a set of sensors enabling them to discover their surroundings better in order to make the right decisions at the right moment. To do this, the vehicle starts by collecting data from its surroundings (location, road conditions, accident, etc.) via the sensors and cameras installed onboard. These data are then stored and processed to retain only the most important ones. Finally, the process ends by transmitting this data to the recipient vehicle belonging to the same network or a completely different network type (Fig. 2).



Fig. 2. Data Collection Process.

C. VANET Features

As mentioned above, VANET network is a specific case of a MANET network with features that differentiate it from other mobile networks, such as:

- **Dynamic topology:** this specificity comes from the fact that the vehicles are constantly changing their position and driving at high speed. They can also move in both directions and join or leave the network at any moment.
- **Frequent disconnection:** this is one of the direct consequences of the network's dynamic topology, making the connection unstable. The link between vehicles can easily be broken, resulting in the loss of the data exchanged.
- **Mobility and prediction:** Although the network's topology is dynamic, vehicles normally follow a mobility model that considers traffic lights, speed limits, road conditions, and driver behavior. However, the vehicle trajectory cannot be predicted in advance[5].
- **Geographical communication:** communication between a source and destination vehicles is possible only if the latter can be reached, a method quite different from those used in other types of networks, which are based more on the vehicle Id.

These different features have led to new challenges in setting up a safe and reliable exchange system that takes into account the heterogeneous aspect of the network and its infrastructure, without forgetting the particularity of the vehicles considered as important storage resources and powerful processing units.

D. Challenges in VANET Network

VANET faces a wide range of challenges (Fig. 3). The most critical ones are data privacy and security, especially as messages are broadcast. Anyone can intercept them, making the network vulnerable to various attacks that threaten road users' safety [6]. Security is an important aspect of any system. Therefore, it must fulfill certain requirements for the system's proper functioning and the continuity of its services. Failure to comply with any of these requirements may lead to irreversible consequences, such as a system crash or data loss.

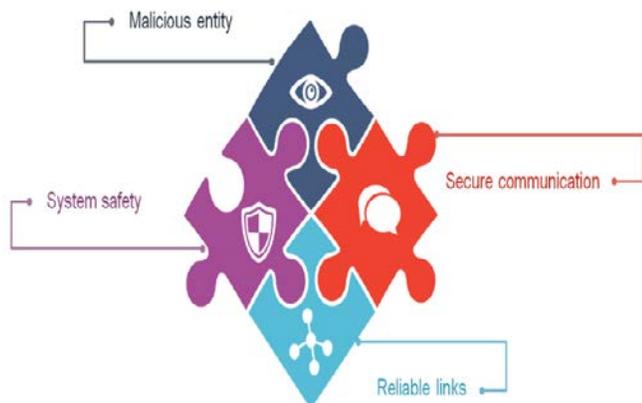


Fig. 3. VANET Challenges.

Moreover, other requirements are exclusive to VANET networks, such as authentication, data integrity, and confidentiality. Authentication in VANET networks aims to monitor the vehicle's authentication levels. Once connected, the vehicle has a unique identifier, which prevents it from passing for multiple vehicles and usurping other vehicles' identities [7]. The second one is integrity; its role is to ensure that the message's data has not been altered or modified fraudulently during its transmission to the receiving node. For confidentiality, its role is to ensure the reliability of the communication. Sending encrypted messages ensures that only the destination vehicle is authorized to consult their content[7]. The security criterion aims to secure communications between vehicles on the one hand and between vehicles and road facilities on the other hand. One of the most dangerous attacks that vehicles face is the Sybil Attack, which is the subject of several research and studies. one of the recommended ways to prevent this attack is to establish a balance between privacy and non-repudiation[8]. Other dangerous attacks can target vehicles, such as jamming and DOS attacks, which require further research to discover the appropriate security mechanisms and algorithms to protect against them.

In some cases, the communication in an open space leads to the infiltration of malicious nodes that can intercept and corrupt the network's data. One of the major challenges that VANET faces is preventing attackers from accessing information to preserve the reliability of the links set up between vehicles and the loss of data packets on the network. The author in [9] have implemented the Reputation-Based Global Trust Establishment (RGTE) scheme to achieve these objectives. The method adopted is based on statistical laws to circulate only trustworthy information. It also allows the detection of malicious nodes through a dynamic mechanism. However, this method cannot be applied in high mobility vehicles due to packet loss.

Finally, the last major challenge for VANET networks is the link reliability between vehicles in V2V and V2I communications used for traffic data collection. Due to the high node mobility, this problem persists until now. One of the proposals is to deploy a system based on IDS (Intrusion Detection Scheme) for detecting malicious nodes involved in illegal activities. In this scheme, each node is linked to a control system in order to monitor each node and check if the network packet is transferred to the next node. Unfortunately, there is a strong probability that the packets are lost and do not reach their destination, and these nodes are consequently considered malicious even if they are not in reality. Note that this type of system has not yet been implemented and is just at the experimentation phase.

III. CELLULAR INFRASTRUCTURE IN THE 5G ERA

Cellular networks appeared at the beginning of the 1990s, its objective is to provide voice service, using a set of base stations installed in the coverage area, each of these stations cover a smaller area exclusively. The addition of traffic and data processing have been introduced into mobile networks from the second generation onwards. Mobile networks have undergone several evolutions, not just in terms of bandwidth but also in terms of standards and technologies. Several

components constitute this network. Among them, there are two that have remained the same throughout the years. The first is the Radio Access Network (RAN), and the second is the Core Network (CN). The RAN is responsible for wireless signal processing (baseband, passband) from the user equipment (UE). Simultaneously, the CN is responsible for the incoming and outgoing traffic flow's reliable routing to their destinations.

A. Main Components of the Network

During the development of mobile networks, several components have evolved many times. The eNodeB (the evolved node B) represents the base station covering the user equipment in a defined area, through which the user equipment can communicate and reach the remote destination. The other component is the Tracking Area (TA), which manages mobility and optimizes user and system management. In practice the area covered is partitioned into several TAs, and each of them includes a set of enodeBs. The Mobility Management Entity (MME) is one of the most important control components that are part of the LTE network; it acts as a signalling node in an EPC control plane. Finally, the Serving Gateway (SGW) component acts as a router and allows the Transfer of the data flow between the BSs and the Packet Data Network Gateway. A typical architecture of a cellular network can be illustrated as follows (Fig. 4).

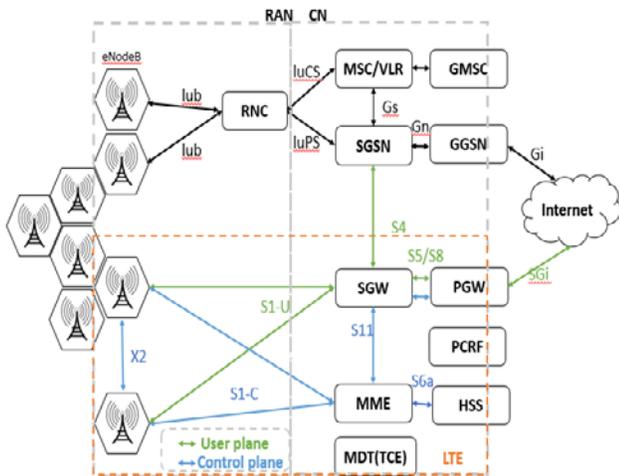


Fig. 4. Cellular Network Architecture.

B. Overview of the 5G

In the case of VANET networks, the data exchange occurs several times and between different actors. The generated data requires special processing, as it contains valuable information on road traffic. Unfortunately, conventional data management systems do not allow the management of a large amount of data in a short period. This new generation of mobile networks promises to solve several problems, including massive data management. Indeed, 5G has started to invade the global market since 2019; several new features will emerge, such as improved response time (estimated at 1ms). 5G will also enable the massive deployment of the Internet of Things (IoT) ecosystem in where networks will satisfy the communication needs of billions of connected devices in a very short time and at a low cost while reducing energy consumption. Several

technologies constitute the pillars of 5G (Fig. 5), enabling it to achieve a level of performance unequaled in the history of mobile networks; the first is the small cell, which is a cellular radio access node with a range from 10 meters to a few kilometers, its role is to increase the data capacity of the network, since, as it is the case with the previous generations, 5G radio waves cannot reach a long-range due to the high frequency used by 5G. The mmWaves (millimeter wave) are localized between 30GHz and 300 GHz; using these waves allows the 5G to reach a high transmission speed. For BeamForming, it is used to steer a radio wave to a specific destination; this is done by tuning the radio waves in order to point in a particular direction. The penultimate technology, which is the Massive mimo, aims to transmit M data flows simultaneously on M transmitting antennas, and each flow is received by N receiving antennas; for this purpose, often several transmitting and receiving antennas are used to improve both the transmission gain and the spectral efficiency. Finally, full duplex is a technology that allows a transceiver to transmit and receive data at the same time, this last technology enables to increase the capacity of the wireless network at the physical layer.

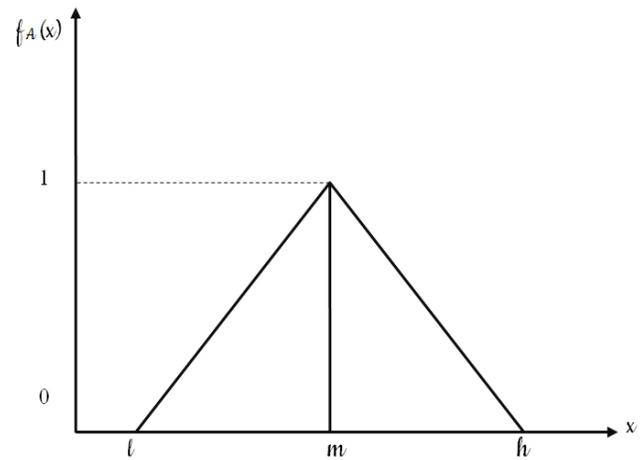


Fig. 5. Triangular Fuzzy Number.

IV. PROPOSED APPROACH TO SELECT THE BEST NETWORK

Roaming between the two infrastructures, road, and cellular is essential for any vehicle belonging to the VANET network qualified as a heterogeneous network. The selection between the two infrastructures is quite an important step. Unlike traditional mobile networks, the nodes' mobility directly impacts the network topology and affects the performance of the used protocols [10]. Also, in urban areas with high towers and a growing number of buildings, the V2X connections can be disturbed due to the phenomenon of channel fading, which is spatially correlated with the number of fixed obstacles locations [11]. In addition, the mobility of vehicles has an impact on the stability of the cluster node due to its effect on the merge and the split of the cluster [12]. As a result, vehicle mobility faces technical and organizational challenges in maintaining connections between vehicles in V2V mode on the one hand and between vehicles and infrastructure (RSU/BS) on the other hand. Moreover, since the vehicle constantly changes

position and moves in and out of the coverages of other vehicles, RSUs, and BSs.

The management of mobility becomes essential to maintain the continuity of shared services and to achieve better QoS in a heterogeneous environment. Hence the importance of using handover strategies and an appropriate network selection scheme. There are two types of handover strategy, the first is horizontal and concerns the data session transmission from one PoA to another in the same network, the second strategy is the vertical handover [13], used when a data transmission session is transmitted from one PoA to another belonging to another network (using a completely different access technology).

A. Choice of Criteria

The selected criteria have been retained since they cover the different aspects that a network should satisfy.

Capacity must be significant to meet the increased demand for bandwidth. In fact, transmitting alerts and information messages requires a moderate capacity since they are transferred occasionally. Nevertheless, autonomous driving systems require a regular exchange with vehicle onboard sensors, in addition to the applications used by road users, which consume a considerable amount of data. For the second criteria, VANET network security faces several challenges, especially with IoT integration, which will increase the number of end-users and connected objects. Different types of attacks have emerged and evolved over time, such as availability and data integrity attacks, which have irreversible consequences on vehicles and people's data and can sometimes lead to road users' death, hence the importance of taking this criterion into account in this study. On the other hand, Coverage is a determining criterion in the infrastructure selection because vehicles will be able to communicate with the infrastructure only if they are within the coverage area of the BS and RSU. Studies have shown that 4G coverage is better than 5G coverage, especially as the latter uses small cell technology. For this reason, during the simulation, the RSU coverage was given an advantage over the 5G coverage. Finally, for the last two criteria, in the VANET context, several constraints slow down the response time (mainly high mobility and dynamic

network topology) and others that affect the reliability of the data exchanged, and what has been noticed in the tests is that improving reliability harms performance and latency and vice versa.

B. MADM Methods

In the context of the heterogeneous environment, there are several types of MADM methods (SAW, GRA, TOPSIS, AHP, MEW) that have proved their efficiency in solving VHO decisions mainly for the simplicity of their implementation. According to researchers, the most efficient methods used in the VHO process are TOPSIS, SAW, and VIKOR[14]. The principle on which they are based is ranking the candidate networks according to their scores. The researchers[15] proposed given weights in order to know the importance of each attribute in the QoS class under consideration. For other researchers, such as [16], they studied using AHP/TOPSIS pair for decision making situation, AHP to calculate the vector weight and TOPSIS for ranking access network, and their results showed the importance of the weights in decision making. Concerning decision-making, algorithms based on intelligent computation are the most performant, using a neural network and intelligent implementation techniques such as fuzzy logic. Indeed,[17] have been interested in the Fuzzy AHP to determine the weight of the evaluation criteria and the Fuzzy TOPSIS for classifying the alternatives. For these reasons, we have included fuzzy logic in the model to ensure the appropriate network selection.

C. System Model

This study focuses on the proposal of a MADM approach to evaluate the two alternatives subject to this paper: cellular and road infrastructure. To do so, we start by identifying the alternatives and the evaluation criteria, and then we build the hierarchy of decisions. The criteria retained to carry out the handover decision are those of capacity, security, coverage, reliability, and latency. Finally, TOPSIS is applied to the weighted matrices to obtain the ranking of the two infrastructures. One thousand iterations covering the different weight vectors have been performed (Fig. 6).

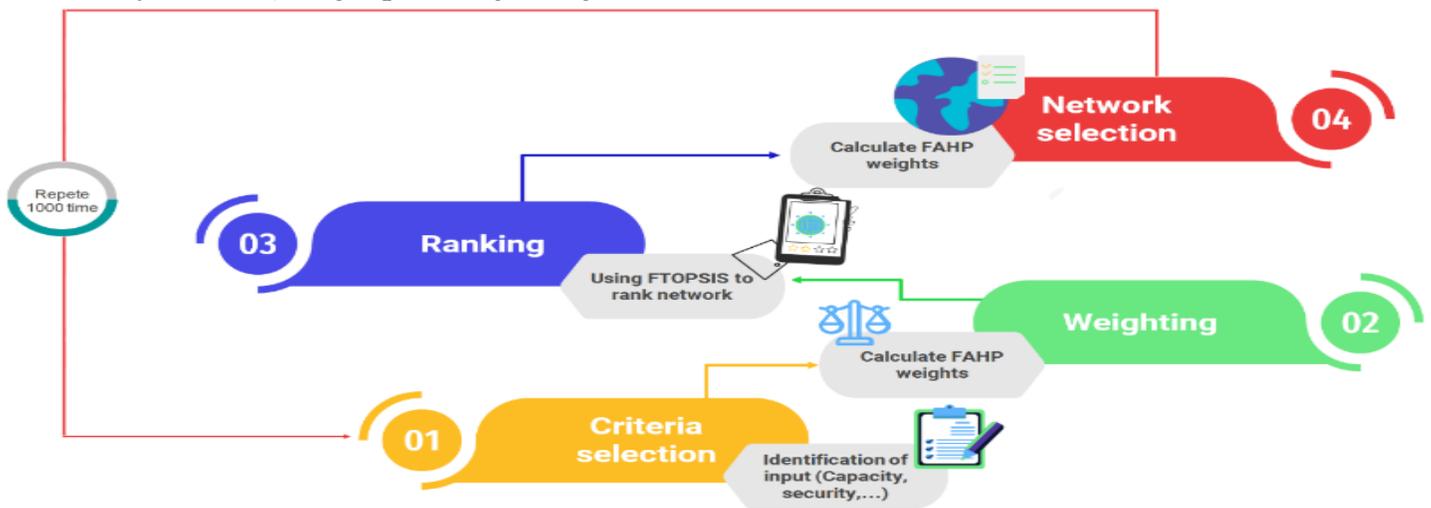


Fig. 6. System Model.

D. Network Selection using Fuzzy MADM Approach

To deal with the uncertainty of human decisions and thoughts, Zadeh [18] introduced the fuzzy set theory to represent the vagueness relative to certain classes of objects mathematically. Its ability to represent vague data is considered an important contribution in the field of mathematics and technology in general.

A fuzzy set (A in X) is characterized by a membership function f which associates for each point belonging to X a real number in the interval [0,1], $f_A(x)$ represents the membership grade of x in A.

$$f_{A(x)} = \begin{cases} \frac{x-l}{m-l}, & l < x < m \\ \frac{h-x}{h-m}, & m < x < h \\ 0, & \text{Otherwise} \end{cases} \quad (1)$$

In this case, TFN (Triangular fuzzy number) will be used in order to present fuzzy relative importance (Fig 7). In the pair-wise comparison the TFNs used are defined by three real numbers expressed as a triplet (l,m,h) with $l \leq m \leq h$ in order to describe a fuzzy event. The choice of TFN is generally related to the number of tuning and classifications (Table 1).

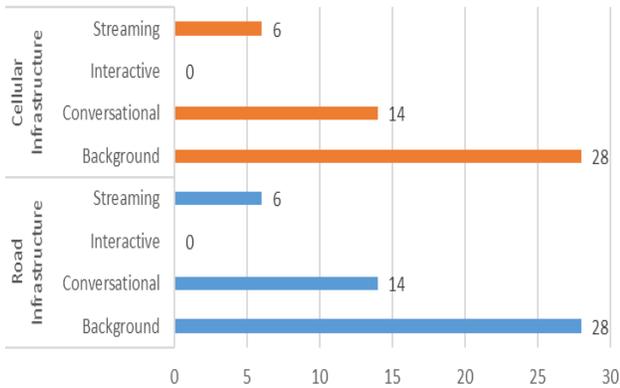


Fig. 7. Number of Vertical Handovers for 1000 Iterations.

TABLE I. MEMBERSHIP FUNCTION OF LINGUISTIC SCALE

Fuzzy Number	Linguistic scales	TFN
$\tilde{1}$	Equally important (Eq)	(1,1,1)
$\tilde{3}$	Weakly important (Wk)	(2,3,4)
$\tilde{5}$	Essentially important (Es)	(4,5,6)
$\tilde{7}$	Very strongly important (Vs)	(6,7,8)
$\tilde{9}$	Absolutely important (Ab)	(9,9,9)

a) AHP: The AHP technique has been introduced to analyze and organize complex decisions. However, research has shown some weaknesses in the AHP Saaty method [19]. FAHP appeared to overcome this deficiency; this new method combines AHP using fuzzy logic and linguistic variables. The most important step in the FAHP process is the generation of the relative fuzzy importance for each pair of factors. Using TFN and through the pair-wise comparison, the following fuzzy valuation matrix is obtained:

$$\tilde{V} = (\tilde{v}_{ij})_{n \times m} = \begin{pmatrix} (1,1,1) & \dots & \tilde{v}_{1m} \\ \vdots & \ddots & \vdots \\ \tilde{v}_{n1} & \dots & (1,1,1) \end{pmatrix} \quad (2)$$

Such as:

$$\tilde{v}_{ij} = (l_{ij}, m_{ij}, u_{ij})$$

Concerning the process of weighing using the fuzzy AHP, we opted for the Bucklet's method [20]. The geometric mean method is used for the calculation and analysis of the resulting vector in the criteria of comparison. From the equation (1), the geometric mean procedure takes the following form:

$$\tilde{r}_i = (\tilde{v}_{i1} \otimes \dots \otimes \tilde{v}_{im})^{1/m} \quad (3)$$

Thereafter the calculation of the weights is given by:

$$\tilde{w}_i = \tilde{r}_i \otimes (\tilde{r}_1 \oplus \dots \oplus \tilde{r}_m)^{-1} \quad (4)$$

The following equation is used to calculate the final resulting vector [21]:

$$\tilde{w}_i^{A/G} = (\tilde{w}_i^{A/C_1} \otimes \tilde{w}_1^{C/G}) \oplus \dots \oplus (\tilde{w}_i^{A/C_m} \otimes \tilde{w}_m^{C/G}) \quad (5)$$

To complete, the fuzzy AHP method is applied to each of the QoS classes and weights are generated for each criterion.

b) TOPSIS: For a system selection requirement, a MADM method called TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) was developed in 1981. It is a ranking technique that is simple to conceive and apply. TOPSIS aims to propose the best alternative that simultaneously has the shortest distance to the ideal positive solution and the farthest distance to the ideal negative solution[22]. the steps of TOPSIS have been detailed in [23] We make the normalized decision matrix of beneficial and non-beneficial criteria:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \quad (6)$$

Concerning the weighted normalized decision matrix, it is calculated as follows:

$$t_{ij} = W_i \cdot n_{ij} \text{ where } \sum_{j=1}^n W_i = 1 \quad (7)$$

The ideal positive and negative solutions are:

$$p^+ = \{v_1^+, \dots, v_2^+, \dots\}, v_j^+ = \max_i(v_{ij}) \quad (8)$$

$$p^- = \{v_1^-, \dots, v_2^-, \dots\}, v_j^- = \min_i(v_{ij}) \quad (9)$$

The distances that separate each Alternative P from the positive ideal solution and the separation of this alternative from the negative anti-ideal solution are calculated as follows:

$$d_i^+ = \sqrt{\sum_{j=1}^M |v_i^+ - v_{ij}|} \quad (10)$$

$$d_i^- = \sqrt{\sum_{j=1}^M |v_i^- - v_{ij}|} \quad (11)$$

E. Finally, we Calculate the Relative Proximity to the Ideal Solution by Applying the Following Formula:

$$0 \leq R_i = \frac{d_i^-}{d_i^+ + d_i^-} \leq 1 \quad (12)$$

a) *QoS Class*: In this study, we have retained three main classes from the IEEE 802.16 Broadband Wireless Access Working Group project: conversational, streaming, interactive, and background class. The distinguishing factor between these classes is the delay sensitivity of the traffic. For example, for the conversational class, it is dedicated to traffic, which is very sensitive to delay, unlike the background class, which is much less delay-sensitive.

The different applications used on the road network are divided into these classes according to their delay sensitivity but also in terms of their need in terms of bandwidth and the level of security required, for this purpose we have ordered each criteria by level of importance for each class by assigning each one a TFN, thereafter we have calculated their weights which we will show in a table in the following section.

V. EXPERIMENTAL RESULT

A. Methodology

The simulations designed for the two infrastructure systems, Road and Cellular based on 5G, involved four application types (background, Streaming, Conversational, and Interactive) that cover different QoS requirements for the user side. For each of the available alternatives, the selected criteria were randomly generated according to the intervals mentioned in Table 3 for the thousand iterations of the experimentation. Finally, each range level was associated with a different vector weight. The criteria weights are calculated for each class of service using the fuzzy AHP method and are listed in Table 2.

TABLE II. ASSOCIATED WEIGHTS FOR EACH QoS CLASS

QoS Class	Capacity	Security	Coverage	Reliability	Latency
Background	0,053	0,403	0,105	0,409	0,030
Streaming	0,097	0,145	0,156	0,316	0,286
Conversational	0,090	0,219	0,197	0,044	0,450
Interactive	0,075	0,177	0,095	0,313	0,340

TABLE III. COMPARISON CRITERIA

Criteria	Capacity (Mbit/s)	Security (%)	Coverage (m)	Reliability %	Latency (ms)
Road Infra.	[10,30]	70	[1,1000]	40	[10,100]
Cellular Infra.	[1000,3000]	50	[1,100]	50	[1,10]

B. Results

The results of the first phase of tests performed on the two infrastructures for the different iterations of each scenario show that the approach adopted ensures an equal distribution of vertical handover decision between the two infrastructures for the different types of applications, except for interactive communications (Fig. 8). In fact, this proves that for V2V and V2I vehicle interactions, the mobile infrastructure is more adapted than the road network for this kind of exchange.

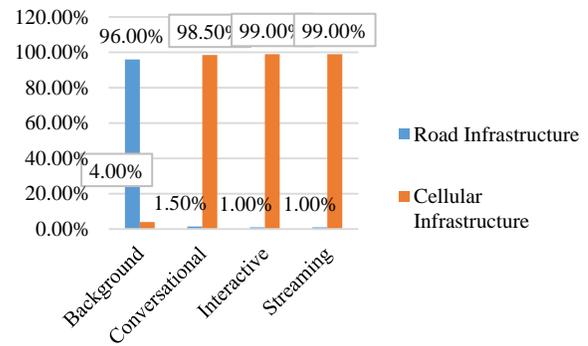


Fig. 8. The Average Number of Infrastructure Selections Per QoS Class.

For the second phase, the simulations showed that the cellular infrastructure exceeded the number of times it was selected, except in background applications, which revealed that vehicles tend to use the road infrastructure to run their programs in this particular case.

The received results demonstrate the potential benefit that 5G has over the other technologies used in VANET's V2X communications. This will have the effect of better managing the different types of communications as well as improving the services offered by each of the stakeholders in the VANET environment.

VI. CONCLUSION AND PERSPECTIVES

This article proposes a new approach based on Fuzzy MADM for selecting the most appropriate infrastructure in a VANET network. The two infrastructures that are the subject of this study are the cellular and the road infrastructure. To conduct this comparative study, five criteria have been considered: capacity, security, coverage, reliability, and latency. We started by presenting the VANET network with its main components, characteristics, and different communication modes. We also introduced the 5G mobile infrastructure, highlighting the main technologies that distinguish it from the previous versions. Then we ended the article by the comparative study while explaining the system model adopted and the different methods used to test and evaluate the two infrastructures, namely the Fuzzy AHP, to generate the weights of the evaluation criteria, and the TOPSIS method was applied to rank the alternatives. The adopted approach also allows us to select better the infrastructure in a shorter delay and less complexity. This approach's advantage for managing vertical handover decision in this type of network characterized by its dynamic topology and its nodes' high mobility has also been demonstrated.

We intend to include vehicle mobility parameters during the tests to have a more realistic simulation in future work. We also plan to test other combinations of MADM methods such as VIKOR and SAW to select the best network and infrastructure.

REFERENCES

- [1] A. Hussein, I. H. Elhadj, A. Chehab, et A. Kayssi, « SDN VANETs in 5G: An architecture for resilient security services », in 2017 Fourth International Conference on Software Defined Systems (SDS), Valencia, Spain, mai 2017, p. 67-74. doi: 10.1109/SDS.2017.7939143.

- [2] A. Houari et T. Mazri, « A comparative study of road infrastructure and 5G- based cellular infrastructure in the case of a VANET network. », p. 5.
- [3] K. MOGHRAOUI, « GESTION DE L'ANONYMAT DES COMMUNICATIONS DANS LES RÉSEAUX VÉHICULAIRES AD HOC SANS FIL (VANETs) », Thesis, L'Université du Québec, Montréal, 2015. [En ligne]. Disponible sur: https://oraprdnt.uqtr.quebec.ca/pls/public/docs/FWG/GSC/Publication/1645/34/1918/1/75448/8/F1689012680_M_moire_FINAL_K_Moghraoui.pdf.
- [4] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, et H. Zedan, « A comprehensive survey on vehicular Ad Hoc network », *J. Netw. Comput. Appl.*, vol. 37, p. 380-392, janv. 2014.
- [5] F. Cunha et al., « Data communication in VANETs: Protocols, applications and challenges », *Ad Hoc Netw.*, vol. 44, p. 90-103, juill. 2016.
- [6] A. Houari, T. Mazri, et A. Addaim, « Using the static cloud vehicle for storing and caching data in VANET network », in 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, oct. 2020, p. 1-7.
- [7] M. N. Mejri, J. Ben-Othman, et M. Hamdi, « Survey on VANET security challenges and possible cryptographic solutions », *Veh. Commun.*, vol. 1, no 2, p. 53-66, avr. 2014.
- [8] M. S. Sheikh, J. Liang, et W. Wang, « Security and Privacy in Vehicular Ad Hoc Network and Vehicle Cloud Computing: A Survey », *Wirel. Commun. Mob. Comput.*, vol. 2020, p. 1-25, janv. 2020.
- [9] X. Li, J. Liu, X. Li, et W. Sun, « RGTE: A Reputation-Based Global Trust Establishment in VANETs », in 2013 5th International Conference on Intelligent Networking and Collaborative Systems, Xi'an city, Shaanxi province, China, sept. 2013, p. 210-214.
- [10] K. Abboud et W. Zhuang, *Mobility Modeling for Vehicular Communication Networks*. Cham: Springer International Publishing, 2015.
- [11] E. Yaacoub, F. Filali, et A. Abu-Dayya, « QoE Enhancement of SVC Video Streaming Over Vehicular Networks Using Cooperative LTE/802.11p Communications », *IEEE J. Sel. Top. Signal Process.*, vol. 9, no 1, Art. no 1, févr. 2015.
- [12] K. Abboud et W. Zhuang, « Stochastic Modeling of Single-Hop Cluster Stability in Vehicular Ad Hoc Networks », *IEEE Trans. Veh. Technol.*, vol. 65, no 1, Art. no 1, janv. 2016.
- [13] S. Fernandes et A. Karmouch, « Vertical Mobility Management Architectures in Wireless Networks: A Comprehensive Survey and Future Directions », *IEEE Commun. Surv. Tutor.*, vol. 14, no 1, Art. no 1, 2012.
- [14] Z. Xu, *Uncertain Multi-Attribute Decision Making*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015. doi: 10.1007/978-3-662-45640-8.
- [15] E. Stevens-Navarro et V. W. S. Wong, « Comparison between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks », in 2006 IEEE 63rd Vehicular Technology Conference, Melbourne, Australia, 2006, vol. 2, p. 947-951.
- [16] K. Savitha et C. Chandrasekar, « Trusted Network Selection using SAW and TOPSIS Algorithms for Heterogeneous Wireless Networks », *Int. J. Comput. Appl.*, vol. 26, no 8, Art. no 8, juill. 2011.
- [17] V. Sasirekha et M. Ilanzkumaran, « Heterogeneous wireless network selection using FAHP integrated with TOPSIS and VIKOR », in 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, Salem, févr. 2013, p. 399-407.
- [18] J. A. Goguen, « L. A. Zadeh. Fuzzy sets. Information and control, vol. 8 (1965), pp. 338-353. - L. A. Zadeh. Similarity relations and fuzzy orderings. Information sciences, vol. 3 (1971), pp. 177-200. », *J. Symb. Log.*, vol. 38, no 4, Art. no 4, déc. 1973.
- [19] R. W. Saaty, « The analytic hierarchy process—what it is and how it is used », *Math. Model.*, vol. 9, no 3-5, Art. no 3-5, 1987, doi: 10.1016/0270-0255(87)90473-8.
- [20] J. J. Buckley, « Fuzzy hierarchical analysis », *Fuzzy Sets Syst.*, vol. 17, no 3, Art. no 3, déc. 1985.
- [21] A. Radionovs et O. Uzhga-Rebrov, « COMPARISON OF DIFFERENT FUZZY AHP METHODOLOGIES IN RISK ASSESSMENT », *Environ. Technol. Resour. Proc. Int. Sci. Pract. Conf.*, vol. 2, p. 137, juin 2017.
- [22] C.-L. Hwang et K. Yoon, *Multiple Attribute Decision Making*, vol. 186. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981.
- [23] A. Vega, J. Aguarón, J. García-Alcaraz, et J. M. Moreno-Jiménez, « Notes on Dependent Attributes in TOPSIS », *Procedia Comput. Sci.*, vol. 31, p. 308-317, 2014.

A Modified Particle Swarm Optimization Approach for Latency of Wireless Sensor Networks

Jannat H. Elrefaei¹, Ahmed H. Madian⁴

Department of Radiation Engineering, NCRRT
Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt

Ahmed Yahya²

Department of Electrical Engineering
Al-Azhar University, Nasr City, Cairo, Egypt

Mouhamed K. Shaat³

Department of Reactors, NRC
Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt

Refaat M. Fikry⁵

Department of Engineering, NRC, Egyptian Atomic Energy
Authority (EAEA), Cairo, Egypt

Abstract—In time-sensitive applications, such as detecting environmental and individual nuclear radiation exposure, wireless sensor networks are employed. Such application requires timely detection of radiation levels so that appropriate emergency measures are applied to protect people and the environment from radiation hazards. In these networks, collision and interference in communication between sensor nodes cause more end-to-end delay and reduce the network's performance. A time-division multiple-access (TDMA) media access control protocol guarantees minimum latency and low power consumption. It also overcomes the problem of interference. TDMA scheduling problem determines the minimum length conflict-free assignment of slots in a TDMA frame where each node or link is activated at least once. This paper proposes a meta-heuristic centralized contention-free approach based on TDMA, a modified particle swarm optimization. This approach realizes the TDMA scheduling more efficiently compared with other existing algorithms. Extensive simulations were performed to evaluate the modified approach. The simulation results prove that the proposed scheduling algorithm has a better performance in wireless sensor networks than the interference degree leaves order algorithm and interference degree remaining leaves order algorithm. The results demonstrate also that integrating the proposed algorithm in TDMA protocols significantly optimizes the communication latency reduction and increases the network reliability.

Keywords—Wireless sensor networks; media access control protocol; scheduling algorithms; meta-heuristics; particle swarm optimization

I. INTRODUCTION

Nuclear energy is an important clean energy source. The probability of radiation releases from nuclear facilities is extremely low due to strictly applied safety measures. But the serious situations that could result from a nuclear accident make it essential to use online detection of radiation [1], [2]. Wireless sensor network (WSN) technology is usually integrated into time-sensitive applications such as online radiation monitoring systems because of its efficiency to deal with low-rate communications, simplicity in construction, reconfigurability, and low cost [3] [4]. An essential role of sensor networks is target coverage, whereas the sensor nodes' function collects data periodically and transmits it to the

destination node in the WSN. This many-to-one communication pattern is known as converge-cast [5], [6]. Although many WSN applications are beneficial, limited resources of WSN cause several challenges that need to be addressed for efficient performance. Energy consumption is the main problem as sensors are usually battery-powered [7], [8]. Minimizing communication latency is a fundamental objective of the alarm-driven WSNs applications or disaster early warning applications [9] - [13]. Communication latency is caused by several factors depending on the design of stack layers of the WSN. To minimize communication latency, the WSN protocols in different layers should overcome the following problems:

- **Transmission interference:** Conflicts are classified into two types: primary and secondary. When a node has several communication tasks in a same time slot, primary conflict arises. When a node switched to a specific transmitter inside the communication range of other communication process considered for it neighbours, a secondary conflict occurred [14]. The existence of a conflict between any two sensors would block these two sensors from transmitting simultaneously [15].
- **The radio transceiver of sensor node:** Sensor nodes use half-duplex transceivers due to hardware limitations as it is impossible for a node to send and receive simultaneously.
- **The topology of WSN:** It is an essential factor as a flat topology results in high latency, and a single-hop-to-sink topology achieves minimum delay, but scalability for its network is limited. The hierarchal or tree-based topology achieves low latency and satisfies power consumption [16].
- **The duty-cycle mechanism:** It controls the listen (transmit or receive) and sleep periods of sensor nodes. This mechanism is executed at the media access control (MAC) layer.
- **Overhearing:** When a node receives data intended for other nodes, it is referred to as overhearing.

Communication latency is affected mainly by the design of the data-link layer and the network layer [17]. The cross-layer strategy is two or more layers communicate to improve the network's overall performance [18]. The cross-layer approach is used here for latency, where the MAC protocol used routing protocol information to determine optimum schedules. MAC protocols conclude mainly in two types; contention-based MAC protocols as carrier sense multiple-access (CSMA) and contention-free MAC protocols as time division multiple access (TDMA) [19]. TDMA protocols are more suitable for heavy traffic conditions to avoid collisions successfully [16] [20]. Interference is an essential factor in determining the TDMA schedule. The sink node organizes the frame's scheduling, whereas time slots are specified to nodes in WSN for data transmission and reception considering interference. The TDMA scheduling remains fixed unless there is no reconfiguration in the WSN [21].

Designing efficient approaches to schedule TDMA transmissions in multi-hop WSNs is an NP-complete problem [22]-[25]. Meta-heuristic optimization algorithms are the most reliable techniques to find near-optimal scheduling that achieve minimum latency for critical WSNs [26] [27].

Eberhart and Kennedy introduced the particle swarm optimization (PSO) approach for swarm intelligence, a meta-heuristic evolutionary technique inspired by bird flocking or fish schooling's social behaviour [28]. A population of random solutions known as particles follows the same idea. These particles fly into a multidimensional search space in the direction of optimum value. Then the PSO algorithm presents a solution to the TDMA scheduling problem and provides the near-optimal schedule off-line, which is then used by the sink node to schedule the sensors in real-time.

This paper aims to discover the best result to the communication latency problem in multi-hop WSNs depending on cross-layer optimization. This paper's primary contribution is as follows:

- Employment of a modified PSO algorithm to optimize TDMA scheduling in a WSN to minimize communication latency.
- Formulation of a fitness function considering communication latency minimization.
- Provision of comprehensive simulation results to demonstrate the advantages of the modified PSO over other relevant algorithms in minimizing the communication latency.

Section II of this article presents a review of related work, and it gives the problem statement. The network model is developed and provided in Section III. The proposed PSO algorithm is discussed in Section IV; also, evaluation metrics are presented in the same section. Section V illustrates the simulation and assessment of performance. Section VI discusses the results, while Section VII concludes this work and suggestions future studies.

II. RELATED WORK

A. TDMA Scheduling Optimization

Several scheduling algorithms for data collection in WSNs have been developed and presented solutions for the aforementioned problems. Each algorithm has one or multi-objective associated with data collection according to its application requirements [29].

Sensor nodes interact via wireless multi-hop routing because of the limited range of radio transmission. One-hop TDMA scheduling is simpler than multi-hop scheduling. There is no requirement for spatial reuse of a time-slot since in one-hop TDMA scheduling, and several nodes can broadcast simultaneously if their receivers are not in conflict [30].

The scheduling algorithms use one of two interference models for evaluation; the protocol model depending on a graphing approach or the physical model depending on the Signal-to-Interference-plus-Noise-Ratio (SINR) as discussed in [25], [31].

Previous research efforts on WSN processing time have focused on a specific problem based on meta-heuristic approaches. Following [15], the authors used the PSO technique to reduce the overall transaction time. The PSO algorithm dealt with it as a graph partitioning problem and maximized the parallel operation of the network's sensors. A multi-objective optimization framework is executed as described in [32], where a genetic algorithm (GA) and PSO algorithm were combined to improve searching for a global optimum. This framework achieved a minimization in latency and a power conservation. TRASA (traffic-aware time slot assignment) algorithm is discussed and presented in [33] to gain minimum scheduling and fair medium access. Scheduling of nodes in different time-slots depends on the node's priority (i.e., a node with a high number of offspring, so it has more data to transmit). There are two versions of TRASA; one slot and many slots. When a node possesses a time-slot, only this time-slot is for this node, and many time-slots are assigned to nodes with high priority to transmit without switching delays. M. Bakshi et al. [25] Proposed an optimum converge-cast schedule in a WSN. They used the SINR model of interference and a TDMA -MAC protocol. The PSO algorithm optimized scheduling in multi-channel and multi-time-slot assignment WSN [11], [34]. This PSO algorithm improved the latency and the length of the frame. The Cross-layer approach, CoLaNet [35], is enhanced as described in [36], as the authors proposed new TDMA scheduling algorithms related to routing to reduce communication latency. They present Rand-LO, Depth-LO, and Depth-ReLO algorithms to improve latency. In the slot scheduling approach, these algorithms demonstrate the necessity of traversing the routing tree. Reference [10] considered the interference degree of sensor nodes in the proposed scheduling methods, IDeg-LO and IDeg-ReLO, and improved the network's latency. [9] Proposed an ETDMA-GA algorithm based on a genetic algorithm and cross-layer approach to obtain optimum TDMA scheduling for minimum latency. They compared the obtained results with [10], [36] and proved that the ETDMA-GA algorithm outperforms Rand-LO, Depth-LO, Depth-ReLO, IDeg-LO, and IDeg-ReLO in terms of average latency and average schedule length.

This paper aims to find the efficient and optimum TDMA schedule for sensor nodes based on routing information. The proposed modified PSO approach schedules sensor nodes efficiently and reduces communication latency for a network.

PSO algorithm has several advantages compared with other meta-heuristic algorithms. For example, its mechanism is more straightforward (few parameters to be adjusted), the computational cost is low, the convergence speed is high, and the quality of solutions [37]. A study comparing various swarm intelligence (SI) approaches used selected thirty benchmark functions that measure these performance approaches'. This study concluded that PSO is a second-best approach next to Differential Evolution (DE); it outperforms or equally performs to the best algorithm in eighteen out of thirty functions [38].

In WSNs, PSO has been used to address several problems such as energy conservation [39], coverage maximization [40], optimal deployment of sensors [41], clustering, clustering head selection, data aggregation [42], and node localization [43].

The PSO algorithm has also been modified to solve a variety of complex optimization problems. For example, it has been used to handle large-scale, constrained, multimodal, multi-objective, and discrete optimization problems, among others [44].

B. Problem Statement

As discussed above in Sections I and II, communication latency is a significant issue in WSN, which imposes challenges in alarm-driven WSN applications such as environmental radiation monitoring networks (ERMNs). This communication latency can be due to transmission interference, network topology, the half-duplex transceiver of the sensor node, duty cycle mechanism, and overhearing. Minimizing the converge-cost is an NP-complete problem, which could be addressed using meta-heuristic optimization algorithms to find a near-optimal off-line solution, which is then used by the sink node in real-time. Compared with other meta-heuristic algorithms, the PSO algorithm has several advantages, including its more straightforward mechanism, higher quality, lower computational cost, and higher conversion speed. The problem that is addressed by this work is the need to minimize the communication latency in ERMNs through optimization of the TDMA scheduling using a modified PSO algorithm.

III. MODELING OF WIRELESS SENSOR NETWORK

A. Network Model and Scheduling Model

The WSN employs static sensor nodes supplied with Omni-directional antennas and single half-duplex transceivers. It has one sink node that contains all information about synchronization between sensor nodes, topology, interference relationships, and determination of each node is parent or child.

A WSN is a graph of vertices (V) and edges (E), $G = (V, E)$ where V represents the sensor nodes and E corresponding to links between nodes. The connectivity model describes the way that the links are connecting nodes. The classical connectivity model based on the unit disk graph (UDG) is adopted. In UDG, any two nodes are considered adjacent if their Euclidean distance is the most [35]. The Euclidean

distance d_{ij} between nodes i and j, denoted by $d(i,j)$, is the least number of hops required to send data from one point to another. The communication range of all nodes is assumed to be the same, and therefore the links in the modeled graph are symmetric. Two matrices describe the topology of the WSN; the symmetric connectivity matrix C_{N*N} , which describes connectivity relations between neighbours as in (1), and the interference matrix I_{N*N} , which represents conflicts between neighbour nodes in the network as in (2)

$$C_{ij} = 1 \quad \text{if} \quad d(i,j) = 1; \quad \text{else} \quad C_{ij} = 0 \quad (1)$$

$$I_{ij} = 1 \quad \text{if} \quad d(i,j) \leq 2; \quad \text{else} \quad I_{ij} = 0 \quad (2)$$

According to these two matrices, the sink node determines the TDMA scheduling, and each node recognizes the time slot assigned to it for transmitting and receiving without interference.

Some constraints restrict the parallel operation of some WSN sensors. Interference between sensor nodes in the WSN forms the primary constraint that decreases successful transmissions. The scheduling algorithm proposed by this paper aims to maximize parallel instead of sequential operation of data transmission. An optimization problem for this solution was modeled.

The normal TDMA scheduling method is used for time-slot allocation for all nodes in the initial phase. For a randomly deployed WSN, the sink node applies a depth-first search [45] algorithm. The sink node constructs a shortest-path routing tree. Then the sink node determines the TDMA schedule and broadcasts it to all sensor nodes in the network. The scheduling technique depends mainly on searching for a first free time-slot for a node; a time-slot is free or suitable for a node if it is not busy with any one-hop or two-hop neighbour nodes. A traversal list, depending on the searching algorithm, orders the sensor nodes in the routing tree.

For each node's traversal list, based on the connectivity and interference matrices, a TDMA scheduling can be deduced as presented in the following slot allocation algorithm[9], [10]:

- 1) The frame length initializes with a size equal to the maximum node interference degree in the network.
- 2) The first time-slot is allocated to the first node in the traversal list for data transmission and all connected nodes for data receiving.
- 3) The sink node schedules the rest traversal list's nodes Similarly. The node's time-slot is for transmission and receiving data.
- 4) If there is no free time-slot for allocation, then an extra time-slot to the frame is added.

B. Optimization Algorithm

PSO algorithm initializes particles randomly and converges to the optimal solution by iterations. Each particle modified its velocity and then updates its position. It depends on its expertise (cognitive) and the expertise of other particles (social) [46]. Every particle has two vectors in the PSO

algorithm: position and velocity. The position vector describes the value and the direction of a particle at all iterations. The position function of a particle changes to a best in each iteration of the PSO algorithm by using (3):

$$\bar{x}_i(t+1) = \bar{x}_i(t) + \bar{v}_i(t+1) \quad (3)$$

Where $\bar{x}_i(t)$ means the position of a particle (i) at iteration t and $\bar{v}_i(t+1)$ presents the velocity of the particle (i) at iteration (t+1). Equation (3) determines the update in a particle's position, which depends mainly on its velocity vector. The velocity vector of the particle (i) at iteration (t+1) is as in (4):

$$\bar{v}_i(t+1) = \omega \bar{v}_i(t) + c_1 r_1 (\bar{p}_i(t) - \bar{x}_i(t)) + c_2 r_2 (\bar{g}(t) - \bar{x}_i(t)) \quad (4)$$

The coefficient ω , known as inertial weight, describes the individual coefficient known as the social coefficient and random numbers [0, 1]. In velocity (2), the equation of velocity (4) contains three parts.

The first part $\omega \bar{v}_i(t)$ keeps the orientation the same as the current velocity of the particle. It tunes exploration and exploitation, which are essential to achieve the exact estimation of a global optimum.

The second part: $c_1 r_1 (\bar{p}_i(t) - \bar{x}_i(t))$ describes the single-particle intelligence by saving in its memory and using the obtained best cost to evaluate and update the particle position. The vector $\bar{p}_i(t)$ is the personal best value of the particle (i) at iteration (t), and it is updated each new iteration if the particle (i) finds a better solution. It affects the final value of the velocity is adjusted using c_1 . The second part keeps a direction in the direction of the personal Best value obtained by a particle.

The last part: $c_2 r_2 (\bar{g}(t) - \bar{x}_i(t))$ describes the population's social intelligence and saves the optimum value obtained by all particles in it. The vector $\bar{g}(t)$ means considering the particles' best solution in search space attracts all particles in the global best solution's direction. The impact of this component can be adjusted using c_2 . These components help in updating the position of a particle in search space. The optimization algorithm evaluates the suitability of population particles depending on a fitness function. The chosen fitness function is related mainly to the aim of optimization. The PSO algorithm structure includes the following steps:

- 1) Initialization: An initial population of particles is generated randomly in the search space with random velocities and positions.
- 2) The algorithm determines each particle's position and velocity as in (3) and (4). Then, according to an evaluation process, it calculates the fitness value for each particle.
- 3) Updating the personal Best and the global Best values, if the fitness value is better than the best fitness value, then change the current value to the new value.
- 4) If it is achieved, the stopping condition will go to step 5; otherwise, go back to step 2.

5) Termination of the algorithm and illustration of the global optimum value is done.

IV. PROPOSED PSO ALGORITHM

The PSO algorithm presented here is modified to suit the time-sensitive application addressed by this work (transmission of data on radiation levels). The flow chart of the modified PSO is illustrated in Fig. 1. After WSN nodes' deployment in a specified area to be monitored, a depth-first search algorithm is applied to the deployed WSN to form a shortest path routing tree. It is possible to produce traversal lists associated with all nodes in this tree and corresponding TDMA schedules considering interference and connectivity between nodes.

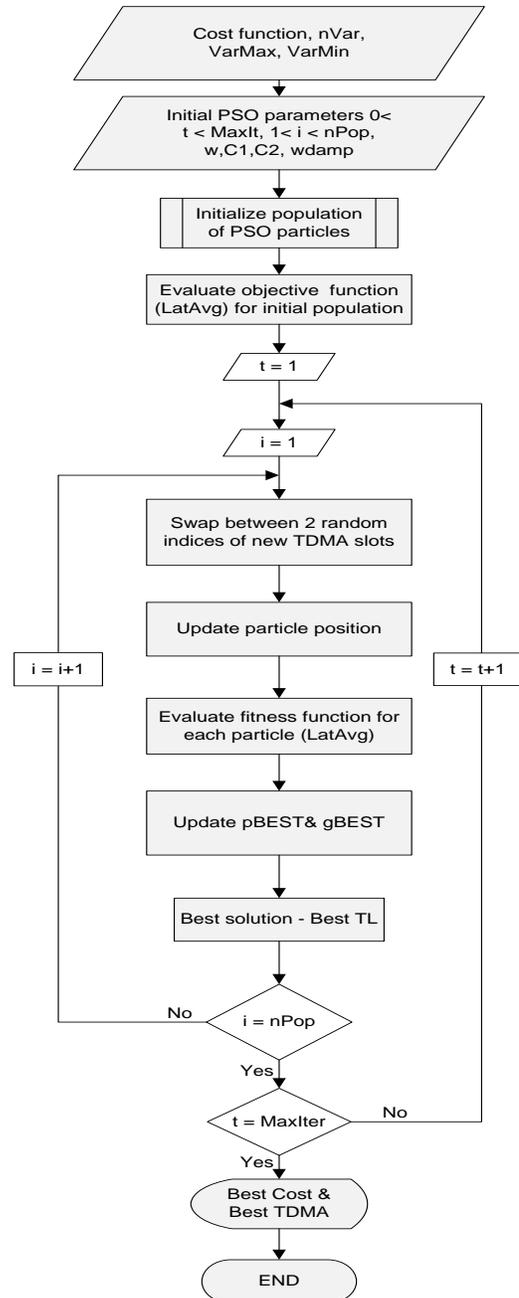


Fig. 1. Flowchart of the Proposed PSO.

A. The Modified PSO Model

The modified PSO re-orders allocated time-slots in the TDMA schedule randomly until optimum scheduling, which reduces the average latency value. The algorithm is as follows:

- 1) Initialization: An initial population of 100 particles (normal TDMA schedules) is generated randomly in the search space with random initial positions.
- 2) The modified PSO algorithm swaps between two random indices of TDMA slots.
- 3) The algorithm determines each particle's position. Then, according to an evaluation process, it calculates the average latency value for each TDMA schedule.
- 4) Keeping personal and global Best values up to date, if the fitness value is better than the best fitness value, then change the current value to the new value.
- 5) If the stopping condition is met, go to step 5; otherwise, go back to step 2.
- 6) Termination of the algorithm and illustration of the optimum value of average latency and its TDMA schedule is done.

The particles in the proposed algorithm are the WSN's TDMA schedules, and their values are optimized by swapping randomly between any two slots, with the children's parent wakes up after it, not before it. As a result, the particle's position is independent of its velocity. Following each TDMA schedule's slot swap, an evaluation is performed to determine the best particle, the TDMA schedule, which has the minimum delay.

B. Evaluation Matrices

The design aim of this paper is to minimize the entire time required to complete a series of tasks based on an optimal time-slot assignment for the TDMA schedule. This objective can be achieved by reducing scheduling length and minimizing average latency. Discussion of these parameters is provided as follows:

- 1) Average latency and average normalized latency: The sum of latencies associated all the WSN's nodes defined the average latency, in according to (5):

$$L_{avg} = \frac{\sum_{i=1}^n L_i}{n-1} \quad (5)$$

Where L_i is the latency related to a node (i)

the average latency per link (the node's delay divided by the number of hops along the shortest routing path between this node and its destination) calculates average normalized Latency (L_{norm}) for all sensors in the network and is given by (6):

$$L_{norm} = \frac{\sum_{i=1}^n L_i / h_i}{n-1} \quad (6)$$

Where (h_i) is the sum of hops along the routing path from a source node (i) to the sink node.

- 2) Schedule length: The number of time-slots in the derived TDMA frame is used to calculate the schedule length. A minimum TDMA schedule length reduces energy consumption as the sleep period for nodes increases. Increasing the usage of the time slots can decrease the size of the schedule. Most algorithms work to maximize frequent synchronous transmissions and allow spatial reuse by development.

- 3) Duty cycle: The ratio between the active mode interval and the whole frame defined the sensor node's duty cycle. Its minimum value results in more power conservation in the sensor node [7]. The duty-cycle is calculated for the TDMA-MAC protocol as the ratio between the total numbers of time-slots engaged with communication to the schedule length.

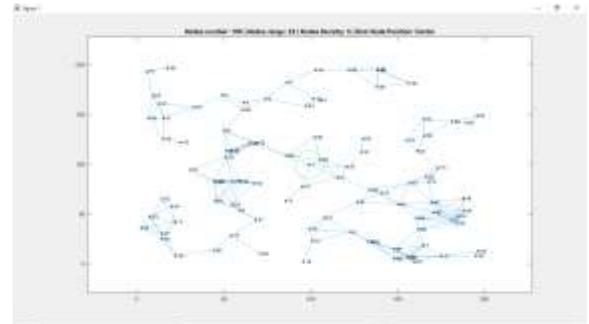
V. SIMULATION AND PERFORMANCE EVALUATION

A. Simulation Setup

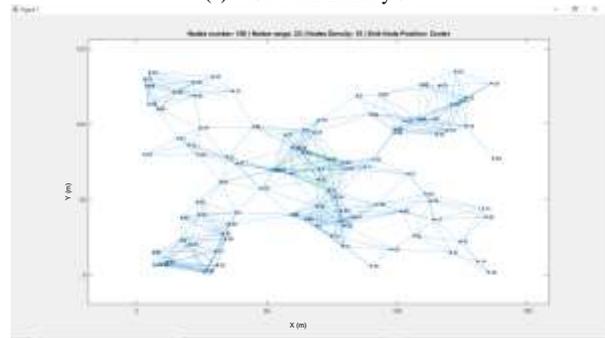
A randomly deployed WSN of a hundred sensor nodes on a square area, and each sensor has the same communication range. This work uses randomly generated WSNs of varying densities achieved by adjusting the scale of the deployment area of the network. Fig. 2(a-d) presents examples of randomly generated WSNs with different densities. The density (δ) describes the average number of neighbours per node in the network as in (7) [47]:

$$\delta = \pi * r^2 * \frac{N}{a^2} \quad (7)$$

r is the communication range of a sensor node, N is the number of nodes, and a^2 is the deployment area.



(a) WSN with Density 5.



(b) WSN with Density 10.

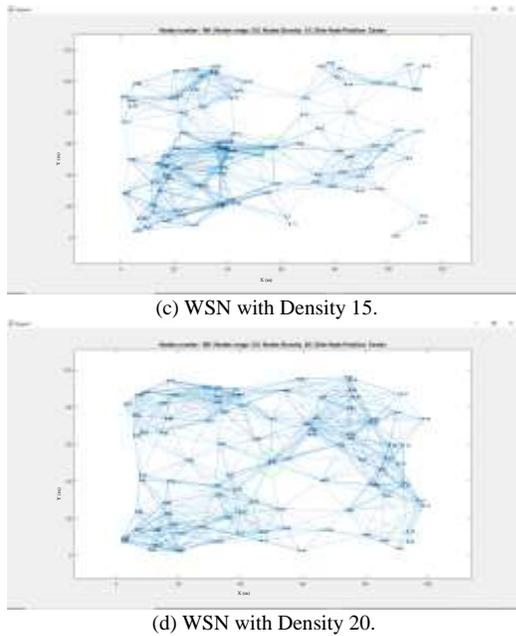


Fig. 2. Randomly Generated WSNs of 100 Sensor nodes with Densities of 5, 10, 15, 20 and Sink Node at Center.

WSN was simulated using numerical simulation of MATLAB in the 1.6 GHz laptop with 4 GB RAM to evaluate the performance of the proposed modified PSO algorithm. Table I presents the simulation setting parameters for the network and the PSO algorithm. The maximum number of particles equals the WSN's TDMA schedules. And these schedules provides using the WSN's traversal lists which equals the WSN's nodes.

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Number of sensors (N)	100 static sensors
Sink node	1
Communication range(r)	25m
Position of the sink node	Center, Corner, MiddleEdge
Network density (δ)	5, 10, 15, 20
Network area (a^2)	Change according to δ
MaxIt	30:80 (according to δ)
nPop	100
Inertia Weight(w)	1
C1	1
C2	2

B. Simulation Results and Evaluation

The best TDMA schedule depends on the initial routing tree generated using the DSF- algorithm, and then the sink node broadcasts this best TDMA schedule in the network. So, the problem of collision is overcome. The proposed modified PSO is a centralized contention-free approach that is achieved by the sink node. Experiments were carried out to compare the proposed algorithm's performance to that of other algorithms.

1) *Convergence time and stopping criterion*: The proposed modified PSO algorithm consists of a swarm of particles exploring the search space for searching for a globally optimum solution. The global optimum solution is challenging to determine, so it is essential to decide on the stopping condition and the convergence for the proposed algorithm. These two coefficients are related to the PSO algorithm [28]. The primary requirement for a convergence of PSO is the stability of the state of its particles. A stable condition means that the distance between the current and the particle's previous position is never more significant than a given threshold value ($\epsilon > 0$), or the difference between fitness functions is minimal. The particle convergence time is the minimum number of steps necessary for the particle to reach its stable state. Its value depends on PSO parameters and the convergence level of the fitness function. The value of the convergence time (t) can be calculated theoretically using (8):

$$t = O(-\log \epsilon) \quad (8)$$

As shown in (8), the relationship between the convergence time and the convergence level of PSO is linear. The convergence time can also be computed experimentally by a hundred experiments on the WSN with a particular density and observe the stable state of the fitness function [48]. Simulations for the modified PSO algorithm are performed using various maximum iteration parameters and monitor the fitness function's stability for each WSN density.

The average latency becomes converged as shown in Fig. 3 at iterations 30, 40, 60, 80 for WSN densities 5, 10, 15, 20, respectively, which are the values to them the proposed modified PSO converged. Fig. 3 shows that the maximum iteration number is proportional to the network density; with the increase in the network density, more iterations are needed to fulfil the process optimization.

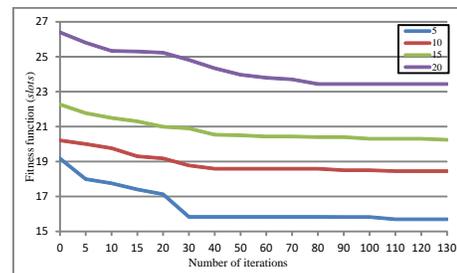


Fig. 3. A Convergence of the Modified PSO Algorithm.

The elapsed time for execution of the proposed algorithm and give a result is presented in Table II for the WSN with a sink node at the center.

TABLE II. THE ELAPSED TIME THE EXECUTION OF THE MODIFIED PSO ALGORITHM

Elapsed time(sec.)	Node density(δ)			
	5	10	15	20
t	74.968	123.3759	159.732	268.857

2) Average latency and average normalized latency: The next experiment compared the outcomes of the normal (initial), IDeg-LO, and IDeg-ReLO algorithms to the optimal value of average latency and average normalized latency. Fig. 4(a-c) and Fig. 5(a-c) illustrate the results of the simulation

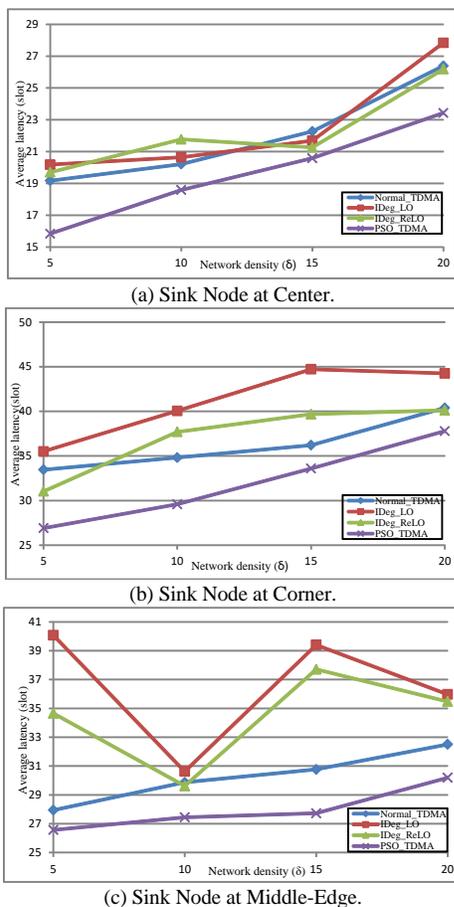


Fig. 4. Average Latency based on the Depth-First Search Routing Tree for Several Scheduling Approaches and Different Sink Node Locations.

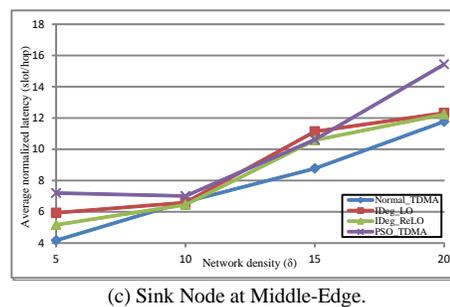
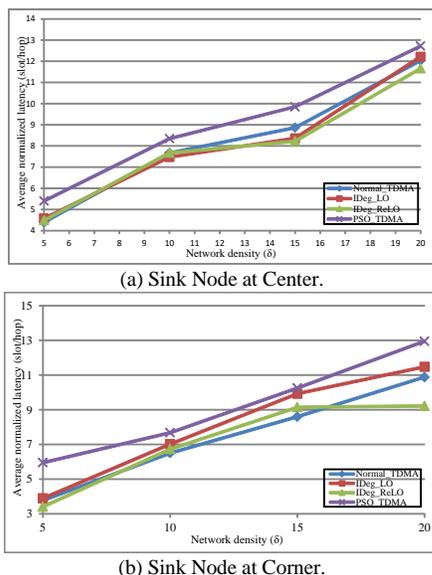


Fig. 5. Average Normalized Latency based on the Depth-First Search Routing Tree for Several Scheduling Approaches and Various Sink Node Locations.

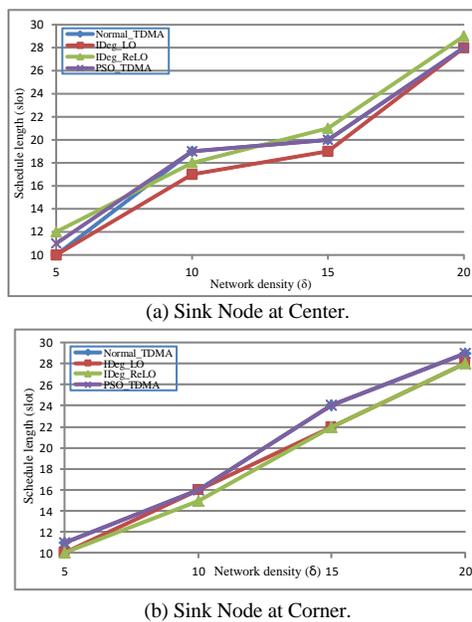
In every situation of network density or sink node placement, the proposed PSO algorithm maximizes the value of average latency and average normalized latency and outperforms existing methods, as shown in Fig. 4 and 5. The gain in average latency reduction of the modified PSO approach compared with Normal, IDeg-LO, and IDeg-ReLO at different densities is averaged and tabulated in Table III.

TABLE III. THE AVERAGE GAIN OF THE MODIFIED PSO IN LATENCY REDUCTION

Sink Node Location	Algorithms		
	Normal	IDEG-LO	IDEG-RELO
Center	14%	29%	17%
Corner	13%	16%	14%
MiddleEdge	8%	31%	23%

The modified PSO from Table III improves the average latency up to 31% better than IDeg-LO and up to 23% compared with IDeg-ReLO.

3) Schedule length: The modified algorithm's schedule length was calculated and compared to other algorithms' results. The outcomes are as follows in Fig. 6.



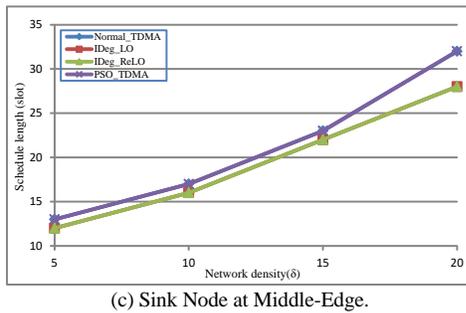
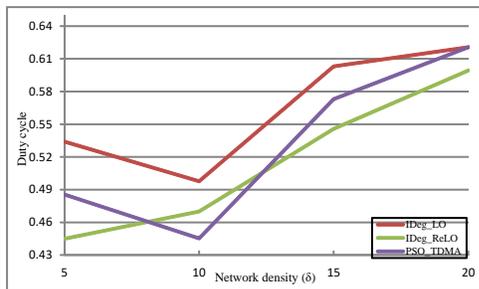


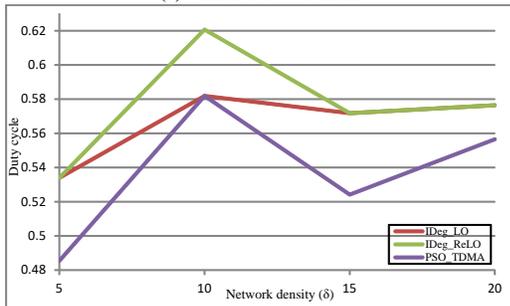
Fig. 6. Schedule Length used in Modified PSO, IDeg-LO, IDeg-ReLO, and Normal Algorithms.

Fig. 6 shows that the modified algorithm's schedule length is the longest one in approximately all cases of the sink node locations and different network densities. It is more than that of IDeg-ReLO by one time-slot, by two time-slots in the case of IDeg-LO, and by four time-slots over these two algorithms for WSN of density 20 and sink node at MiddleEdge.

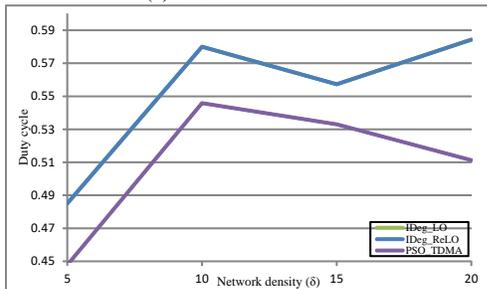
4) *Duty cycle*: The duty cycle of the modified algorithm was evaluated and compared with that of the IDeg-LO and IDeg-ReLO. Fig. 7 shows the experiment results in all cases of the sink node locations.



(a) Sink Node at Center.



(b) Sink Node at Corner.



(c) Sink Node at Middle-Edge.

Fig. 7. Duty-Cycle of Modified PSO, Normal, IDeg-LO, and IDeg-ReLO Algorithms with the Three Locations of the Sink Node.

Fig. 7 shows that the modified PSO has a lower duty cycle value than that for the IDeg-LO and IDeg-ReLO, reaching up to a 14.2% reduction value. This value is acceptable in the WSNs' time-sensitive applications.

VI. DISCUSSION

The TDMA schedule has been optimized to remove all possible transmission interferences, which are the leading cause of communication latency in the network. A modified PSO algorithm was developed to provide near optimum interference-free TDMA scheduling, resulting in minimum WSN latency.

The results of simulation and evaluation show that the modified-PSO algorithm outperforms the competition. The modified PSO outperforms IDeg-LO and IDeg-ReLO in minimizing the WSN average latency, respectively, by 31 and 23%. The results show also that the scheduling length of TDMA is the largest in most cases of the WSN. Using the modified PSO, the scheduling length reaches up to 8% over IDeg-LO and IDeg-ReLO addition of time-slots to overcome interference. The value of the duty-cycle using the modified PSO is lower than that of IDeg-LO and IDeg-ReLO at all densities. In middleaged sink node WSN, it scores a reduction of 14.28%. While in a Center sink node WSN, IDeg-ReLO results in a decrease in the duty-cycle value except at a WSN density of 10. The modified PSO is more effective than IDeg-LO and IDeg-ReLO when the sink node is at the corner of the WSN. This efficiency is demonstrated by achieving a reduction in the duty cycle of up to 9%.

The results show also that the value of the duty-cycle depends on the schedule length. The increase in the schedule length results in minimizing the duty-cycle, resulting in more power conservation.

It is also important to note that the results of this work contribute towards a more precise understanding of the relation between the schedule length and the duty cycle as the effect of the sink node location, which is an essential parameter in minimizing latency.

This work confirmed that the simplicity of execution of the algorithm and the fast convergence of the PSO meta-heuristic for optimization of the TDMA-scheduling for WSN communication latency problem; are significant advantages when using WSN in time-sensitive applications. The modified PSO algorithm is executed in the sink node, which has sufficient memory for algorithm implementation. The obtained results are beneficial for the effective design of WSNs for time-sensitive applications. This approach may be used to acquire data from the surrounding environment with minimal latency in a WSN with static sensors. It has applications in the military, climate monitoring, and the supervision and management of nuclear power facilities. This modified PSO may be used in a variety of ways, including adding mobile sensors and underwater, subterranean sensors.

VII. CONCLUSION AND FUTURE WORK

This study proposed a centralized contention-free TDMA scheduling algorithm. It is a meta-heuristic, a modified PSO algorithm that optimizes the latency reduction and the

network's duty cycle. A comprehensive series of simulations was established, and a comparison was made on the performance of the Modified PSO, IDeg-LO, and IDeg-ReLO. This work demonstrated that the Modified PSO algorithm provides a shorter latency than IDeg-LO and IDeg-ReLO algorithms.

The modified PSO algorithm improves the average latency by 31% compared with IDeg-LO and by 23% compared with IDeg-ReLO depending on the routing tree. Additionally, it improves the duty cycle by 14.2% compared with that of IDeg-LO and IDeg-ReLO algorithms. Therefore, the modified PSO approach results in the most acceptable WSN applications because of its simplicity and low computational cost. It is also more suitable for time-sensitive applications as it offers the expected speed of convergence with high quality of solutions.

Future work could be performed to study optimization techniques for calibrating the power consumption to the lowest possible level while ensuring the connectivity and reliability of the transmitted data.

REFERENCES

- [1] M. V. Karthikeyan and R. Manasa, "Nuclear radiation detection using low cost wireless system: Protection of environment against nuclear leakage and dump," presented at the Recent Advances in Space Technol. Services and Climate Change 2010 (RSTS & CC-2010), Chennai, 2010, pp. 25-28, doi: 10.1109/RSTSCC.2010.5712792.
- [2] J. Lúley, Š. Čerba, B. Vrban, F. Osuský, and O. Šfuka, "RADIATION MONITORING SYSTEM USING UNMANNED AERIAL VEHICLES," *Radiat. Prot. Dosimetry*, 2019 Dec 31; 186 (2-3):337-341. doi: 10.1093/rpd/ncz229. PMID: 31846036.
- [3] R. P. Hudhajanto, N. Fahmi, E. Prayitno, and Rosmida, "Real-time monitoring for environmental through wireless sensor network technology," presented in the Int. Conf. Appl. Eng. (ICAE), Batam, Indonesia 2018, pp. 1-5, Oct.
- [4] H. R. Galappaththi and G. T. Weerasuriya, "Survey on Wireless Sensor Networks (WSNs) Implemented for Environmental Sensing," presented in the 3rd Int. Conf. Info. Tech. Res. (ICITR), Moratuwa, Sri Lanka, pp. 1-6, 2018, doi: 10.1109/ICITR.2018.8736151.
- [5] J. Mao, Z. Wu, and X. Wu, "A TDMA scheduling scheme for many-to-one communications in wireless sensor networks," in *Comput. Commun.*, vol.30, Issue 4, pp 863-872, 26 February 2007.
- [6] M. Bakshi, B. Jaumard, and L. Narayanan, "Optimum convergecast scheduling in wireless sensor networks," in *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5650-5661, Nov. 2018.
- [7] J. H. Elrefaie, H. Kunber, M.K. Shaat, A.H. Madian, and M.H. Saad, "Energy-efficient wireless sensor network for nuclear radiation detection," in *Radiat. Res. and Appl. Sciences*, vol. 12, no. 1, pp. 1-9, 2019.
- [8] F. Wang, W. Liu, T. Wang, M. Zhao, M. Xie, H. Song, X. Li, and A. Liu, "To reduce delay, energy consumption and collision through optimization duty-cycle and size of forwarding node set in WSNs," in *IEEE Access*, vol. 7, pp. 55983-56015, 2019.
- [9] Osamy, W., El-Sawy, A.A., and Khedr, A.M., "Effective TDMA scheduling for tree-based data collection using genetic algorithm in wireless sensor networks," *Peer-Peer Netw. Appl.* 13, May 2020, pp. 796-815. <https://doi.org/10.1007/s12083-019-00818-z>
- [10] Louail and V. Felea, "Routing-aware TDMA scheduling for wireless sensor networks," presented in the 12th Annu. Con. on Wireless On-demand Net. Sys. and Serv. (WONS), Cortina d'Ampezzo, 2016, pp. 1-8.
- [11] M. A. Hussain and M. J. Lee, "End-to-End Delay Minimization in Multi-Channel, TDMA Wireless Sensor Networks by Particle Swarm Optimization," in *Int. Conf. on Advances in Comput. and Commun. Eng. (ICACCE)*, Paris, June 2018, pp. 97-104, doi: 10.1109/ICACCE.2018.8441732.
- [12] Y. G. Kim, B. S. Park, and H. H. Choi, "An End-to-End Delay-based Scheduling Algorithm in IEEE 802.15.4e Networks," *Int. J. of Future Gener. Commun. and Netw., IJFGCN*, vol. 9, no. 7, July 2016, pp. 287-296, http://article.nadiapub.com/IJFGCN/vol9_no7/27.pdf.
- [13] Y. Chang, X. Yuan, B. Li, D. Niyato, and N. Al-Dhahir, "Machine-Learning-Based Parallel Genetic Algorithms for Multi-Objective Optimization in Ultra-Reliable Low-Latency WSNs," in *IEEE Access*, vol. 7, 10 Dec. 2019, pp. 4913-4926.
- [14] J. Ma, W. Lou and X. Li, "Contiguous Link Scheduling for Data Aggregation in Wireless Sensor Networks," in *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 7, pp. 1691-1701, July 2014, doi: 10.1109/TPDS.2013.296.
- [15] K. Veeramachaneni, and L. A. Osadciw "Optimal Scheduling in Sensor Networks Using Swarm Intelligence," presented in the CISS 2004: 38th Annu. Conf. on Inf. Sciences and Syst.: Princeton, New Jersey.; Proc., Princeton, New Jersey, pp. 17-19, 2004.
- [16] Rothe P.R., Rothe J.P. "Medium Access Control Protocols for Wireless Sensor Networks," In: Singh P., Bhargava B., Paprzycki M., Kaushal N., Hong W.C. (eds) *Handbook of Wireless Sensor Networks: Issues and Challenges in Current Scenario's, Advances in Intelligent Systems and Computing*, vol. 1132. Springer, Cham, 2020, pp.35-51, https://doi.org/10.1007/978-3-030-40305-8_3.
- [17] Babar Ali, et al., "Study and Analysis of Delay Sensitive and Energy Efficient Routing Approach" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 10(8), 2019. <http://dx.doi.org/10.14569/IJACSA.2019.0100803>.
- [18] AL-wazedi, and A. K. Elhakeem, "Cross layer design using adaptive spatial TDMA and optimum routing for wireless mesh networks," *AEU - International Journal of Electronics and Communications*, vol. 65, Issue 1, January 2011, pp. 44-52.
- [19] Q. Wang, K. Jaffrès-Runser, Y. Xu, J. Scharbag, Z. An, and C. Fraboul, "TDMA Versus CSMA/CA for Wireless Multihop Communications: A Stochastic Worst-Case Delay Analysis," in *IEEE Trans. Ind. Informat.* vol. 13, no. 2, pp. 877-887, 2017.
- [20] amiullah Khan, et al., "Effect of Increasing Number of Nodes on Performance of SMAC, CSMA/CA and TDMA in MANETs" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 9(2), 2018. <http://dx.doi.org/10.14569/IJACSA.2018.090241>.
- [21] T. Kaur, and D. Kumar, "TDMA-based MAC protocols for wireless sensor networks: A survey and comparative analysis," presented in the 5th Int. Wireless Net. and Embed. Sys. (WECON), Rajpura, India, October 2016, pp.1-6,doi: 10.1109/WECON.2016.7993426.
- [22] M. Hashimoto, N. Wakamiya, M. Murata, Y. Kawamoto and K. Fukui, "End-to-end reliability- and delay-aware scheduling with slot sharing for wireless sensor networks," presented in the 8th Int. Conf. on Commun. Syst. and Netw. (COMSNETS), Bangalore, 2016, pp. 1-8, doi: 10.1109/COMSNETS.2016.7439984.
- [23] F. Dobsław, T. Zhang, M. Gidlund, "End-to-End Reliability-Aware Scheduling for Wireless Sensor Networks" in *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 758-767, 2016.
- [24] F. De Rango, A. F. Santamaria, M. Tropea and S. Marano, "Meta-Heuristics Methods for a NP-Complete Networking Problem," presented in the IEEE 68th Vehicular Tech. Conf.(VTC), Calgary, Canada, 21-24 Sep.2008, pp. 1-5, doi: 10.1109/VETECF.2008.279.
- [25] M. Bakshi, B. Jaumard, M. Kaddour and L. Narayanan" On TDMA Scheduling in Wireless Sensor Networks," in *IEEE Canadian on Elec. and Comp. Eng. (CCECE)*, 2016.
- [26] Tejas M. Vala, Vipul N. Rajput, Zong Woo Geem, Kartik S. Pandya, Santosh C. Vora, "Revisiting the performance of evolutionary algorithms," *Expert Systems with Applications*, Volume 175, 2021.
- [27] Cuevas E., Rodríguez A., Alejo-Reyes A., Del-Valle-Soto C. (2021) Metaheuristic Algorithms for Wireless Sensor Networks. In: *Recent Metaheuristic Computation Schemes in Engineering. Studies in Computational Intelligence*, vol 948. Springer, Cham. https://doi.org/10.1007/978-3-030-66007-9_7.
- [28] Kennedy J. and Eberhart, R., "Particle Swarm Optimization," presented in the IEEE Int. Conf. on Neural Netw., Perth, Australia, 1995.

- [29] M. Bagaa , Y. Challal , A. Ksentini , A. Derhab , and N. Badache, "Data Aggregation Scheduling Algorithms in Wireless Sensor Networks: Solutions and Challenges," *IEEE Comms. Surv. & Tut.*, vol. 16, Issue: 3 , pp. 1339 – 1368, 2014 .
- [30] S.C. Ergen, and P. Varaiya, "TDMA scheduling algorithms for sensor networks," *Wireless Netw* (2010), Springer-Verlag, New York; Berlin, Germany; Vienna, Austria, 16:985–997, DOI 10.1007/s11276-009-0183-0.
- [31] Durmaz Incel, A. Ghosh, B. Krishnamachari and K. Chintalapudi, "Fast Data Collection in Tree-Based Wireless Sensor Networks," in *IEEE Trans. Mobile Comput.*, vol. 11, no. 1, pp. 86-99, Jan. 2012, doi: 10.1109/TMC.2011.22.
- [32] Mao J, Wu Z, and Wu X, "A TDMA scheduling scheme for many-to-one communications in wireless sensor networks," in *Computer Communications*, vol. 30, Issue 4, 26 February 2007, pp. 863-872.
- [33] Amdouni, R. Soua, E. Livolant and P. Minet, "Delay optimized time slot assignment for data gathering applications in wireless sensor networks," presented in the *Int. Conf. on Wireless Commun. in Underground and Confined Areas (ICWUCA.2012)*, Clermont Ferrand, united states, 2012, pp. 1-6, doi: 10.1109/ICWUCA.2012.6402475.
- [34] Y. G. Kim and M. J. Lee, "Scheduling multi-channel and multi-timeslot in time constrained wireless sensor networks via simulated annealing and particle swarm optimization," in *IEEE Commun. Mag*, vol. 52, no. 1, pp. 122-129, January 2014, doi: 10.1109/MCOM.2014.6710073.
- [35] C.-Fu Chou, and K.-Ting Chuang, "CoLaNet: A Cross-Layer Design of Energy-Efficient Wireless Sensor Networks," *2005 Systems Communications (ICW'05, ICHSN'05, ICMCS'05, SENET'05)*, Montreal, Quebec, Canada, 2005, pp. 364-369, doi: 10.1109/ICW.2005.35.
- [36] L. Louail and V. Felea, "Routing-Aware Time Slot Allocation Heuristics in Contention-Free Sensor Networks," Springer-Verlag, New York; Berlin, Germany; Vienna, Austria, pp. 271–283, 2016.
- [37] R. V. Kulkarni and G. K. Venayagamoorthy, "Particle Swarm Optimization in Wireless-Sensor Networks: A Brief Survey," in *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 41, no. 2, pp. 262-267, March 2011, doi: 10.1109/TSMCC.2010.2054080.
- [38] M. N. Ab Wahab, S. Nefti-Meziani, and A. Atyabi, "A comprehensive review of swarm optimization algorithms," *PLoS One* 10(5): e0122827, vol. 10, no. 5, pp. 1-36, 2015, e0122827. doi:10.1371/j.pone.0122827.
- [39] GR and Gowrishankar, "An Energy aware Routing Mechanism in WSNs using PSO and GSO Algorithm," presented at the 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN) Noida, 2018, pp. 7-12, doi: 10.1109/SPIN.2018.8474140.
- [40] M. Kumar and V. Gupta, "Benefits of using particle swarm optimization and Voronoi diagram for coverage in wireless sensor networks," presented at the *Int. Conf. on Emerg. Trends in Comput. and Commun. Technol. (ICETCCT)*, Dehradun, 2017, pp. 1-7, doi: 10.1109/ICETCCT.2017.8280300 .
- [41] Metiaf and Q. Wu, "Particle Swarm Optimization Based Deployment for WSN with the Existence of Obstacles," presented at the 2019 5th Int. Conf. on Control, Automat. and Robot. (ICCAR), Beijing, China, 2019, pp. 614-618, doi: 10.1109/ICCAR.2019.8813498.
- [42] P. C. Srinivasa Rao, Prasanta K. Jana, Haider Banka "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks", *Wireless Netw.*, Springer-Verlag, New York; Berlin, Germany; Vienna, Austria, 18 April 2016.
- [43] V. Nagireddy, P. Parwekar, and T. K. Mishra "Velocity adaptation based PSO for localization in wireless sensor networks," *Evol. Intel.* (2018), Springer-Verlag, New York; Berlin, Germany; Vienna, Austria, 14 September 2018.
- [44] Essam H. Houssein, Ahmed G. Gad, Kashif Hussain, Ponnuthurai Nagarathnam Suganthan, Major Advances in Particle Swarm Optimization: Theory, Analysis, and Application, *Swarm and Evolutionary Computation*, Volume 63, 2021.
- [45] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, "Elementary Graph Algorithms," in *Introduction to Algorithms*, 3rd ed., Cambridge, Massachusetts and London, England, MIT Press and McGraw-Hill, 2009, ch.22, Sec. 3, pp. 540–549.
- [46] S. a. Mirjalili, J. S. Dong, A. Lewis and A. S. Sadiq, "Particle Swarm Optimization: Theory, Literature Review, and Application in Airfoil Design," in *Nature Inspired Optimizers Theories, Literature Reviews and Applications*, 1st ed., vol 811, Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland, Springer Nature Switzerland AG, Cham, Switzerland, Springer, 2020, pp. 167-183.
- [47] N. Bulusu, D. Estrin, L. Girod, and J. S. Heidemann, "Scalable coordination for wireless sensor networks: Self-configuring localization systems," in *Proc. Int. Symp. Commun. Theory Appl. (ISCTA)*, Ambleside, UK, 2001, pp. 1–6.
- [48] Chen CH, Chen YP (2011) Convergence time analysis of particle swarm optimization based on particle interaction. *Adv Artif Intell* pp. 1–7, Article ID 204750. doi:10.1155/2011/204750.

Case Study of Self-Organization Processes in Information System Caching Components

Pavel Kurnikov¹, Nina Krapukhina²
National University of Science and Technology MISiS
IT and Automated Control Systems
Moscow, Russia

Abstract—Most modern Information Systems (IS) are designed with Object-Relational Mapping components (ORM). Such components bring down the number of queries to the database server and therefore increase the system performance. Caching mechanisms in software components are complex dynamic systems of open type. A simulation model of cache-application interaction has been created to assess the critical modes of the system. The authors suggest accumulating the cache elements states during the application run at discrete moments of time and present them as multi-variable time series. This work also suggests a method for reconstruction of phase-plane portrait of the system with multidimensional dynamics of the cache elements states. The article shows self-organization processes in caching components of information systems and illustrates the variability of system states for various initial conditions followed by transition to steady-state conditions. In particular, the paper illustrates dissipative structures as well as deterministic chaos with the complete determinism of queries in simulated information systems.

Keywords—Caching mechanisms; ORM; phase space reconstruction; self-organization; dissipative structures; deterministic chaos

I. INTRODUCTION

Modern Information Systems use caching mechanisms, as the process of their design almost always use the technology for mapping relational data into object model [1]. This technology allows leveling down semantic gaps during system development, but also creates additional expenses in processor resource and memory consumption during the operational phase [2, 3, 4, 5].

Works [6, 7, 8, 9] demonstrated that caching mechanisms in IS are non-linear open-type systems. For instance, article [7] identifies the caching element functioning with the Brusselator model, described through a system of differential equations of high order with non-linear terms. Analytical solution of these systems is not always possible though. The research conducted in work [8] revealed statistical relation between dynamic system running modes and the time for processing user's query. The openness of the system lies in information exchange with the environment. The users working with the IS become sources of "energy", and their actions helps build up structural inhomogeneity of the system. In this case "energy" is the flow of queries to the server that enables modification of the IS. It is inevitable that this process results in cache build-up. Cache release is the zone for dispersion of energy which leads to smoothing the structural inhomogeneity of the system.

Open system of such type can absorb external actions and remain in constant change at the same time.

Caching systems have feedback of both positive and negative types. Input flow of queries from a user can change the inner state of the system. In this case the system can give feedback to the input flow to increase or decrease its force, depending on the inner state of the system. In this regard it is important to take into account the development of processes in the caching system itself, as it is not only the external actions that make it change, but also inner processes that bring the system to self-organization.

In the previous works [4, 9] the authors demonstrated the possibility for extraction of cache system modes while IS model is running and expressing the results as multi-variable time series. Based on the research in works [10-14] the authors suggest using phase space reconstruction methods that are based on Takens embedding theorem [15]. This article aims at giving quality assessment to processes in the observed dynamic system. Attractor reconstruction is made on the basis of changes in caching system variables available for observation in time.

The rest of the paper is organized as follows: Section II briefly describes the caching mechanisms in ORM systems. Section III presents a simulation model of interaction between the caching mechanism and the application. Sections IV and V present the research results and their discussion, respectively.

II. CACHING MECHANISM

Modern ORM frameworks (object-relational mapping) build IS conceptual layer, and then every stored table is assigned an object from the conceptual model. Caching mechanism [16] handles a list of entities, each in one of the following states:

- Detached (from the cache).
- Added.
- Unchanged.
- Modified.
- Deleted.

The object state manager maintains the state of each entity in the cache – detached (from the cache), added, unchanged, modified, deleted and tracks their state transitions.

After the users' interaction with the information in IS the process of transformation of modified entities corresponding with the stored tables of the database is initiated by method SaveChanges. A list of changes is then formed for each entity. This list is formed straight from cache. Next, the list is transformed into a series of algebraic expressions (relational operations with the tables corresponding to the entities) to be realized on the database server [16]. After the transactions are complete, cache mode is synchronized with the new state of the database.

This article uses discrete-event model for simulating scenarios of multiple users interacting with the database via high-level interface [9]. Such a model allows simulating the load of several users on a server and serializes the cache modes in discrete moments of time.

III. DISCRETE-EVENT MODEL

Let us take as R (the data stored in the database) the multitude of stored states of the initial database. ORM creates an entities scheme E (entity schema) from the database scheme (relational schema). Let us assign entity E on the object layer of the application to each object from R . There is a cache vector C for each type of entity. Test on range of entities is defined by predicate formula from range F^E . Transformation between relational and object data is done through bi-direct connection illustrated by formulas 1 and 2.

$$\Delta R_i = \Delta UV(E_i, \Delta^{+/-} E_i), i = 1, \dots, I, \Delta^{+/-} E_i \neq \emptyset \quad (1)$$

where $\Delta^{+/-} E$ is a result of the test from range F^E on entity E_i .

Here the difference in entities state is assessed with *UpdateViews*, which shows how to extract information about the modified state of the stored database [16]. (In Step 1, view maintenance algorithms are applied to update views. This produces a set of delta expressions, *UpdateViews*, which tell us how to obtain delta tables from delta entities and a snapshot of Entities).

In the case when conceptual (entity) level test from F has not modified the entity due to the presence of the modification object in cache, there will be no query to the server. (no update is required).

$$\Delta E = QV(R), \quad (2)$$

here QV – *query views* transforms objects from relational scheme to entities; (The query views express entities in terms of tables).

The general rule which a user uses to work with the common resource C can be presented as:

$$C_i(t_k) = UpdateObjectCache(F_j^E(E_i, t_{k-1}) - F_j^E(E_i, t_k)), \\ j = 1, \dots, J, k = 1, \dots, K \quad (3)$$

This process can be compared to negative and positive feedback. If the list of entities changes received from cache after a certain input test is empty, there will be no query sent to the server. This equals restriction of input flow (negative

feedback). On the other hand, if the result of a certain test run lead to cache flush, that can potentially increase the input flow, and this will result in extra costs for the server.

Let us define a user from range $U = \{u_1, \dots, u_M\}$. A user is defined by a subset of tests $F_{U_m}^E \subset F_E$ which he works with as a full group of disjoint events with probabilities $\langle P \rangle$ as well as a law of test distribution ω and time of work T with the system:

$$U = \langle F_U^E, P, \omega, T \rangle \quad (4)$$

In other words, user U_m is a source of events flow in $[0, T]$. The events flow is an external influence towards the system. Let us set a multi-layer density of probability for intervals between test completion time:

$$\omega_{U_m}(\tau_1, \dots, \tau_K), \tau_k = t_k - t_{k-1}, t_0 = 0, t_K = T \quad (5)$$

where ω_{U_m} is an n -dimensional density function for user U_m defining the law of random vector distribution $\|\tau_k\|, k = 1, K$ for user U_m .

Multi-dimensional random variable is simulated on an ECM with Neumann construction generalized for a multi-dimensional event [17]. The choice of test to be run by a user at the moment of time t_k is made based on a random discrete variable generated in the interval $[0, 1]$:

$$X = \begin{pmatrix} x_1 & \dots & x_L \\ p_1 & \dots & p_L \end{pmatrix}, \sum_{l=1}^L p_l = 1 \quad (6)$$

where $P = \{p_1, \dots, p_L\}$ is a probability vector for a full set of events (tests from range F_U^E) with indexes $\{1, \dots, L\}$.

Test with index x_l where $r_k \in \Delta_l$ will be performed at the moment in time t_k .

$$r \in \begin{cases} \Delta_1, & r \leq p_1 \\ \Delta_l, & p_1 + \dots + p_{l-1} \leq r \leq p_1 + \dots + p_l, (l \geq 2) \end{cases} \quad (7)$$

where Δ_l is a part of a segment $[0, 1]$ with length equal p_l

Server part storing the database is a shared resource for the system. Inner state of the system is cache vector C . This resource is also shared for every user from U .

Let us analyze the dynamics of cache vector C . For this we will inquire the state of vector C at $[0, T]$ every Δt . As a result we have a phase space described by D -dimensional time series X_1, X_2, \dots, X_D where $X_d = (x_{d,1}, x_{d,2}, \dots, x_{d,K})$, $d = 1, 2, \dots, D$, $K = T / \Delta t$, $x_{d,k}$ is the size of d element of cache C at moment in time $k\Delta t$. $D = \dim(C)$. For the systems containing D variables $x_d(t) = [x_{d,1}, x_{d,2}, \dots, x_{d,N}]^T \in \mathbb{R}^D$, $d = 1, \dots, D$ phase space is suggested to be simulated with the method described in work [13]. This approach allows quality simulation of phase space from initial data, and it also has relatively low

calculation complexity compared to the methods described in works [10, 11]. The process involves consecutive calculation of time lapses [18, 19] for each series. Correlation dimensions [20, 21] are found for each variable based on the calculated parameters.

IV. RESULTS

The authors suggest using the previous results [4, 9] for simulating caching mechanisms in the IS. Architecture of the database on the server is illustrated by Fig. 1.

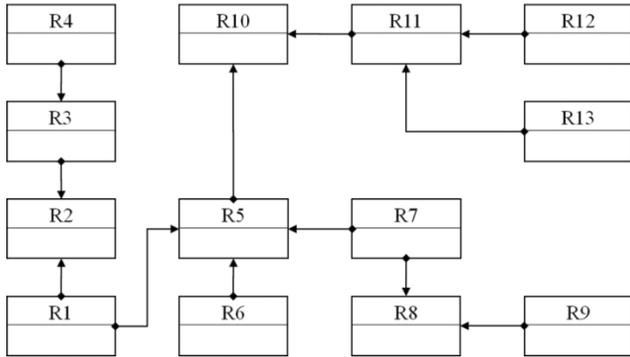


Fig. 1. Database Diagram.

The database consists of 13 tables with multiple links. Three basic tests were taken as functional base: adding data into a table or a set of tables (*insert_data*), removing data from a table or a set of tables (*remove_data*), and getting data from a table or a set of tables (*get_data*). Initial conditions of the model are the number of users working with the system simultaneously. Every user is described by a set of tests which he works with, as well as a function of their distribution in the interval $[0, T]$. A set of tests for each user is a group of disjoint events. Each event is defined by its probability. Another initial condition is the state of cache vector before the run of the model. Initial state of the database was the same for all models described in this work.

Let us look at the model with one user (U_0) working with one database table (R_7). Let us identify the set of test for the user (U_0) as a sub-set $\{insert_data, get_data\}$. Time for modeling was 1000 seconds, Table I.

TABLE I. INITIAL CONDITIONS FOR MODEL 1

Model	Users	Tables	Distribution of basic tests			Time of simulation T, sec
			Insert data	Remove data	Get data	
1	1	R7	0.5	0	0.5	1000

Dynamics for cache size associated with table R_7 is shown in Fig. 2.

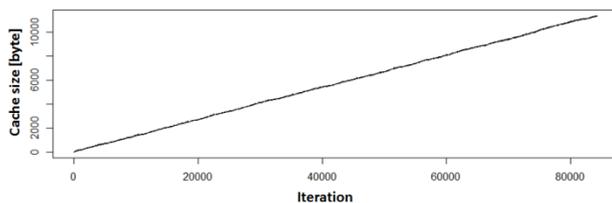


Fig. 2. Cache Dynamics for Table R7 (Model 1).

Cache dynamics for table R_7 is presented by linear growth of space. It is related to the fact that non-linear processes do not occur in single-user mode in the scenario with one object. Additionally, as framework 1 shows obvious domination of *insert_data* and *get_data* tests, each query for an entity is accompanied with caching the object. In 1000 seconds of simulation more than 80,000 queries to cache were executed. The total cache volume reached 11347 bytes. In other words, there is no reaction of the system to external factors, because there is no release of cache and input flow to the database server does not get cut. System does not compensate for these processes and stays in an unbalanced state, meaning there is no self-organization.

Let us simulate a model with three users. The object of the queries is still *table R7*. Let us add test *remove_data* to the *model 1* set of tests. New initial conditions are shown in *Table II*.

TABLE II. INITIAL CONDITIONS FOR MODEL 2

Model	Users	Tables	Distribution of basic tests			Time of simulation T, sec
			Insert data	Remove data	Get data	
2	3	R7	0.3	0.35	0.35	1000

It is worth mentioning that Table II only shows the general distribution of test running in the model during time period T . Fig. 3 illustrates the cache growth dynamics, additionally there are visuals on results of calculations for correlation dimension and time lapse for the simulated attractor. All the calculations were done in RStudio in language R with non-linear Tseries package. Overview of numeral methods for this package is presented in work [22].

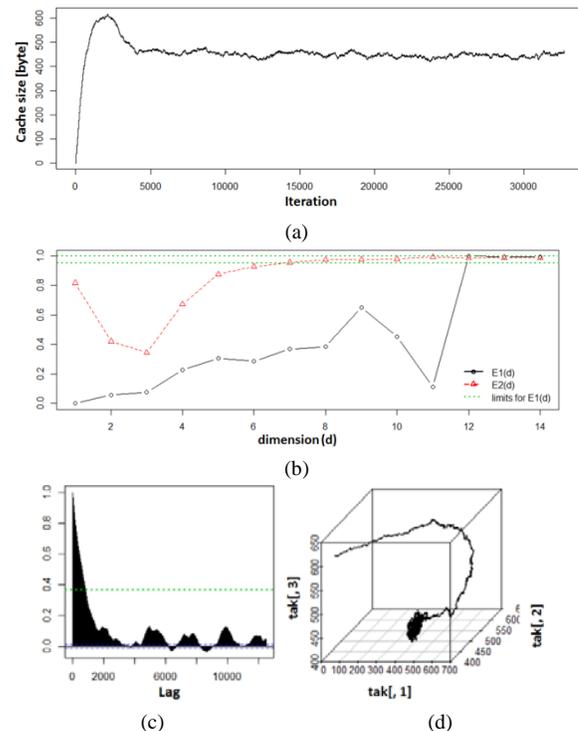


Fig. 3. Simulation Results for 3 users in Accordance with Initial Data for Table II: (a) Cache Dynamics for Table R7; (b) Auto-Correlation Function; (c) Embedding Dimension Calculation; (d) Attractor Visualisation.

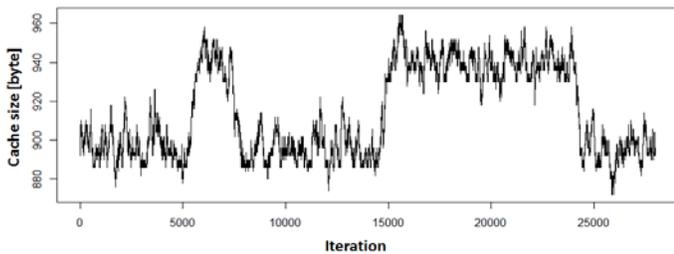
Here non-linear dynamics in cache behavior can be seen. This is due to parallel input flows of queries to the system and inner non-linear caching mechanism. Apart from that, this model shows stable increase of allocated memory, as there is a cache release mechanism at work. This mechanism is a zone for information dispersion, and as a result of its work the system structural inhomogeneity is smoothed. Open system of such type processes external influence, but remains in permanent change. Peak memory use in this case only reached 657 bytes. This system is in unstable stationary mode of work.

Let us add another influence to demonstrate the ‘fragility’ of this balance, as shown in Table III. In *model 3* the new external influence to caching mechanism is user *U4* with an only test *get_data*. This time the object of queries will be new table R8 linked to R7 by the key. Time of simulation was kept unchanged. Query *get_data* to object R8 linked to R7 will introduce changes to cache associated with object R7. This will lead to the increase in query flow to the system.

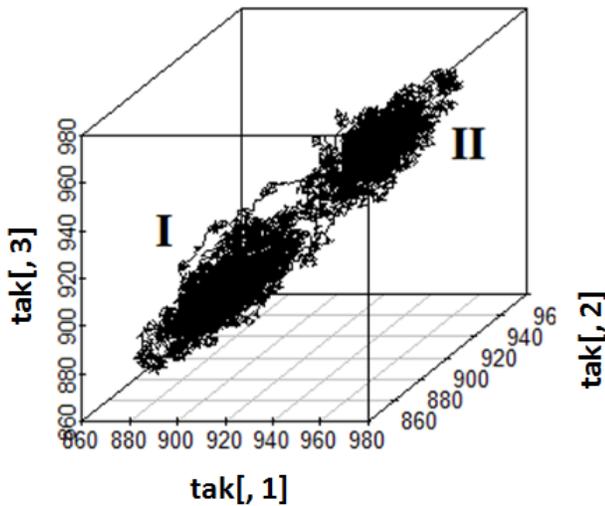
TABLE III. INITIAL CONDITIONS FOR MODEL 3

Model	Users	Tables	Distribution of basic tests			Time of simulation T, sec
			Insert data	Remove data	Get data	
3	4	R7, R8	0.3	0.35	0.35	1000

Cache dynamics for table R7, R8 is shown on Fig. 4.



(a)



(b)

Fig. 4. Simulation Results For 4 users in Accordance with Initial Data for Table III: (a) Shared Cache Dynamics; (b) Attractor Visualisation.

Cyclic transitions of the dynamic system from zone I to zone II are shown in Fig. 4. The system has two attractors. The received data is not stationary as the system is bi-stable. This is an open system, so part of external queries is redirected to cache release, which results in caching mechanism self-regulation.

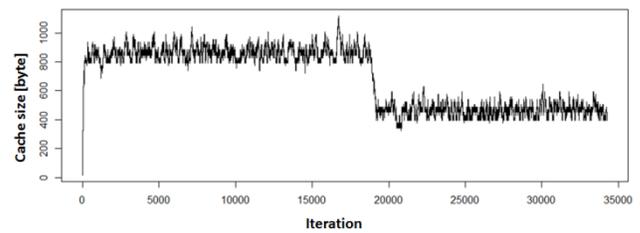
Let us increase input queries flow to the system by adding new users while keeping the test distribution and query objects the same, see Table IV.

TABLE IV. INITIAL CONDITIONS FOR MODEL 4

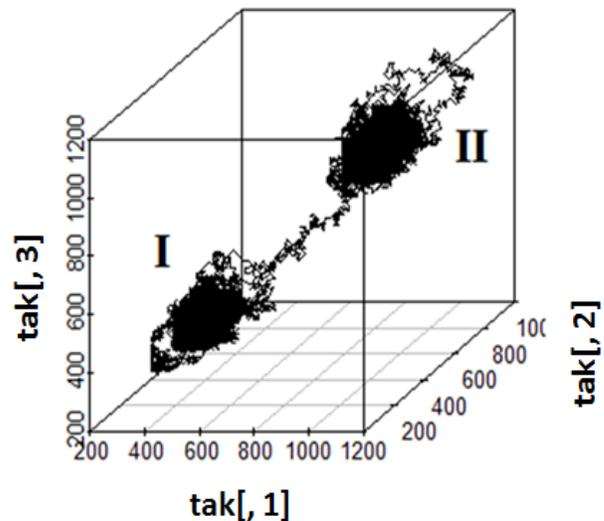
Model	Users	Tables	Distribution of basic tests			Time of simulation T, sec
			Insert data	Remove data	Get data	
4	10	R7, R8	0.3	0.35	0.35	1000

Increase in number of users from 4 to 10 affected the total number of completed transactions on the server and queries to cache (Fig. 5a). This did not though change the number of equilibrium conditions of the system (Fig. 5b).

Let us look at the system with two groups of users. Within its group each user will be described by a group of tests it interacts with, as well as a unique function of their distribution in the interval $[0, T]$. Let us also change the objects of the users' queries, see Table V.



(a)



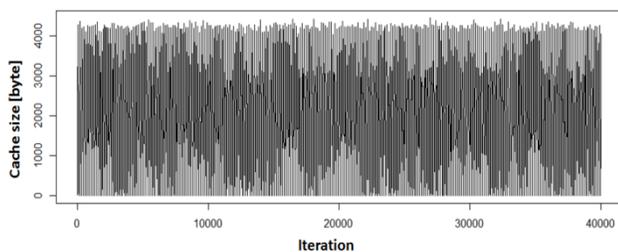
(b)

Fig. 5. Simulation Results for 10 users in Accordance with Initial Data for Table IV: (a) Shared Cache Dynamics; (b) Attractor Visualisation.

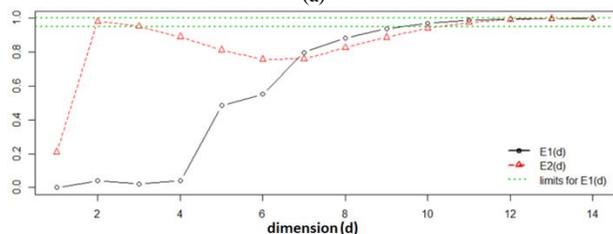
TABLE V. INITIAL CONDITIONS FOR MODEL 5

Model	Users	Tables	Distribution of basic tests			Time of simulation T, sec
			Insert data	Remove data	Get data	
5	1-4	R11 – R13	0.35	0.3	0.35	5000
	5	R1 – R4	0.2	0.4	0.4	

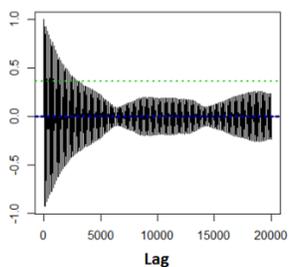
Here a group of users U1-U4 interacts with tables R11-R13. User U5 interacts with tables R1-R4.



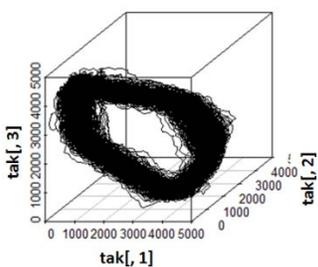
(a)



(b)



(c)



(d)

Fig. 6. Simulation Results for 20 users in Accordance with Initial Data for Table V: (a) Shared Cache Dynamics; (b) Auto-Correlation Function; (c) Embedding Dimension Calculation; (d) Attractor Visualisation.

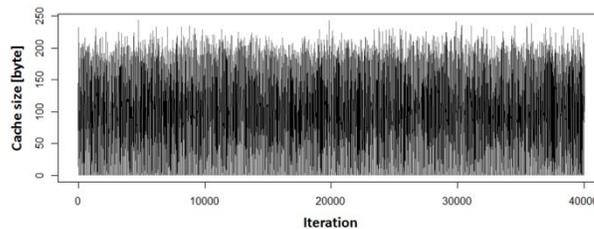
Despite the presence of inquiries to linked entities (R12, R13 and R1, R3, R4), Fig. 6d. shows closed-loop isolated trajectory in phase space corresponding with cyclic movement. This movement is asymptotically stable and in the process of self-organization the system will be attracted to this trajectory regardless of the initial cache state. Despite minor presence of chaotic processes caused by a large number of users interacting with the system, they do not disturb the end cycle. In other words, the system is far from the point of stochastic bifurcation.

Let us introduce additional inquiry objects to the system, as well as new groups of users for running the tests. The time of simulation will be kept unchanged, see Table VI.

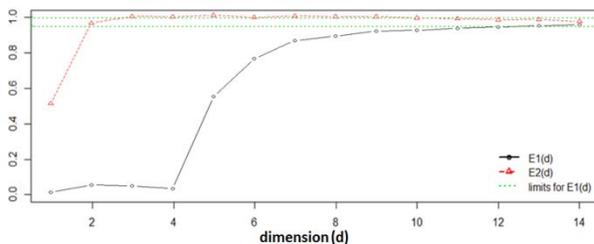
TABLE VI. INITIAL CONDITIONS FOR MODEL 6

Model	Users	Tables	Distribution of basic tests			Time of simulation T, sec
			Insert data	Remove data	Get data	
6	1-4	R11–R13	0.35	0.3	0.35	5000
	5	R1–R4	0.2	0.4	0.4	
	6-14	R1, R5, R6	0.25	0.25	0.5	
	15-20	R1, R5, R6	0	0	1	

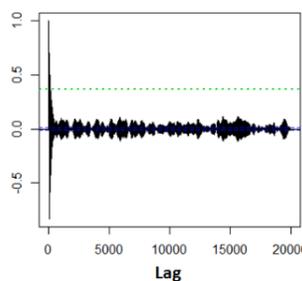
New groups of users U6-U14 and U15-U20 interact with tables R1, R5, R6, but the latter only accesses the database for information without making any changes.



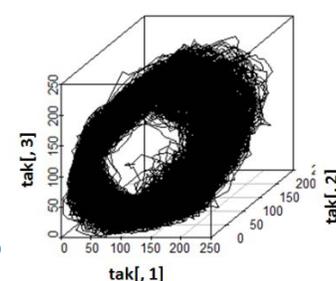
(a)



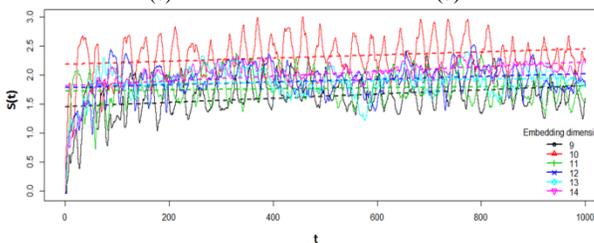
(b)



(c)



(d)



(e)

Fig. 7. Simulation Results for 20 users in Accordance with Initial Data for Table VI: (a) Shared Cache Dynamics; (b) Embedding Dimension Calculation; (c) Autocorrelation Function; (d) Attractor Visualisation; (e) Estimating Maximal Lyapunov Exponent.

Information dispersed and injected into the system is strictly compensated, despite minor presence of chaotic

processes caused by a large number of users interacting with the system. This points at dissipative type of the observed system. Fig. 7(e) demonstrates the estimates for maximal Lyapunov exponent [23, 24, 25] for space several special dimensions. Despite the high dispersion of this exponent, it has positive value which is another confirmation for the presence of chaotic processes.

V. CONCLUSION AND FURTHER RESEARCH

A discrete-event model was developed as a result of this research. The model allows simulation of various scenarios for users interacting with the database server through caching mechanisms of object-relative representation. Data on states of each system cache valuable in time were reflected as variables in multi-dimensional time series. Each component of the time series was assigned necessary parameters for phase space reconstruction. Analysis of trajectories data revealed the mode of work for simulated dynamic systems.

The authors demonstrated linear processes of cache dynamics for single user (*U0*) interacting with one database table (*model 1*). With the randomly requested data in test *get_data* and the absence of *remove_data* all the entries were fully cached. In case of collective work of users with the system, combined with non-linear connections in caching mechanism, the latter demonstrates complex non-linear dynamics both on macro- and micro-levels. This is the key element in self-organization process.

Unstable stationary dynamic mode was demonstrated for the system that three users were interacting with (*model 2*). Set of tests and distribution were selected in a way that the model does not show stable increase of allocated memory. It can be described as dissipation of information by the caching mechanism, which leads to the open system of this type processing external influence, while staying in permanent change.

Non-equilibrium of the stationary state of *model 2* was confirmed by the data from *model 3* run, where user *U4* with the sole test *get_data* became the additional external influence on the system. The system demonstrated bi-stable transitions between two attractors I and II, but it still remained in constant self-organization. Adding new users in *Model 4* affected the total number of transactions completed on the server during the simulation time *T*, but it did not affect the number of equilibrium conditions of the system.

Models 5 and 6 demonstrated presence of self-oscillating processes. High load on the caching mechanism was provided by several groups of users simultaneously interacting with most of the objects of the initial database. As a result, the system had stable limit cycle without a stable focus. Observations of *model 5* showed that convergence of phase trajectory does not depend on the initial state of cache. Besides, even though the users in *model 6* had determined behavioral rules, non-linear dynamic system was subject to determined chaos, which was revealed with calculating highest Lyapunov exponent. Despite this, such behavior of caching mechanism is an optimum option for multi-user regime of work with IS due to the system's resistance to initial conditions.

To reproduce a larger number of information systems, one should expand the functional test base of user-application server interaction. Research into self-organization processes in caching components of information systems will allow comparative analysis of function composition at IS development stage and enable choice of a more stable version.

REFERENCES

- [1] S. Bhatti, Z. Abro, F. Rufabro. "Performance evaluation of java based object relational mapping tool," Mehran University Research Journal of Engineering and Technology, 2013, pp. 159-166.
- [2] A. Gruca, P. Podsiadlo. "Performance Analysis of NET Based Object-Relational Mapping Frameworks," International Conference: Beyond Databases, Architectures and Structures, Springer, Cham, 2014, pp.40-49.
- [3] R. Jacobson, S. Misner, Microsoft SQL Server (TM) 2005 Analysis Services Step by Step, Microsoft Press, 2006.
- [4] N. Krapuhina, P. Kurnikov, I. Tarkhanov. "Multi-criteria method of estimating performance ORM components in information systems," Trudi Instituta sistemnogo analiza Rossiyskoy akademii nauk ISA RAN, 65(22), 2015, pp. 105-109.
- [5] D. Zmaranda, L. Pop-Fele, C. Gyorodi, R. Gyorodi, G. Pecherle. "Performance comparison of crud methods using net object relational mappers: A case study," International Journal of Advanced Computer Science and Applications, (11), 2020.
- [6] S. Melnikov. "Formation of dissipative structures in ORM-components of high-loaded portals," Modern problems of science and education. Surgery, 6(5), 2013.
- [7] S. Kudzh, S. Melnikov. "Unsteady operation of a software system with the active "caching" to access the database," Perspectives of Science and Education, 8(2), 2014, pp. 51-55.
- [8] S. Melnikov. "Dissipative structures of the "cycle" in the software, which exists in the mode of multi-threaded execution of the executable code," Modern problems of science and education. Surgery, 10(6), 2012.
- [9] P. Kurnikov, N. Krapuhina. "Phase space reconstruction of high-loaded caching mechanism dynamics in information systems," Journal of information technologies and computing systems, 2019, pp. 49-65.
- [10] S. Garcia, J. Almeida. "Multivariate phase space reconstruction by nearest neighbor embedding with different time delays," Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 72(2), 2005.
- [11] Y. Hirata, H. Suzuki, K. Aihara. "Reconstructing state spaces from multivariate data using variable delays," Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 74(2), 2006.
- [12] J. Barnard, C. Aldrich, M. Gerber. "Embedding of multidimensional time-dependent observations," Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, 64(4), 2001.
- [13] R. Wang, J. Gao, Z. Gao, X. Gao, H. Jiang, L. Cui. "Data Fusion Based Phase Space Reconstruction from Multi-Time Series," International Journal of Database Theory and Application, 8(6), 2015, pp. 101-110.
- [14] S. Palit, S. Mukherjee, D. Bhattacharya. "A high dimensional delay selection for the reconstruction of proper phase space with cross auto-correlation," Neurocomputing, 113 (2013), pp. 49-57.
- [15] F. Takens. "Detecting strange attractors in turbulence, Dynamical systems and turbulence," Warwick 1980, Springer, Berlin, Heidelberg, 1981, pp. 366-381.
- [16] A. Adya, P. Bernstein, S. Melnik. "Generation of query and update views for object relational mapping," U.S. Patent No. 7,647,298. Washington, DC: U.S. Patent and Trademark Office, 2010.
- [17] V. Bykov. "Digital modeling in statistical radio engineering," STIN, 75 (1973), p. 10294.
- [18] J. Theiler. "Spurious dimension from correlation algorithms applied to limited time-series data," Physical Review A, 34(3), 1986, p. 2427.
- [19] H. Kantz, T. Schreiber. "Nonlinear time series analysis," vol. 7, Cambridge University Press, 2004.
- [20] D. Turcotte. "Fractals and Chaos in Geology and Geophysics," Cambridge University Press, 1992.

- [21] G. Malinetskii, A. Potapov, A. Rakhmanov. "Limitations of delay reconstruction for chaotic dynamical systems," *Physical Review E*, 48(2), 1993, p. 904.
- [22] T. Parker, L. Chua. "Practical Numerical Algorithms for Chaotic Systems," *Mathematics of Computation*, 56(193), 1991, pp. 377-379.
- [23] L. Arnold, W. Kliemann, E. Oeljeklaus. "Lyapunov exponents of linear stochastic systems," In *Lyapunov exponents*. Springer, Berlin, Heidelberg, 1986, pp. 85-125.
- [24] M. Rosenstein, J. Collins, C. De Luca. "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D: Nonlinear Phenomena*, 65(1-2), 1993, pp. 117-134.
- [25] J. Eckmann, S. Kamphorst, D. Ruelle, S. Ciliberto. "Lyapunov exponents from time series," *Physical Review A*, 34(6), 1986, pp. 4971-4979.

Dynamic Management of Security Policies in PrivOrBAC

Jihane EL MOKHTARI¹, Anas ABOU EL KALAM², Siham BENCHADDOU³, Jean-Philippe LEROY⁴

LISER Laboratory, IPI, Paris, France^{1,4}

LRI Laboratory, ENSEM, Casablanca, Morocco^{1,3}

TIM Laboratory, ENSAM, Marrakesh, Morocco²

Abstract—This article is a continuation of our previous work on identifying and developing tools and concepts to provide automatic management and derivation of security and privacy policies. In this document we are interested in the extension of the PrivOrBAC model in order to ensure a dynamic management of privacy-aware security policies. Our approach, based on smart contracts (SC) and the WS-Agreement Specification, allows automatic agents representing data providers and access requesters to enter into an access agreement that takes into consideration not only service level clauses but also security rules to protect the privacy of individuals. Our solution can be deployed in such a way that no human intervention is required to reach this type of agreement. This work shows how to use the WS-Agreement Specification to set up a process for negotiating, creating and monitoring Service Level Agreements (SLAs) in accordance with a predefined access control policy. This article concludes with a case study accompanied by a representative implementation of our solution.

Keywords—Access control; privacy; PrivOrBAC; PrivUML; smart contract; WS-agreement

I. INTRODUCTION

Our work is related to the management and development of security and privacy policies. Our goal is to be able to identify security and privacy needs and integrate them early into the design of complex systems. This practice has yet to be widely adopted; generally, security requirements are expressed separately in the form of complementary modules and integrated late into the systems. The principal difficulty which designers encounter concerns the methods and tools used to design functional and security models, as well as to describe the business demands which the system must satisfy. While functional design methods have become increasingly effective so as to allow one to design models formal enough to provide a base of refinement down to code, security design methods are not yet at the same level.

As a first step, we focused our efforts on defining the security policy and modeling tools allowing the integration of privacy protection into the system's design models. It is in this context that we relied on PrivOrBAC [1], an access control model dedicated to the protection of privacy. It is based on the OrBAC model [2] which introduced the notions of organization, context and object views as attributes of access management. Thus, access permission is granted to a role within an organization to perform an activity on an object view in a specific context. PrivOrBAC takes this logic and enriches it with attributes specific to privacy. Access permission in

PrivOrBAC is therefore granted to a role in an organization to perform an activity, justified by a purpose in a specific context, on a granular object view following the consent of its priority. To model PrivOrBAC, we have proposed the new PrivUML metamodel [3] which is an extension of the UML language, and which makes it possible to integrate all the attributes of PrivOrBAC necessary for the protection of privacy into a class diagram. Our implementation of PrivUML under XACML (eXtensible Access Control Markup Language) [4] allowed us to set up a privacy-aware security policy. However, at this level, our solution remains static.

The second part of the article is therefore devoted to automating the management of this policy to make the access control process dynamic. As it stands, any modification to be made in the security policy requires manual interventions, which diverge from the aspirations of complex systems known by their high rates of interactivity and which are therefore penalized by the obligation of human involvement in this type of task. The idea is to put in place intelligent agents capable of replacing human skills. These agents will not only update the security policy, but also manage the entire upstream process responsible for investigating the incoming request and negotiating, creating and monitoring the access agreements in accordance with the policy. It is in this sense that we propose to set up smart contracts managed through the WS-Agreement Specification [5].

We organize the remainder of this article into three sections. The first is dedicated to the presentation of our previous work. We devote the second section to the mechanism for automating the management of our security policy based on smart contracts and the WS-Agreement Specification, and then we present a discussion of our contribution in the third section.

II. PREVIOUS WORK

Our previous work, [3] and [6], focused on providing the means and tools to integrate the principles of privacy protection into IT systems. To achieve our objectives, we have opted, throughout this work, for a certain number of choices which we present in what follows.

1) *PrivOrBAC (Privacy-Aware Organization Based Access Control*: In OrBAC model [2], the organization is the central component; privileges are not applied directly to subjects, they are assigned to roles within an organization [7]. Permission is granted for a role to perform a subset of activity on a number of views of an organization's objects in a specific

context. PrivOrBAC [1] extends the OrBAC model to covers Privacy management requirements. The first change consists in setting up a hierarchy of views to better manage the granularity of the data consulted. The same view of objects can have different levels of detail from one organization to another. Regarding consent and purpose, they were introduced as new attributes of the context. The organization in PrivOrBAC continues to play its role as a central component which makes it possible to configure the access policy according to the other abstract entities (Role, Activity, View and Context). The role and activity entities behave the same as in OrBAC; privileges are assigned to a role within an organization in order to perform an activity.

2) *PrivUML Metamodel*: In accordance with the MDA approach (Model Driven Architecture), we have defined our CIM model (Computation Independent Model) composed of the pair (ends, means) corresponding to the protection of privacy through the PrivOrBAC access control model. The next step is to go to the second level of MDA which is the PIM (Platform Independent Model). Translating CIM to PIM requires finding the modeling tool capable of integrating all of our Privacy requirements into the model. After a study on the various security modeling tools focused on modeling access control requirements (SecureUML [8], UMLSec [9], PaML [10] and Privacy UML Profile [11]), we noted that none of these tools is completely adapted to our needs. Based primarily on roles and enriched by purpose, they do not, in their current state, allow the PrivOrBAC access control to be modeled. We therefore proposed PrivUML [3], which is an extension of UML enriched by the following concepts:

- The access modalities allowing to express the authorization of access, the refusal or the obligation.
- The hierarchy of object views to control the granularity of the data to be consulted.
- The purpose justifying the access request.
- The explicit consent of the data owner granted on the basis of the purpose.

We have therefore set up a meta-model capable of integrating security policies, meeting privacy requirements, from the design phase of the system. In PrivUML, an access request is refused or accepted with or without conditions to a subject, having a role in an organization, to perform an action justified by a purpose on a specific set of data whose owner has given his prior consent.

III. AUTOMATING OF THE MANAGEMENT AND PROTECTION OF PRIVACY

The work presented in the previous section makes it possible to set up a privacy policy in a static manner. Any new changes require manual updating of this security policy. Complex systems therefore have to dedicate personnel to this task, the cost of which increases according to the volumes of their interactions. This also increases the risk of error and data leaks. Take for example the French Health Management

System (FHMS) which manages all the health information of adherents. Any medical entity not listed in the system must first go through the stage of negotiating a new agreement access, which must comply with the security policy in force, and which will result in updating the security policy of FHMS. The implementation of smart contracts is necessary to ensure automatic management of requests for negotiation and establishment of access contracts. This management method will be piloted by intelligent agents responsible for negotiating contracts as well as updating security policies. In what follows, we will define the notion of smart contract, then we will present the solution chosen for the management of smart contracts, and finally we will explain through a case study how we ensured the automation of management and the protection of privacy in an interactive system based on the elements mentioned above.

A. Smart Contract

The concept of a smart contract first appeared in the 1990s and has evolved over the years. Programmers tend to see it as a solution that replaces traditional contracts and contract law [12]. Another vision of the smart contract considers it as a mechanism to express calculations on a blockchain [13]. The author in [14] defines the smart contract as an automatable and enforceable agreement: automatable, although some parts may require human intervention and control; and enforceable either by legal application of rights and obligations, or by tamper-proof execution of computer code. For our work, we agree with [15] on its definition of the smart contract and consider it as the formalization of an agreement, the terms of which are automatically applied by relying on a transaction protocol, while minimizing the need for an intermediary.

B. Access Agreement

Data providers and consumers operate in a dynamic environment governed by a set of rules, conditions, obligations and guarantees formalized in a contract or access agreement. Several solutions exist for the management of this type of electronic contract, from negotiation to monitoring. Among these solutions we find:

- WSLA (Web Service Level Agreement): this specification, created by IBM in 2003 [16], allows the creation and monitoring of SLA (Service Level Agreement) contracts. WSLA defines a flexible and extensible language to allow consumers and providers of web services to define and specify a set of parameters such as technical-functional descriptions of web services, supplier commitments, etc.
- WSOL (Web Service Offering Language): this specification, based on XML language, allows suppliers to define several service levels or SLOs (Service Level Objectives) for the same web service. A consumer can therefore choose the desired web service and select the SLOs that interest him according to the level of service that suits his needs [17].
- WS-Agreement (Web Service Agreement): this specification defines a protocol and a language for negotiating, renegotiating, creating and monitoring bilateral access contracts (between consumer and

provider) in distributed systems [5]. The WS-Agreement Specification and the WS-Agreement Negotiation Protocol [18] are the only specifications standardized and accepted by a large community [19], [20]. Therefore, our study will be based on this specification to establish the protocol for negotiating and creating smart contracts that frame access agreements.

1) *WS-Agreement Specification*: WS-Agreement is a specification that enables exchanges based on XML language between providers and consumers of web services. An agreement is an XML file created from a template. The agreement and the template have the same structure. An agreement is made up of three sections:

- Name: this section contains the name and ID of the agreement.
- Context: the context contains meta-information about the agreement such as the identifiers of the initiator and the responder of the agreement, the identifier of the template that served as the basis for creating the agreement, references to other agreements, the period of validity, etc.

- Terms: The terms of an agreement include terms of service and, optionally, guarantee terms that define the constraints applicable to the services set out in the agreement. A service term is made up of several sections of Service Descriptions Term (SDT), Service References (SR) and Service Properties (SP). A guarantee term for a service consists of a section describing the scope of the service, a section describing the service level objectives (SLO) and possibly a section describing the business values of the service.

2) *WS-Agreement Negotiation Protocol*: Automating negotiations requires formalizing the definition of each term of the access contract. Much work has been done in recent years to automate contract negotiations in various fields such as connected objects ([19], [21], [22]), cloud computing platforms ([23], [15], [24], [25]) as well as distributed environments ([20], [26], [27], [28]). The WS-Agreement Negotiation Model [18], shown in Fig. 1, consists of three layers: the negotiation layer, the agreement layer, and the service layer.

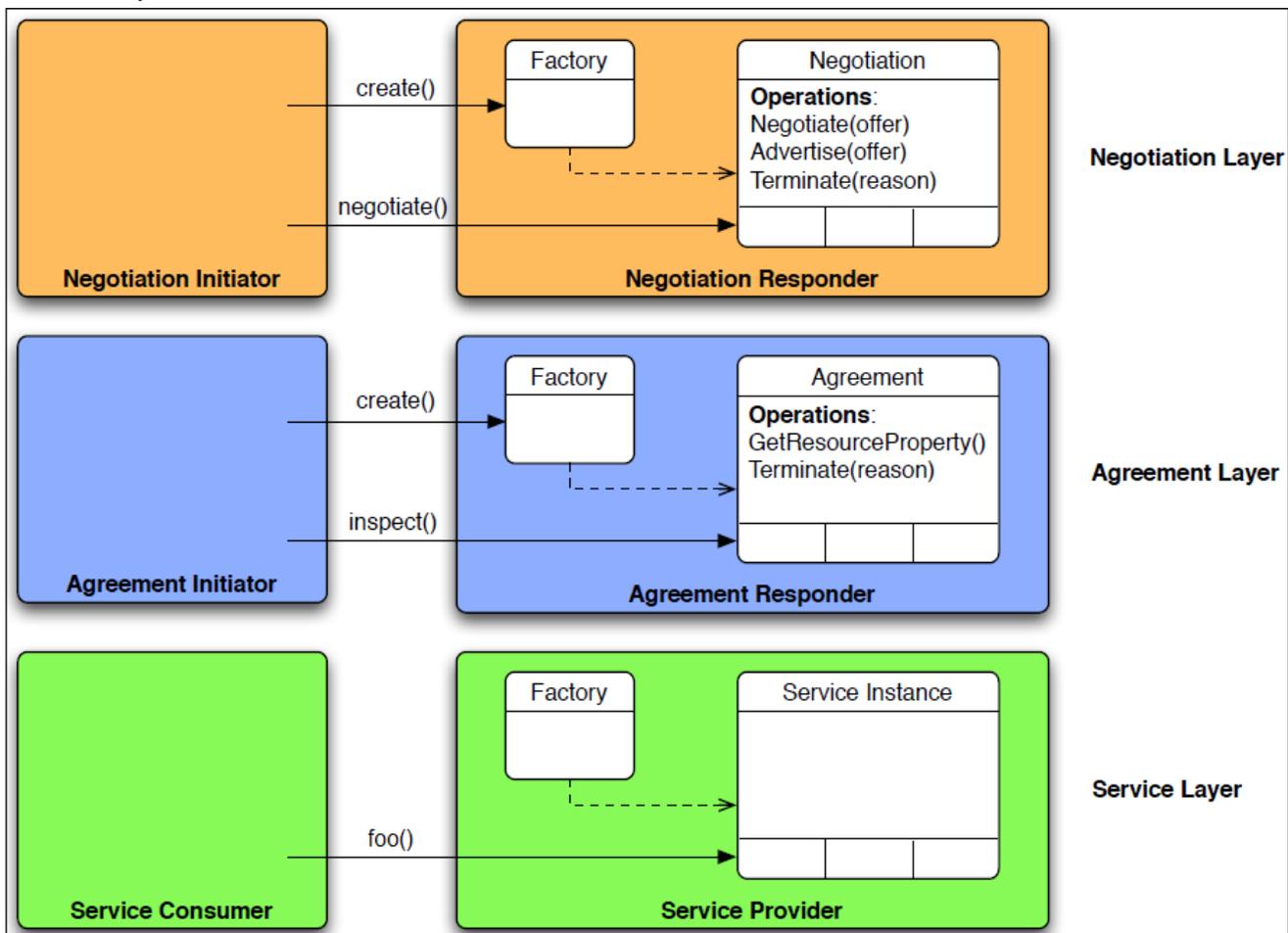


Fig. 1. WS-Agreement Negotiation Model [18].

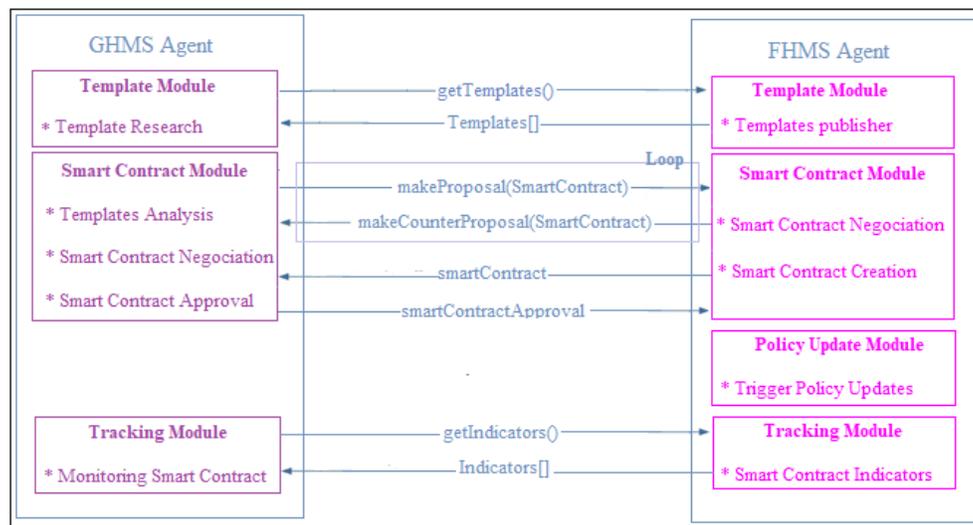


Fig. 2. Protocole De Négociation De Smart Contrat.

The negotiation layer provides a protocol and language for negotiating offers that indicate the willingness of both parties to enter into a subsequent agreement. The negotiation process includes the exchange of negotiating offers and counter-offers until an agreement is reached between the two parties. The agreement layer, on the other hand, provides a protocol and language that provides the basic functionality to create and monitor agreements. Finally, the service layer defines the services offered by the data provider. The execution of the service on the service layer is governed by the agreement layer. Fig. 2 illustrates the protocol for negotiating a smart contract between the French Health Management System (FHMS), as data provider, and the German Health Management System (GHMS) as data consumer.

The consumer agent (GHMS) initiates the negotiation process by requesting the retrieval of the services modeled, described and published by the data provider, FHMS, by calling the public "getTemplates()" method. Once the templates have been received, the consumer agent analyzes them and chooses the one that corresponds to its needs. It then fills in the fields and builds his offer, then sends it to the provider agent via the "makeProposal(smartContract)" method. Upon receipt of this proposal, the provider agent checks the values entered and their compliance with the predefined constraints of contracts creation. In case of invalidity of the proposal made by the consumer agent, the provider agent is responsible for adapting the non-conforming values and making a counter-offer by calling the "makeCounterProposal(smartContract)" method. The negotiation process can see several back and forth between the two agents until a common agreement is reached or the failure following the consumption of all the allowed iterations. The common agreement is reflected in the creation of the smart contract signed by the provider agent and sends it to the consumer agent who in turn signs it and communicates his proof of approval. Then, the provider agent updates the security policy according to the access rules negotiated in the contract. It also provides the consumer agent with the possibility of monitoring his contract.

3) *WS-Agreement Implementation:* Many projects implement the WS-Agreement Specification and Protocol to set up access agreements. Some frameworks, such as WSAG4J [29] and SLA-Framework [30], make it possible to simplify this implementation. WSAG4J is a generic implementation of the WS-Agreement Protocol. It supports common functionality to create and monitor agreements generically and allows users to quickly create and deploy services based on WS-Agreement Specification. WSAG4J follows a declarative approach to support and manage the entire lifecycle of an agreement, from the definition of an agreement template, through the deployment of models in *Factories*, to the management of the agreement. SLA-Framework, as for it, is another implementation of the WS-Agreement Specification. It is an open source project that helps manage the lifecycle of access agreements (negotiation, creation and monitoring agreements). Currently in version V1.1, this framework only supports one-shot negotiation. The core SLA-Framework provides a language and protocol to define and advertise the capabilities of service providers in SLA templates, create agreements based on the templates, and track compliance with agreements at runtime.

C. Case Study

France's national health security organization manages the medical records of all adherents of this state public service. All information relating to an adherent is listed in the French Health Management System (FHMS). This is personal information that the owner has designated as private, for example the Person of Confidence to Contact (PCC) or medical order that describes everything related to the Patient's Care History (PCH). All access to this information is controlled by the system which grants or denies it according to the predefined access control policy. A national or foreign medical entity can therefore issue a request for access to the system to consult all or part of a medical file. These requests are normally issued by entities already recognized by the system, but it is quite possible for a new organization (new private

sector clinic, foreign hospital, new research laboratory, etc.) to request access. In this case, a collaborative access contract is required. Smart Agents, representing consumers and data provider, handle the negotiation and commissioning of the smart contract using the web services responsible for accessing the data. These data concern the personal information of a patient (surname, first name, age, address, telephone number, etc.) as well as all medical information relating to his/her state of health and his/her care path (results of analysis, treatments, chronic diseases, allergies, blood pressure and heart rate measurements, etc.), which are collected by the various portable and implantable sensors of the WBAN (Wireless Body Area Networks) or entered directly by doctors and subsequently stored in the Cloud (Fig. 4) in an encrypted format. The encryption is performed by a local server based on the access control rules specific to each type of data. Fig. 3 shows an excerpt from the access control policy applied to encrypt the PCC and PCH of Mr. Jean Dupont.

Mr. Jean Dupont, a French traveling in Germany, is transported to the emergency room of the K Hospital in the city of Berlin following a serious accident. Dr. Karl Schmidt, who

takes care of him in the emergency department, needs to consult the information on the patient's care path as well as his contact details. Dr. Schmidt connects to the German Health Management System (GHMS) and asks through it to establish a link with its French equivalent (FHMS) in order to retrieve information from Jean Dupont. The GHMS thus sends a request to establish access to the FHMS which in return requests a certain amount of information relating to the requester and the access purpose. Following the sending of this information, the FHMS and GHMS negotiate and enter into a contract which grants the right of access to users attached to the GHMS governed by the access control policy in force in the FHMS. A token is therefore provided by the GHMS to Dr. Schmidt over a fixed period as well as the identifier of the established access contract. The FHMS, for its part, updates its security policy with the attributes relating to the new contract. Dr. Schmidt connects via the token to the PEP of FHMS, which is the user entry point to the system. The PEP transmits the access request to the PDP which studies it based on the input data and access control policies stored in the PAP database and grants Dr. Schmidt access to Jean Dupont's PCH and PCC.

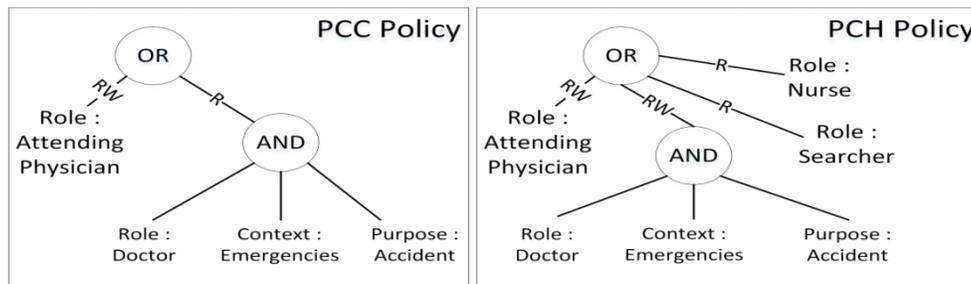


Fig. 3. PCC and PCH Access Control Policy of Mr. Jean Dupont.

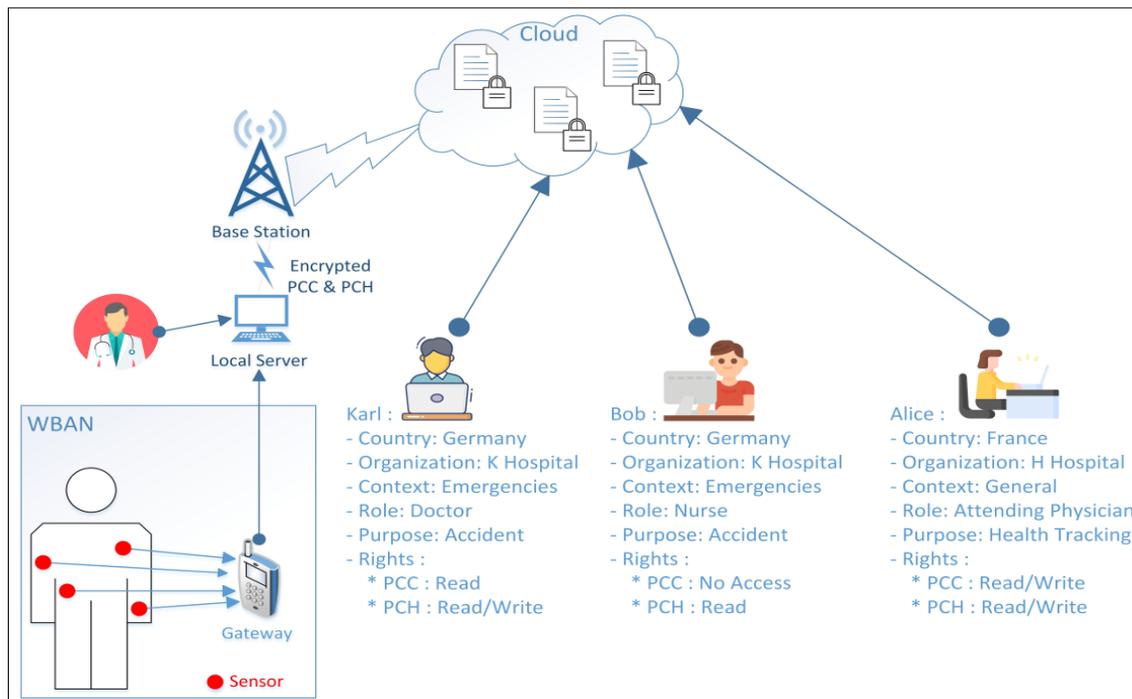


Fig. 4. Case Study Illustration.

Bob, a nurse in the emergency department at K Hospital in Germany, can also connect to the PEP thanks to the signed contract covering all staff attached to the GHMS. He made the same request as Dr. Karl Schmidt but did not obtain the same authorizations. According to the access policy in Fig. 3, Bob cannot access the PCC. Although the context (Emergency) and the purpose (Accident) for the access request matches PCC policy, Bob does not have the "Doctor" role and therefore does not meet all the criteria for access. However, he can have read-only access to PCH data in accordance with its policy. Finally, Alice, Jean Dupont's attending physician attached to the FHMS, benefits from permissions to access, read and write to all of his file (PCC and PCH), granted on the basis of her role (attending physician) in organization (H Hospital).

D. Solution Implementation

Fig. 5 shows the sequence diagram of the implementation of our case study.

This sequence diagram summarizes the technical steps taken so that Dr. Karl Schmidt can access Jean Dupont's

medical (PCH) and private (PCC) data. Dr. Karl Schmidt logs into his session in the German Health Management System (GHMS) and formulates a need for access to the French Health Management System (FHMS). The GHMS checks whether a smart contract (SC) already exists and makes it possible to respond positively to the need of Dr. Schmidt. In the case of the absence of a SC or the non-coverage of the need by the existing SCs, the GHMS initiates the automatic process of establishing a new SC with the FHMS as already detailed in the description of Fig. 2. The existence of a SC or the establishment of a new SC triggers the generation of a token by the GHMS which communicates it in addition to the identifier of the SC to Dr. Schmidt. These elements allow the doctor to activate a collaborative session at the FHMS and to formulate an access request to its PEP (Policy Enforcement Point) in which he specifies the attributes necessary to process his request. On the basis of this information, the PEP constructs a XACML request which it transmits to the PDP (Policy Decision Point) to calculate the decision.

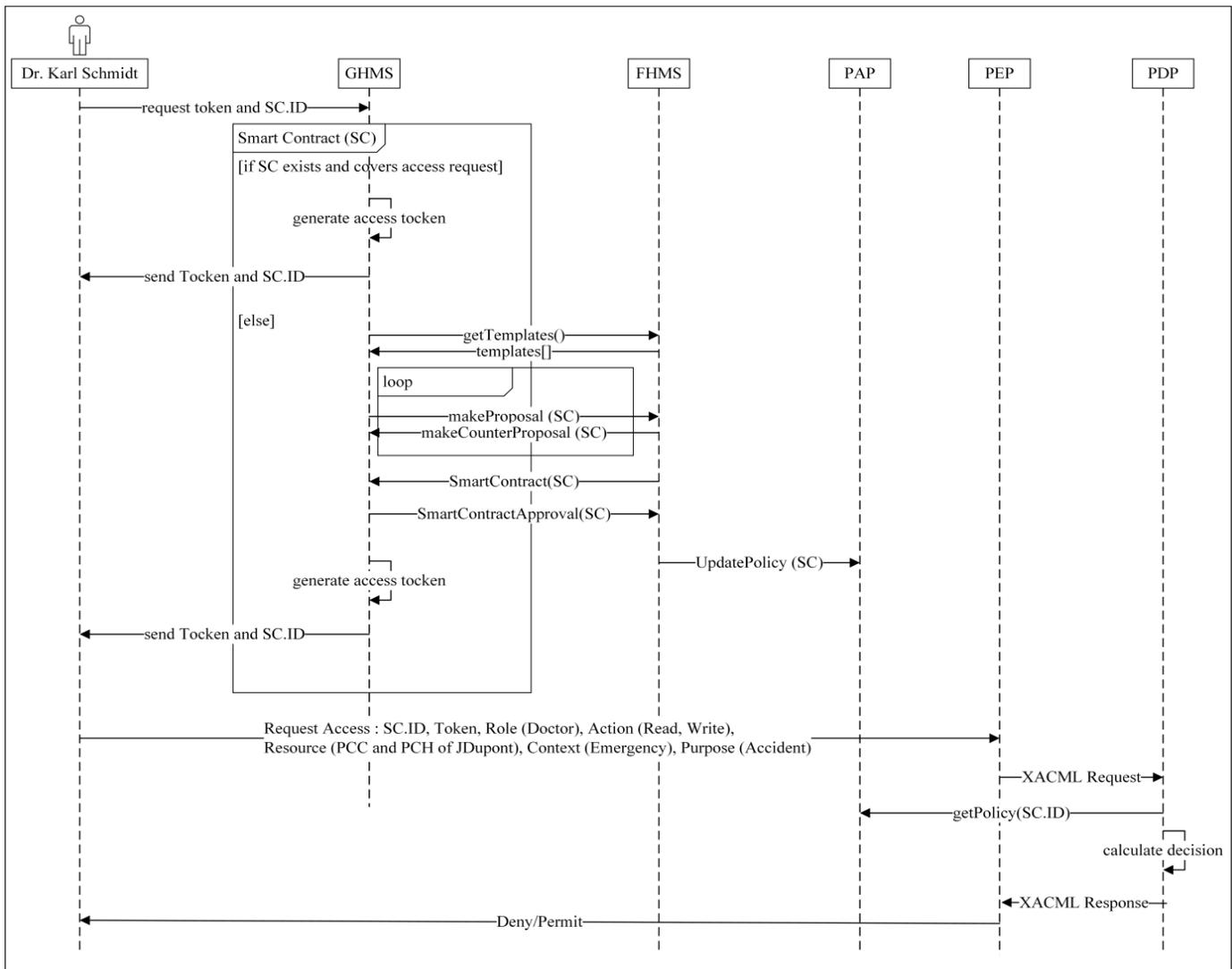


Fig. 5. Negotiation and Establishment of the Smart Contract and Collaborative Access to Data.

Here is an example of a XACML request:

```
<?xml version='1.0' encoding='UTF-8'?>
<Request>
  <Subjects>
    <Subject>
      <AttributeId>RSO.identity</AttributeId>
      <AttributeValue>Karl Schmidt</AttributeValue>
    </Subject>
    <Subject>
      <AttributeId>RSO.role</AttributeId>
      <AttributeValue>Doctor</AttributeValue>
    </Subject>
  </Subjects>
  <Resources>
    <Resource>
      <AttributeId>VOO.Owner</AttributeId>
      <AttributeValue>Jean Dupont</AttributeValue>
    </Resource>
    <Resource>
      <AttributeId>VOO.Identity</AttributeId>
      <AttributeValue>PCC</AttributeValue>
    </Resource>
  </Resources>
  <Actions>
    <Action>
      <AttributeId>AO.Type</AttributeId>
      <AttributeValue>Read</AttributeValue>
    </Action>
  </Actions>
  <Environments>
    <Environment>
      <AttributeId>Context</AttributeId>
      <AttributeValue>Emergency</AttributeValue>
    </Environment>
    <Environment>
      <AttributeId>Purpose</AttributeId>
      <AttributeValue>Accident</AttributeValue>
    </Environment>
  </Environments>
</Request>
```

In our case study, Dr. Karl Schmidt receives permission to read and write to Jean Dupont's treatment path. He has also been granted read-only access permission to his private data in accordance with the "PCC: Access_modality" policy in which the affected person gives his consent to consult this data by a doctor in the event of an accident:

```
<Policy PolicyId="PCC:Access_modality"
RuleCombiningAlgId="&rule-combine;permit-
overrides">
  <Target>
    <!-- this policy concerns the reading of PCC
patient J. Dupont who consented to doctors to
access private information for specific purposes-->
    <Resources>
      <Resource>
        <AttributeId>VOO.Identity</AttributeId>
        <AttributeValue>PCC</AttributeValue>
      </Resource>
      <Resource>
        <AttributeId>VOO.Owner</AttributeId>
        <AttributeValue>JeanDupont</AttributeValue>
      </Resource>
    </Resources>
    <Actions>
      <Action>
        <AttributeId>AO.Type</AttributeId>
        <AttributeValue>Select</AttributeValue>
      </Action>
```

```
</Actions>
</Target>
<Rule RuleId="PCC:Access_PCC" Effect="Permit">
  <Target>
    <Resources>
      <Resource>
        <AttributeValue>Consent.Y</AttributeValue>
      </Resource>
    </Resources>
    <Actions>
      <Action>
        <AttributeId>AO.Type</AttributeId>
        <AttributeValue>True</AttributeValue>
      </Action>
    </Actions>
  </Target>
  <Condition FunctionId="function:and">
    <Apply FunctionId="&function;string-equal">
      <AttributeId>RSO.role</AttributeId>
      <AttributeValue>doctor</AttributeValue>
    </Apply>
    <Apply FunctionId="&function;string-is-in">
      <AttributeId>Purpose.title</AttributeId>
      <AttributeValue>accident</AttributeValue>
    </Apply>
  </Condition>
</Rule>
</Policy>
```

We have therefore seen through this example of an implementation how a subject (Dr. Karl Schmidt) having a role in an organization (Doctor attached to the GHMS) succeeded in exercising an activity (reading and / or writing) on views of objects managed by another organization (PCC and PCH of Jean Dupont, member of the FHMS) in a specific context (Emergencies) and for a specific reason (Accident) following the consent of the data owner (Jean Dupont). This case study and this implementation therefore demonstrates how we were able to implement the process of automating the management of security and privacy policies based on PrivOrBAC using smart contracts and the WS-Agreement Specification.

IV. DISCUSSION

Complex systems allow entities to exchange information even though it is unknown to each other. This type of exchange is generally governed by access contracts negotiated in advance between the providers and consumers of the data. Today, this "static" approach is no longer appropriate. According to [31], Web Services are widely used in the automation of decision making while the use of WS-Agreement asserts itself in the automation of service level agreements (SLA). Automating access negotiation processes and integrating security rules into SLAs seems to be a good solution to make access control dynamic. Stankov et al. [32] consider that SLAs can be used as an instrument to build trust between providers and consumers of services in Cloud Computing platforms even before a relationship is established between them. The WS-Agreement Specification appears to be one of the most promising solutions to achieve this goal. Ludwig et al. [33] use WS-Agreement to negotiate SLA contracts to share resources in Grid Computing. In this same context, Smith et al. [34] also use WS-Agreement and Web Services to set up a security configuration mechanism according to the requirements provided by the user such as the configuration of resources and the quality of service, as well as

security parameters such as the level of encryption and sandboxing. In our work, we use WS-Agreement and Web Services to extend the PrivOrBAC model by incorporating access agreements that cover security rules to protect Privacy.

The WS-Agreement Specification can therefore be combined with access control models to allow data providers and consumers to specify their security rules. This is the case of Li et al. [35] which propose to integrate the attributes and their values in an SLA contract through OrBAC rules. This approach consists in that, thanks to the negotiation protocol of WS-Agreement, the providers and consumers of services exchange offers and counter-offers until an agreement is reached. However, this approach does not explicitly implement the WS-Agreement Specification to eliminate human intervention when discovering services and negotiating terms of the agreement. Our solution, on the other hand, makes it possible to respond to this problem and allows intelligent agents (or connected objects) to discover each other, discover services, negotiate service level clauses as well as clauses related to access security. However, the explicit management of trust based on transaction history and the management of inferences are, among others, two aspects that still needed to strengthen the model and avoid derivation. Indeed, the application of security rules seems to be a good solution to block unauthorized access and even more if these rules are negotiated automatically, but the dynamic reinforcement of these rules throughout the life of the information system is an essential requirement in today's systems.

V. CONCLUSION AND PERSPECTIVES

A. Conclusion

In this article, we first started by explaining our directions in terms of privacy protection. We have given a definition to this concept and we have subsequently explained the approach followed, the choices adopted and the contributions made in our previous work to achieve the integration of privacy protection from the early design phases of complex systems. This integration in its static form is not entirely satisfactory to large organizations equipped with systems with high volumes of interactivity. The most awaited by these organizations is a dynamic administration of security policies and privacy protections. It is in this context that we have proposed in this article an automatic process for managing these policies in interactive systems. Our mechanism is based on the negotiation and establishment of smart contracts through the WS-Agreement Specification. This automation allows a system to autonomously manage access requests from external organizations, negotiate access contracts in accordance with current policies with these organizations, and update these policies with attributes of new contracts. Thus, we have provided the means to convert our integration of security and privacy policies from a static to a dynamic form.

B. Perspectives

We have taken care in our work to offer the various means capable of covering all aspects of privacy protection, ranging from the definition of this concept to the automation of the process guaranteeing it. "Management of inferences" is another aspect that remains to be taken into account. It is true that

PrivOrBAC is a powerful model which allows defining and configuring access controls adapted to different situations, but the risk of inference in this model is not reduced to zero. A user with access to a set of data can always combine them to derive private data that they cannot directly consult. In our case study, suppose that Jean Dupont is HIV positive and he chose to keep this information private. Bob, the nurse at K Hospital, does not have access to Jean Dupont's private data, but can view medical data. Bob can therefore see that Jean Dupont is being treated with a protein called "interferon". This protein is used in the treatment of viral diseases (AIDS, hepatitis, papillomavirus, etc.), in oncology (sarcoma) or in preventive treatment. This information alone does not allow deducing with certainty that Jean Dupont is a carrier of HIV, but by combining it with the results of analyzes on the P24 antigen which are also part of the medical data that Bob can consult, the private data can therefore be revealed. Our perspectives therefore focus on proposing ways to combine with PrivOrBAC to guarantee privacy-aware access control without risk of inference.

REFERENCES

- [1] N. Ajam, N. Cuppens-Bouahia, and F. Cuppens, "Contextual Privacy Management in Extended Role Based Access Control Model," in Springer-Verlag Berlin Heidelberg, 2010, pp.121-135. https://doi.org/10.1007/978-3-642-11207-2_10.
- [2] A. Abou El Kalam, R. El Baida, P. Balbiani, S. Benferhat, F. Cuppens, Y. Deswarte, A. Miège, C. Saurel, and G. Trouessin, "Or-BAC : un modèle de contrôle d'accès basé sur les organisations, " in Cahiers francophones de la recherche en sécurité de l'information, 2003, pp. 30-43.
- [3] J. EL MOKHTARI, A. ABOU EL KALAM, S. BENHADDOU, and H. MEDROUMI, "PrivUML: A Privacy Metamodel," in Proceedings of the 10th International Conference on Ambient Systems, Networks and Technologies (ANT 2019), 2019, pp.53-60. <https://doi.org/10.1016/j.procs.2019.04.011>.
- [4] OASIS , "eXtensible Access Control Markup Language (XACML) Version 1.0," 2003. [www.oasis-open.org > committees > oasis-xacml-1.0.pdf](http://www.oasis-open.org/committees/oasis-xacml-1.0.pdf).
- [5] A. Andrieux, K. Czajkowski, A. Dan, K. Keahey, H. Ludwig, T. Nakata, J. Pruyne, J. Rofrano, S. Tuecke, and M. Xu, "Web services agreement specification (WS-Agreement), " in Global Grid Forum, 2004. <http://www.ogf.org/documents/GFD.192.pdf>.
- [6] J. EL MOKHTARI, A. ABOU EL KALAM, S. BENHADDOU, and J.P. LEROY, "Transformation of PrivUML into XACML Using QVT," in Proceedings of the 12th International Conference on Soft Computing and Pattern Recognition (SoCPAR), 2021, pp.1-13. https://doi.org/10.1007/978-3-030-73689-7_93.
- [7] F. Cuppens, and N. Cuppens-Bouahia, "Modeling contextual security policies," in Springer-Verlag, 2007, pp. 285–305. DOI 10.1007/s10207-007-0051-9.
- [8] T. Lodderstedt, D. Basin, and J.Doser, "SecureUML: A UML-Based Modeling Language for Model-Driven Security," In UML 2002 - Springer, Berlin, Heidelberg, 2002, pp.426-441. https://doi.org/10.1007/3-540-45800-X_33.
- [9] J. Jürjens, "UMLsec: Extending UML for Secure Systems Development," In UML 2002 - Springer, Berlin, Heidelberg, 2002, pp.412-425. https://doi.org/10.1007/3-540-45800-X_32.
- [10] P. Colombo, and E. Ferrari, "Towards a Modeling and Analysis Framework for Privacy-aware Systems," In International Conference on Privacy, Security, Risk and Trust and International Conference on Social Computing, 2012, pp.81-90. DOI: 10.1109/SocialCom-PASSAT.2012.12.
- [11] T. Basso, L. Montecchi, R. Moraes, M. Jino, and A. Bondavalli, "Towards a UML Profile for Privacy-Aware Applications," In IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic

- and Secure Computing; Pervasive Intelligence and Computing, 2015, pp.371-378. DOI: 10.1109/CIT/IUCC/DASC/PICOM.2015.53.
- [12] M. Kõlvart, M. Poola, and A. Rull, "Smart Contracts," In *The Future of Law and eTechnologies* - Springer, Cham, 2016. https://doi.org/10.1007/978-3-319-26896-5_7.
- [13] I. Sergey, A. Kumar, and A. Hobor, "Scilla: a Smart Contract Intermediate-Level Language," in *DBLP Computer Science Bibliography*, 2018. arXiv:1801.00687.
- [14] C.D. Clack, V.A. Bakshi, and L. Braine, "Smart Contract Templates: foundations, design landscape and research directions," in *Barclays Bank PLC*, 2017. <http://arxiv.org/abs/1608.00771>.
- [15] V. Scoca, R.B. Uriarte, and R. De Nicola, "Smart Contract Negotiation in Cloud Computing," in *IEEE 10th International Conference on Cloud Computing (CLOUD)*, 2018, pp.592-599. DOI: 10.1109/CLOUD.2017.81.
- [16] H. Ludwig, A. Keller, A. Dan, R. King, and R. Franck, "Web Service Level Agreement (WSLA) Language Specification," in *IBM Corporation*, 2003, pp.815-824.
- [17] V. Tomic, K. Patel, and B. Pagurek, "WSOL — Web Service Offerings Language," 2002, pp.57-67. DOI: 10.1007/3-540-36189-8_5.
- [18] O. Wäldrich, D. Battré, F. Brazier, K. Clark, M. Oey, A. Pappaspyrou, P. Wieder, and W. Ziegler, "WS-Agreement Negotiation Version 1.0," 2011. <http://www.ogf.org/documents/GFD.193.pdf>.
- [19] F. Marino, C. Moiso, and M. Petracca, "Automatic contract negotiation, service discovery and mutual authentication solutions: A survey on the enabling technologies of the forthcoming IoT ecosystems," in *Computer Networks*, 2018, pp.176-195. DOI: 10.1016/j.comnet.2018.11.011.
- [20] Y. Li, N. Cuppens-Bouahia, J.M. Crom, F. Cuppens, and V. Frey, "Reaching Agreement in Security Policy Negotiation," in *IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, 2014, pp.98-105. DOI: 10.1109/TrustCom.2014.17.
- [21] F. Li, and S. Clarke, "A Context-Based Strategy for SLA Negotiation in the IoT Environment," in *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2019. DOI: 10.1109/PERCOMW.2019.8730752.
- [22] F. Li, C. Cabrera, and S. Clarke, "A WS-Agreement Based SLA Ontology for IoT Services," in *Internet of Things – ICIOT 2019*, 2019, pp.58-72. DOI: 10.1007/978-3-030-23357-0_5.
- [23] B. Shojaiemehr, A. Rahmani, and N. Qader, "A Three-phase Process for SLA Negotiation of Composite Cloud Services," in *Computer Standards & Interfaces*, 2019. DOI: 10.1016/j.csi.2019.01.001.
- [24] T. Labidi, A. Mtibaa, W. Gaaloul, and F. Gargouri, "Ontology-Based SLA Negotiation and Re-Negotiation for Cloud Computing," in *IEEE 26th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, 2017, pp.36-41. DOI: 10.1109/WETICE.2017.24.
- [25] R. Baig, W. Khan, I. Haq, and I. Khan, "Agent-Based SLA Negotiation Protocol for Cloud Computing," in *International Conference on Cloud Computing Research and Innovation (ICCCRI)*, 2017, pp.33-37. DOI: 10.1109/ICCCRI.2017.13.
- [26] S.S. Tseng, H.C. Chen, L.L. Hu, and Y.T. Lin, "CBR-based negotiation RBAC model for enhancing ubiquitous resources management," in *International Journal of Information Management*, 2017, pp.1539-1550. DOI: 10.1016/j.ijinfomgt.2016.05.009.
- [27] I. Castro, A. Panda, B. Raghavan, S. Shenker, and S. Gorinsky, "Route Bazaar: Automatic Interdomain Contract Negotiation," in the *15th Workshop on Hot Topics in Operating Systems (HotOS)*, 2015.
- [28] C. Coutinho, A. Cretan, C.F. da Silva, P. Ghodous, and R. JardimGonçalves, "Service-based Negotiation for Advanced Collaboration in Enterprise Networks," in *Journal of Intelligent Manufacturing - Springer Verlag*, 2016, pp.201-216. DOI: 10.1007/s10845-013-0857-4ff.
- [29] WSAG4J framework V2.0.0, 2012. <http://wsag4j.sourceforge.net/site/wsag/overview.html>.
- [30] SLA Framework v1.1, 2016. <https://github.com/FIWARE/ops.Sla-framework>.
- [31] A. Lenk, L. Bonorden, A. Hellmanns, N. Roedder, and S. Jaehnichen, "Towards a Taxonomy of Standards in Smart Data," in *IEEE International Conference on Big Data*, 2015. DOI: 10.1109/BigData.2015.7363946.
- [32] I. Stankov, R. Datsenka, and K. Kurbel, "Service Level Agreement as an Instrument to Enhance Trust in Cloud Computing – An Analysis of Infrastructure-as-a-Service Providers," in *Proceedings of the Eighteenth Americas Conference on Information System (AMCIS)*, 2012. <http://aisel.aisnet.org/amcis2012/proceedings/HCIStudies/12>.
- [33] H. Ludwig, T. Nakata, O. Wäldrich, P. Wieder, and W. Ziegler, "Reliable Orchestration of Resources Using WS-Agreement," in *HPCC 2006 - Springer, Berlin, Heidelberg*, 2006. https://doi.org/10.1007/11847366_78.
- [34] M. Smith, M. Schmidt, N. Fallenbeck, C. Schridde, and B. Freisleben, "Optimising Security Configurations with Service Level Agreements," 2007. <https://www.semanticscholar.org/paper/Optimising-Security-Configurations-with-Service-Smith-Schmidt/bb18eca8f629edb494cb8dd8c7fab9027cd23dcf>.
- [35] Y. Li, N. Cuppens-Bouahia, J.M. Crom, F. Cuppens, and V. Frey, "Expression and Enforcement of Security Policy for Virtual Resource Allocation in IaaS Cloud," in *IFIP Advances in Information and Communication Technology - Springer*, vol 471, 2016. https://doi.org/10.1007/978-3-319-33630-5_8.

Hyperspectral Image Classification using Convolutional Neural Networks

Shambulinga M¹, G. Sadashivappa²

Dept. of Electronics and Telecommunication Engineering
RV College of Engineering, Bengaluru

Abstract—Hyperspectral image is well-known for the identification of the objects on the earth's surface. Most of the classifier uses the spectral features and does not consider the spatial features to perform the classification and to recognize the various objects on the image. In this paper, the hyperspectral image is classified based on spectral and spatial features using a convolutional neural network (CNN). The hyperspectral image is divided into a small number of patches. CNN constructs the high level spectral and spatial features of each patch, and the multi-layer perceptron helps in the classification of image features into different classes. Simulation results show that CNN archives the highest classification accuracy of the hyperspectral image compared with other classifiers.

Keywords—Convolutional neural network; hyperspectral image; classification

I. INTRODUCTION

Latest developments in optics and photonics have made a hyperspectral imaging sensor with better spectral and spatial resolution. The spatial and spectral information is efficiently exploited to identify the materials and objects on the earth surface. The spectral signatures are modelled in such a way that they will differentiate the various objects and materials. It is possible to see the identification of different materials, objects and surface ground cover classes based on their reflectance properties as a classification task, i.e. the classification of image pixels based on their spectral characteristics. Classification of hyperspectral imaging has been used in a broad range of applications, such as agriculture, environmental science, astronomy, surveillance, astronomy and biomedical imaging. However, the classification of the hyperspectral images has its own particular problems, in addition to (i) high dimensionality, (ii) the small number of samples which have been labelled, and (iii) significant spatial variation of spectral signatures.

Much of the recent work on hyperspectral data classification follows the traditional pattern recognition method, which comprises of two distinct steps: first, detailed handcrafted features are obtained from the original data input, and secondly, classifiers like Support Vector Machines (SVM), Neural Networks (NN) [1], maximum likelihood [2], parallelepiped classification, k nearest neighbours [3], minimum distance [4], and logistic regression [5] have been used to learn the extracted features. The “curse of dimensionality” affects the majority of the algorithms mentioned above. Any dimensionality reduction-based classification approaches [6] have been suggested to manage

the large dimension complexity and small training samples of hyperspectral data. Band selection and transformation are the other methods available to deal with dimensionality. In general, statistical learning techniques have been used to solve the large dimensionality and variability of high dimensional hyperspectral data and when few training samples are accessible.

Support vector machine, a popular classification method for hyperspectral data classification, is presented in [7]. SVM is resistant to the Hughes phenomena and has poor sensitivity to high dimensionality. In certain situations, SVM-based classifiers perform better compared to other commonly used pattern recognition strategies in terms of classification accuracy. These classifiers were cutting edge technological tools for a long time.

In recent years, spatial information has become highly relevant for hyperspectral data classification. In terms of efficiency, spatial-spectral classification methods provide substantial benefits. Many new methods aim to bring spatial information into account to deal with spectral signature spatial variability. In [8], applying SVM and a guided image filter, a technique for classifying hyperspectral images has been presented. The guided image filter is used to incorporate the spatial features into the SVM classifier. In [9], the edge-preserving filters such as bilateral filter and guided image filter are included to incorporate the spatial features to the SVM classifier.

However, it is essential to know which features are most relevant for the classification algorithms due to the high diversity of represented materials. The various deep learning models [10,11] have been developed for classification purposes. These models are trained from various levels of features. The high levels features required for training the model are obtained from low-level features. These models automate the extraction of features for any problem compared with any convolutional pattern recognition technique. In addition, deep learning systems tend to match and resolve the classification issue more efficiently with wider datasets and large images with high spatial and spectral resolution. Deep learning approaches have already shown promising results for detecting real artefacts, such as man-made objects or vehicles, as well as for classifying hyperspectral data.

More precisely, a deep learning system for the classification of hyperspectral data with accurate results was used in [12].

In precise, the auto encoder's helps in the design of the deep architecture, which hierarchically extracts the high-level spectral features required for the classification of each pixel in the image. Spectral features were coupled with spatial-dominated information in a separate stage and fed to a logistic regression classifier as input.

In the same way, for the classification of hyperspectral data into various classes, we suggest a deep learning system. Our method, however, is based on a coherent structure that incorporates spectral and spatial information in a single stage, creating high-level spectral-spatial characteristics at the same time.

In specifically, we suggest the use of a Modified Convolutional Neural Network (CNN) that performs the operation of constructing large-level features and a Multi-Layer Perceptron (MLP) which is used for the classification of the image. The evolved framework builds spectral-spatial characteristics at once under this kind of design and, at a similar time, performs real-time predictions of the various classes in the image because of the existence of feed forward network in CNNs and MLPs.

This paper is organized into five sections: Section II gives a brief introduction and background of the convolution neural network. Section III presents the implementation methodology of the convolution neural network for hyperspectral image classification. Sections IV and V discuss the results and discussion of the image classification and conclusion.

II. CONVOLUTIONAL NEURAL NETWORK (CNN): BACKGROUND

Convolution neural network (CNN) is the most standard deep learning algorithm for image classification and image recognition problems. Along with these applications, CNN is widely used to recognize human faces and classify objects.

CNN algorithm takes the input image and extract the features in the learning phase and classify it to various class or

categories. The algorithm sees the input image as an array of a matrix ($h \times w \times d$), and it will depend on the image resolution. CNN consider the ($h \times w \times 3$) matrix for the RGB image and ($h \times w \times 1$) for the grayscale image. In principle, create the CNN model bypassing the input image series through various convolution layers with filters or kernels to extract features, pooling, fully connected layers, and finally applying the softmax function to detect the objects probabilistic value between 0 and 1. Fig. 1 shows the entire CNN follow diagram process.

A. Convolution Layer

The features existing in the input images are extracted through a convolution layer. The relationship that exists between the image's pixels is preserved by taking a small square of input data at the learning phase. The convolution is the mathematical operation given by the multiplication of an image matrix and a filter or kernel. The convolution operation of the image is shown in Fig. 2.

Consider an example of an image pixel whose matrix is 5×5 with values 0 and 1 and the filter matrix or kernel 3×3 . The convolution layer multiplies the image matrix with a filter matrix to provides the convolved feature or feature map. The same convolved operations are shown in Fig. 3 and Fig. 4.

Use the various types of filters or kernel listed in the Table I for convolution operation. These filters can also perform different mathematical calculations like sharpening of the image, blur the image and to detect the edges in the image.

B. Strides

Strides are used in convolution operation, and it determines the shift in the filter kernel matrix by certain number of pixels on the input image. For example, if stride value is 1 in convolution operation, then the filter kernel in convolution is shifted by one pixel at a time on the image matrix. If stride value is 2, then a filter kernel is shifted by two pixels at a time on the image matrix. Fig. 5 shows the operation strides by 2 pixels.

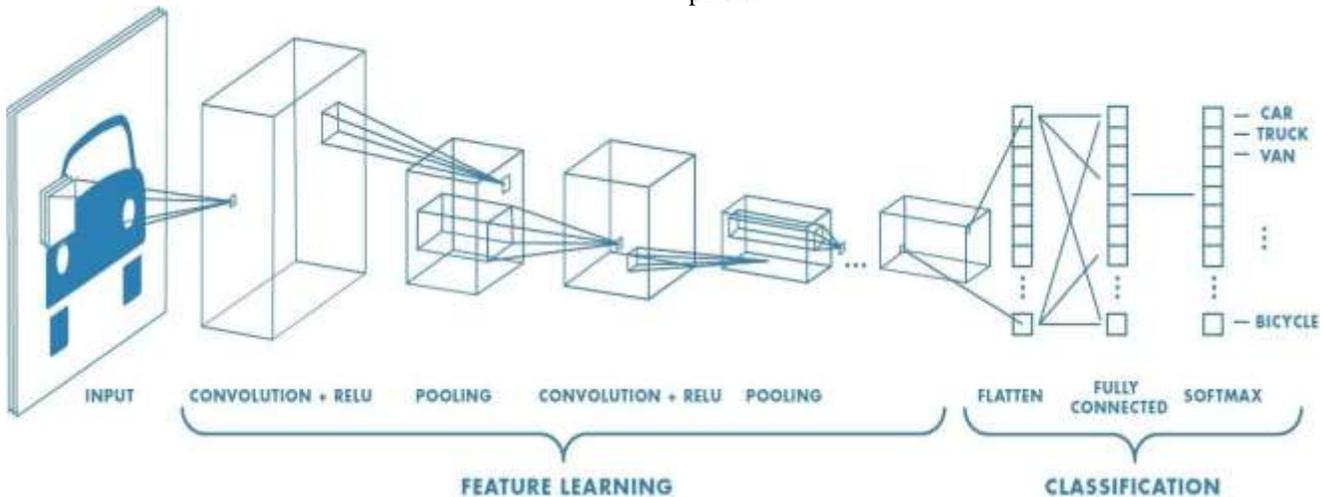


Fig. 1. CNN Flow Diagram.

- An image matrix (volume) of dimension $(h \times w \times d)$
- A filter $(f_h \times f_w \times d)$
- Outputs a volume dimension $(h - f_h + 1) \times (w - f_w + 1) \times 1$

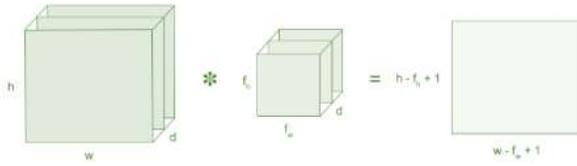


Fig. 2. Convolution Operation of the Image using Filter Kernel.

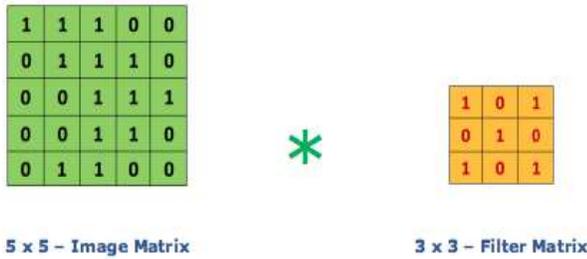


Fig. 3. Image Matrix Multiplies Kernel or Filter Matrix.

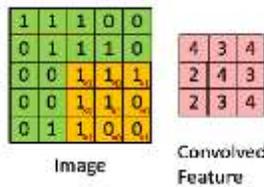


Fig. 4. 3 x 3 Output Matrix.

TABLE I. FILTERS AND ITS KERNEL FOR CONVOLUTION OPERATIONS

Operation	Filter
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Edge Detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$
Box blur	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$
Gaussian blur	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

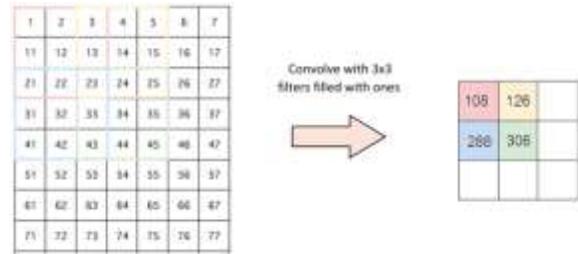


Fig. 5. Stride by 2 Pixels in Convolution Operations.

C. Padding

In most of the cases, the chosen filter does not fit to the input image. In those cases, the following options are used to fit the input images.

- Zero paddings: input image is padded with zeros to fit the input images.
- Drop a few pixels in the image to fit the input images.

D. Non Linearity (ReLU)

Rectified Linear Unit (ReLU) performs the non-linear computation in convolution network. The output of ReLU is $f(x) = \text{Max}(0, x)$.

The goal of introducing the non-linearity in ReLU is to learn the non-negative linear values in the convolution network. Fig. 6 shows the ReLU operation in CNN.

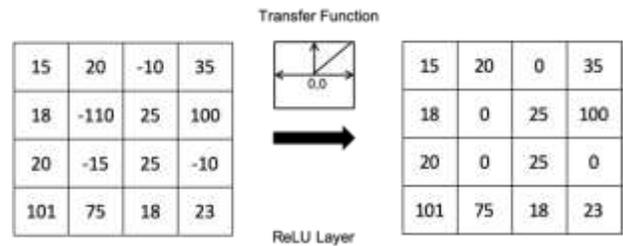


Fig. 6. ReLU Operation in CNN.

Tanh and sigmoid are the some of the non-linear functions which can be employed instead of ReLU. But the performance of ReLU is better compared to the other two non-linear operations

E. Pooling Layer

The pooling layer decreases the output pixels of the convolution layer and reduces the complexity of large images. The spatial pooling is also called subsampling or downsampling of pixels, and it decreases the spatial dimension of the image by retaining most of the information in the image. Following are three different categories of spatial pooling.

- Max Pooling
- Average Pooling
- Sum Pooling

Max pooling selects the greatest value in the convolved feature value, as shown in Fig. 7. The average pooling selects the average pixel value of all the convolved feature value. The sum pooling selects the addition of all the components in convolved feature value.

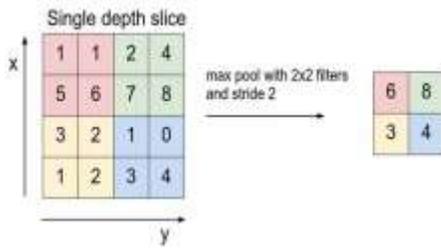


Fig. 7. Max Pooling Operation.

F. Fully Connected Layer

The two-dimensional matrix of the pooling layer is converted into a single dimension matrix. This single layer vector is feed into a fully connected layer similar to a neural network. Fig. 8 shows the operation of a fully connected layer. The output feature matrix is converted to vector as x_1, x_2, x_3 . These features are merged to construct a model using a fully connected layer and finally use softmax or sigmoid function to classify the outputs.

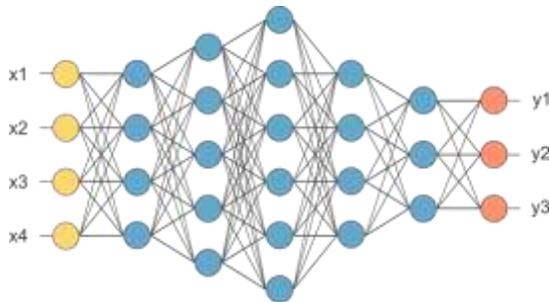


Fig. 8. Fully Connected Layer in CNN.

III. METHODOLOGY

The implementation methodology of hyperspectral image classification using the CNN approach is shown in Fig. 9. The hyperspectral image is defined as a 3-D matrix of size $h \times w \times c$, where h and w represent the image's height and width and c represents the spectral channel in the image. The hundreds of channels in the hyperspectral image enhance the calculating time and memory resources of the training and prediction process. On the other hand, using statistical analysis, it is observed that the spectral response variance is minimal for the pixel that belongs to each class.

It means for every channel, pixels with the same class labels have related spectral values, and pixels with different class labels have different spectral values. A dimensionality reduction procedure can be used, based on these properties, to decrease the dimensionality of the input data, to optimize the training and classification processes. PCA is the dimensionality reduction algorithm that reduces the hyperspectral image's

spectral dimensions without losing any image information. After dimensionality reduction, split the hyperspectral image into small patches to compatible with CNN's basic nature. Each created patch contains the spectral and spatial features of a single pixel.

More precisely, square patch of dimension $s \times s$ centred at a pixel $P_{x,y}$ is used to classify the pixel at location (x, y) on the image plane and combine spectral and spatial feature data. Let us denote the class label of the pixel at the given location as $l_{x,y}$ and patch centred at pixel $p_{x,y}$ as $w_{x,y}$. then create a dataset $D = \{w_{x,y}, l_{x,y}\}$ where $x = 1, 2, \dots, w$ and $y = 1, 2, \dots, h$. The patch $w_{x,y}$ is the 3D matrix with size $s \times s \times c$, which has spatial and spectral data for each pixel. Furthermore, this matrix $w_{x,y}$ is split into c matrixes of size $s \times s$, which are given as an input to the CNN algorithm. This CNN develops the large features that encapsulate the pixel's spectral and spatial information. The CNN architecture contains many layers, as shown in Fig. 10. The first layer in CNN structure is the convolution layer with $C_1 = 3 \times c_r$ a trainable filter of dimensions 3×3 . This layer gives C_1 matrixes of dimensions 3×3 . The second layer is also a convolution layer with $C_2 = 3 \times C_1$ trainable filter. The output of the convolution layer is given to flatten layer and fully connected layer. The flow chart for the implementation of hyperspectral image classification using a convolutional neural network is shown in Fig. 11.

IV. RESULTS AND DISCUSSIONS

The Hyperspectral image is used to simulate the proposed algorithm. Indian pines dataset is the hyperspectral image captured at north western Indiana test site using AVIRIS sensors, and its specification is shown in Table II. The overall accuracy, test loss, precision, recall and f1-score are parameters used to analyze the performance of algorithm.

In the simulation, PCA reduces the spectral bands, and 30 principal components are chosen for the classification. After dimensionality reduction, each patch contains $s \times s \times c_r$ dimensions. During the simulation, the patch size s is defined as 5 to consider the nearest 24 neighbours of each pixel. Each patch is given to the CNN architecture for classification.

The classification accuracy obtained for the 5×5 patch is 84.12%, with a simulation time of 1142.26 seconds. The screenshot of the results and the classified image is shown in Fig. 12. The test loss for the same is 47.33%. The precision, recall and f1-score of the individual class is shown in Fig. 12. The patch size s is increased to 7, 9, 11 and 13 to increase the number of neighbours for each pixel.

TABLE II. SPECIFICATION OF INDIAN PINES DATASET

Image size	145x145 pixels
Spectral bands	224
Spectral range	(0.4-2.5) μ m
No. of Class	16

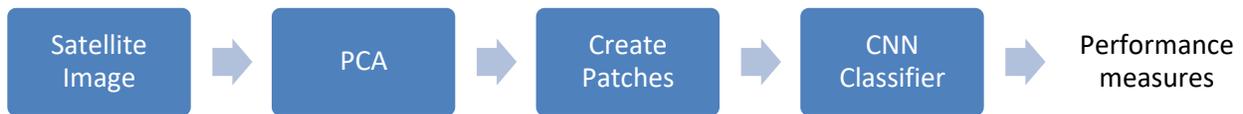


Fig. 9. Block Diagram of the Implementation of Image Classification using Convolutional Neural Networks.

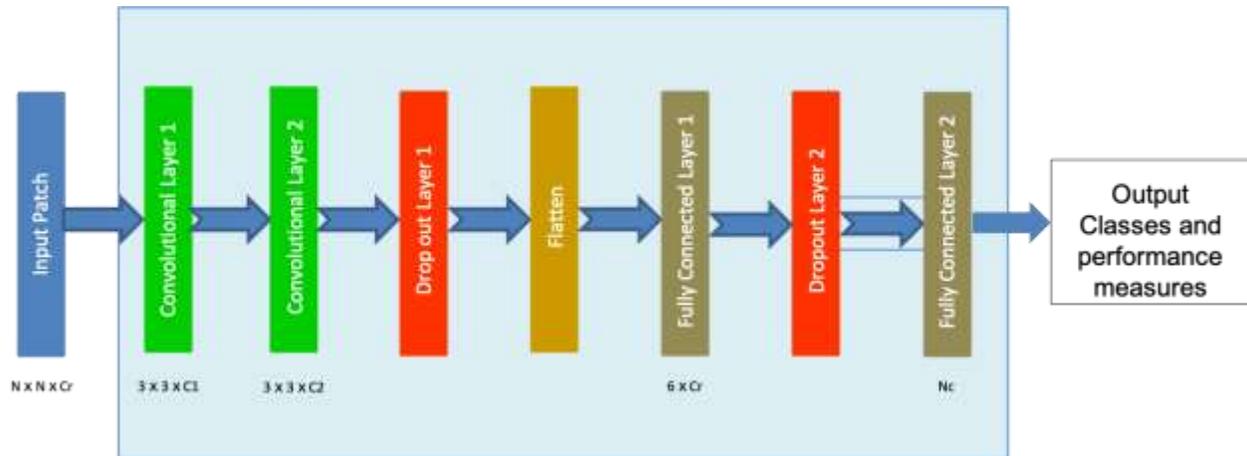


Fig. 10. CNN Architecture.

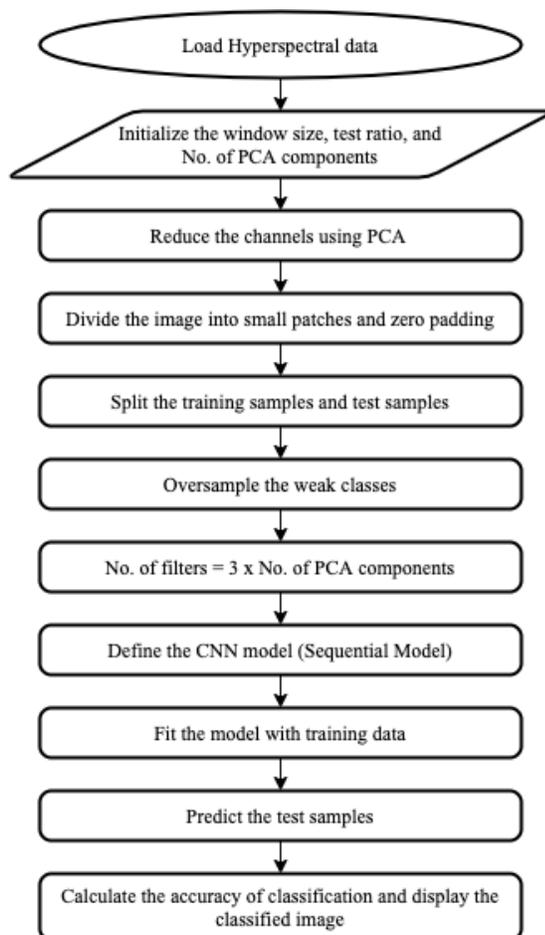


Fig. 11. Flow Chart of an Implementation of Hyperspectral Image Classification.

47.336000204086304 Test loss (%)
84.12017226219177 Test accuracy (%)

	precision	recall	f1-score	support
Alfalfa	1.00	1.00	1.00	11
Corn-notill	0.73	0.77	0.75	357
Corn-mintill	0.74	0.88	0.80	208
Corn	0.77	0.93	0.85	59
Grass-pasture	0.96	0.95	0.95	121
Grass-trees	0.98	0.97	0.98	183
Grass-pasture-mowed	1.00	1.00	1.00	7
Hay-windrowed	1.00	1.00	1.00	120
Oats	0.83	1.00	0.91	5
Soybean-notill	0.68	0.84	0.75	243
Soybean-mintill	0.88	0.66	0.75	614
Soybean-clean	0.76	0.84	0.80	148
Wheat	1.00	1.00	1.00	51
Woods	0.99	0.97	0.98	316
Buildings-Grass-Trees-Drives	0.84	0.94	0.89	97
Stone-Steel-Towers	0.96	1.00	0.98	23
accuracy			0.84	2563
macro avg	0.88	0.92	0.90	2563
weighted avg	0.85	0.84	0.84	2563

Fig. 12. Screenshot of CNN Classification with 5 x 5 Patch Size.

Increasing the patch size enhance the classification accuracy and increases the simulation time. The results of CNN classification with different patch size is shown in Fig. 13. The accuracy and computation time of CNN classification is given in Table III. From the table, it can be seen that CNN achieves the highest classification accuracy and computation time for the patch size 13 x 13. No further improvements are observed in classification accuracy for the patch size of more than 13, and in fact, the accuracy of the classifier deteriorates and increases the computational resources. The proposed CNN algorithm for the classification of the hyperspectral image is compared with the Support vector machine classifier as shown in Table IV. Support vector machine uses the spectral features for classification and spatial features are add to SVM output by using guided image filter and bilateral filter.

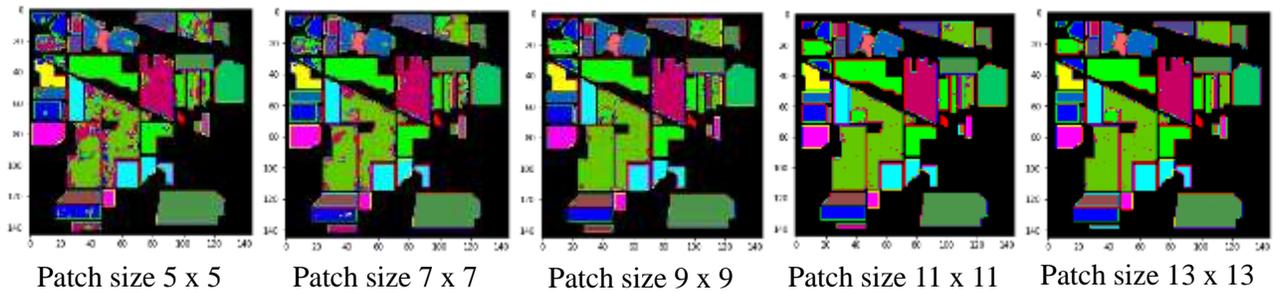


Fig. 13. Results of CNN Classification with different Patch Size.

TABLE III. ACCURACY AND COMPUTATION TIME OF CNN CLASSIFICATION WITH DIFFERENT PATCH SIZE

Patch Size	5 x 5	7 x 7	9 x 9	11 x 11	13 x 13
Overall Accuracy(%)	84.12	91.10	95.00	97.42	98.28
Computation time (sec)	1142.26	2111.512	3047.524	4263.925	5002.114

TABLE IV. COMPARISON OF SVM AND CNN CLASSIFICATION METHODS

Methods	SVM [8]	SVM with PCA [8]	SVM with LDA [9]	SVM with GIF [8]	SVM with PCA & GIF [8]	SVM with LDA and GIF [9]	SVM with BF [9]	SVM with PCA & BF [9]	SVM with LDA and BF [9]	CNN
Overall Accuracy (%)	91.28	88.1	86.53	99.04	98.99	96.37	98.89	98.93	96.53	98.28
Computation Time (sec)	4117.7	568.11	423.58	4118.15	568.25	424.1	4118.03	568.32	424.08	5002.114

Along with SVM, the dimensions of the spectral features are reduced using PCA and LDA. The SVM method along with guided image filter and bilateral filter increases the classification accuracy and SVM method along with PCA and LDA decreases the computational time. CNN algorithm combines both spectral and spatial features at the same time without using guided image filter and bilateral filter. CNN algorithm archives the highest classification accuracy compared to the SVM classification accuracy with increasing computation time. SVM algorithm consumes less computational time compared with CNN algorithm.

V. CONCLUSION

The hyperspectral image is classified using a convolutional neural network. Both spectral and spatial features are considered to classify the image. Spatial features are included in the classification to improve the classification. CNN archives the highest classification accuracy of 98.28% compared to support vector machine classifier and other methods. CNN accuracy depends on the patch size considered for the classification. Patch size indicates the number of spatial features considered for classification. It is observed that the patch size of 13 x 13 is enough to achieve the highest accuracy. CNN consumes more computation time for training and testing compared with other classifiers. The proposed method avoids the usage of an edge-preserving filter to incorporate spatial features into the classification.

REFERENCES

- [1] S.Li, W.Song, L.Fang, Y.Chen, P.Ghamisi and J.A.Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57(9), pp. 6690-6709, 2019.
- [2] D. Saqui, J. H. Saito, L. A. D. C. Jorge, E. J. Ferreira, D. C. Lima and J. P. Herrera, "Methodology for Band Selection of Hyperspectral Images Using Genetic Algorithms and Gaussian Maximum Likelihood Classifier," *International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, 2016, pp. 733-738.
- [3] B. Tu, J. Wang, X. Kang, G. Zhang, X. Ou and L. Guo, "KNN-Based Representation of Superpixels for Hyperspectral Image Classification," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 4032-4047, Nov. 2018.
- [4] W. Li, S. Prasad, E. W. Tramel, J. E. Fowler and Q. Du, "Decision fusion for hyperspectral image classification based on minimum-distance classifiers in the wavelet domain," *IEEE China Summit & International Conference on Signal and Information Processing*, Xi'an, 2014, pp. 162-165.
- [5] S. Bajpai, H. V. Singh and N. R. Kidwai, "Feature extraction & classification of hyperspectral images using singular spectrum analysis & multinomial logistic regression classifiers," *2017 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, Aligarh, 2017, pp. 97-100.
- [6] Q. Sun, X. Liu and M. Fu, "Classification of hyperspectral image based on principal component analysis and deep learning," *International Conference on Electronics Information and Emergency Communication*, 2017, pp. 356-359.
- [7] S. Zhong, C. Chang and Y. Zhang, "Iterative Support Vector Machine for Hyperspectral Image Classification," *International Conference on Image Processing*, 2018, pp. 3309-3312.

- [8] Shambulinga M and Sadashivappa G., "Hyperspectral Image Classification using Support Vector Machine with Guided Image Filter" International Journal of Advanced Computer Science and Applications, vol. 10(10), 2019.
- [9] Shambulinga M and G. Sadashivappa, "Supervised Hyperspectral Image Classification using SVM and Linear Discriminant Analysis" International Journal of Advanced Computer Science and Applications, Vol. 11(10), 2020.
- [10] M.S.Aydemir and G.Bilgin, "Semisupervised Hyperspectral Image Classification Using Deep Features," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12(9), pp. 3615-3622, 2019.
- [11] I.Bidari, S.Chickerur, H.Ranmale, S.Talawar, H.Ramadurg and R.Talikoti, "Hyperspectral Imagery Classification Using Deep Learning," *Conference on Smart Trends in Systems, Security and Sustainability*, 2020, pp. 672-676.
- [12] S.Liu, Q.Shi and L.Zhang, "Few-Shot Hyperspectral Image Classification With Unknown Classes Using Multitask Deep Learning," in *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

Skeletonization of the Straight Leg Raise Movement using the Kinect SDK

Hustinawaty¹, T. Rumambi², M. Hermita³

Department of Informatics^{1,2}

Department of Psychology³

Gunadarma University, Indonesia

Abstract—Biomechanics is widely used as a basis for research in studying disorders of the movement system in the human body. The development of biomechanics has an effect on the field of physiotherapy. Various physiotherapy tests or studies have been developed by researchers to identify and analyze the causes of movement disorders in the human body. Physiotherapy tests performed to detect movement disorders in the human body include: Hawkin Test, Standing Hip Flexion, Standing Trunk Sidebend Test and Straight Leg Raise Test. Straight Leg Raise test is a test to lift one leg straight up in a lying down position. In the medical field, Straight Leg Raise (SLR) movement is very specific for lower lumbar disc herniation and through the use of a device called Kinect which is a color based camera (RGB) and depth camera that can detect movement and track the human skeleton. To get the human body skeleton is through the stages from RGB image acquisition, pose calibration, depth image, skeletonization to feature extraction of the two legs that can form ROM. The result of this study, Kinect is able to track the user's foot frame in SLR pose. Kinect can be relied upon to be used as a tool for detecting and tracking skeletons, and can then be used to measure SLR ROM in real time and analyze low back pain problems and measure effectiveness in physiotherapy.

Keywords—Straight leg raise; range of motion; skeleton; kinect

I. INTRODUCTION

Biomechanics is the study of the use of mechanical forces and includes the interaction effects of forces or systems of movement, nerves, muscles, and forces arising from within and outside the body. Biomechanics is widely used as a basis for research in studying disorders of the movement system in the human body. The development of biomechanics has an effect on the field of physiotherapy. Various physiotherapy tests or studies have been developed by researchers to identify and analyze the causes of movement disorders in the human body in order to increase locomotion in the human body. Physiotherapy tests carried out to detect movement disorders in the human body include: Hawkin Test, Standing Hip Flexion, Standing Trunk Sidebend Test and Straight Leg Raise Test [1].

The straight leg raise test also called the Lasegue test, is a fundamental neurological maneuver during physical examination of the patient with lower back pain aimed to assess the sciatic compromise due to lumbosacral nerve root irritation [2]. The Straight Leg Raise test is used to establish tension in the sciatic nerve to assist in the diagnosis of nerve root compression from the lower lumbar nerve root. Patients can perform Straight Leg Raise (SLR) movements so that the

hip flexion Range of Motion (ROM) is obtained. The test must be done passively and the patient's knee is maintained in full extension and the hip is in neutral rotation. For each test, the doctor records every symptom produced during the test and also the degree of hip flexion. Positive test results require reproduction of foot symptoms that the patient has known between 30° and 70° from hip flexion. Positive contralateral straight leg test results are very specific for lower lumbar herniation. [3][4]. This study describes the clinical markers for Low Back Pain, in which doctors examine and observe the patient's posture and motion and assess range of motion including flexion, extension and rotation. SLR tests should be performed on patients with evidence of sciatica or radicular pain. The SLR test is specifically aimed at detecting lumbar nerve root irritation and positive identification when sciatica is produced between 30 and 60 degrees of foot elevation. The SLR Test is widely used as one of the primary physical diagnostic tests, patients who have low back, lower back and leg pain [5]. But in line with current technological developments, doctors can perform SLR tests with the help of a technology that is using motion capture technology. Motion capture is a terminology used to describe the process of recording motion and interpreting the movement into a digital model using the Kinect sensor that is controller-free gaming and video game platforms by Microsoft. Kinect has an RGB camera, depth sensor and microphone that run on special software devices, which provide the ability to capture motion in 3D, recognize faces and recognize sounds. The depth sensor records 3D video data in any lighting conditions. and can access Kinect depth information, process raw depth images to detect players (users) and find joint positions in 3D and track body skeleton [6]. Kinect technology and tracking markers of human gestures can be developed by several virtual reality applications using Kinect sensors [7]. Research [8] uses Kinect in physical rehabilitation tests to measure a person's ability to walk within a range of 10 meters and to measure the range of motion of the neck, shoulder, elbow, upper thigh, knee joints. Then research using depth images obtained by the Kinect sensor can bring innovative ways to build intelligent fall detection systems that can be used to observe parents at home, and notify guards by raising alarms in the event of a fall event, a fall detection system in real-time scenarios based on the extraction of skeletons from human figures with the help of a Kinect depth sensor[9]. The Microsoft Kinect sensor has the potential to be a low-cost solution for clinical and home-based assessment of movement symptoms in people with Parkinson's disease. The aim of this study is to establish

Kinect accuracy in measuring clinically relevant movements in people with PD, namely a series of movements including standing, reaching in all directions and walking and walking in place, holding hands, tapping fingers, feet, leg dexterity, riding a chair and hand pronation. [10]. In this research combine several image processing techniques with the depth images captured by a Kinect sensor to successfully recognize the five distinct human postures of sitting, standing, stooping, kneeling, and lying.[11].

II. RELATED WORK

A. Kinect Component

Microsoft Kinect is an input device for motion detection and Kinect is a RGB-Depth sensor from Microsoft that uses Light Coding technology PrimeSense, an Apple Inc. company. Light Coding is a technology that can reconstruct a 3-dimensional depth map of the state in realtime and detail. Kinect 640 x 480 pixel depth resolution. Kinect sensors include the following main components, shown in Fig. 1 [12]:

- Camera RGB (color).
- Infrared (IR) emitter and an IR sensor depth.
- Motor Tilt.
- Array Microphone.
- Light Emitting Diode.

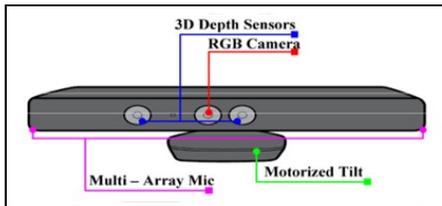


Fig. 1. Kinect Component.

B. Kinect for Windows SDK

Kinect for Windows SDK is a programming toolkit for application developers. This allows academics and the community to access the capabilities offered by Microsoft Kinect devices connected to computers with the Windows 7 OS. This SDK is equipped with drivers, APIs for raw sensor streams, skeletal tracking, installation documentation and other resources. This SDK also provides Kinect capabilities for developers who will create applications with C ++, C # and Visual Basic using Microsoft Visual Studio [13].

C. Skeleton

The development of RGB-Depth camera technology, opens new research opportunities in computer science, such as computer vision, games, motion-based control and virtual reality. Research [14] introduces a method for predicting the 3D position of the joints of the human body by extracting image depth information. Then calculate the estimated 3D position of the joints of the body using a search approach based on an average shift with a Gaussian kernel weight. By using a very large training set and various classifications of body parts, body shape, clothing and so on, classification can

be done appropriately. PrimeSense extracts features from joint data from the RGB-D Kinect camera and uses a learning method approach to infer activities that are being carried out by humans.

D. Microsoft API Skeleton

The Natural User Interface (NUI) Skeleton Application Programming Interface (API) provides information about the location of up to two players standing in front of the Kinect Sensor, with detailed information on position and orientation. The data is given to the application code as a set of points called the position of the framework, which forms the skeleton, as shown in Fig. 2. This template is the position when the user is posing. To use skeleton data, an application will display the data when initializing an NUI, then trace the skeleton. The skeleton dataset used in this study is the result of the algorithm which has been integrated by Microsoft into the Windows API operating system. Skeleton data generated by the Windows API divides the skeleton structure into 20 sections as shown in Fig. 2.

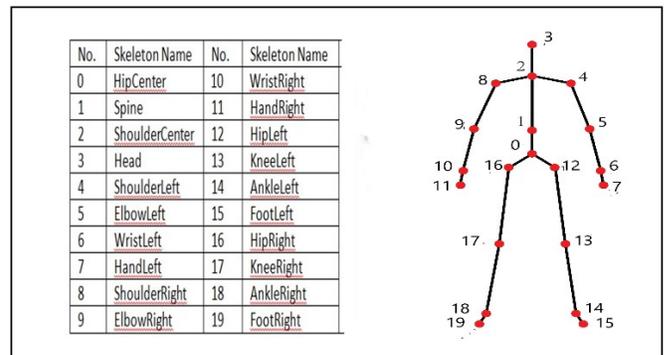


Fig. 2. Skeleton Dataset.

Each skeleton control point is defined by the X, Y, Z coordinate positions expressed in the skeleton space. Skeleton space is defined around the sensor located at the point (0, 0, 0) where the X, Y, and Z axes meet. In units of meters with the midpoint of the coordinates located on the HipCenter skeleton point and depth data (depth) which is the Z axis whose value is the distance between the Kinect device and the object [12].

E. Straight Leg Raise (SLR)

Straight Leg Raise Test is the movement of lifting one leg up straight in the lying position. The movement of lifting the legs straight up slowly will form a hip angle between the legs as shown in Fig. 3. The SLR movement is limited by pain, this test is positive and shows a sciatica of the lumbosacral (spine) [5] [15].

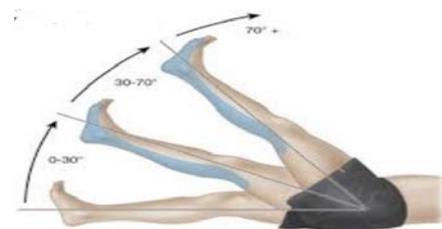


Fig. 3. Straight Leg Raise [1].

III. METHODOLOGY

A. Diagram of the Skeletonization Process Steps

Stages of forming skeleton at the time of the SLR can be described in the flowchart diagram in Fig. 4.

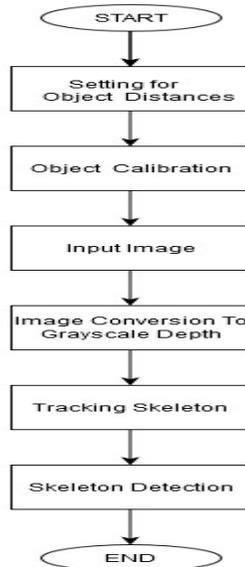


Fig. 4. Diagram of Methodology.

The initial stage is the acquisition and calibration stage. Kinect as an image acquisition device is placed at a height of approximately 3 meters from the floor. The patient in the supine position is the object of image acquisition by a Kinect device that has been equipped with a camera. The image format recorded by the camera is in RGB format. The calibration stage is the alignment stage between the camera position and the patient's body position in a supine pose so that it can be detected according to the expected position standard. The results of calibrated image acquisition are converted into grayscale depth. The second step is tracking the skeleton. Based on the depth data obtained and data on Kinect as machine learning, searching is done to detect humans in visualization of 3-dimensional space and Kinect has the ability to see in 3D through depth sensors that have skeletal tracking capability, namely processing depth images to get the joint position of the human skeleton and optimize the tracking of objects facing the Kinect to be used as a skeleton. skeleton tracking is successful which forms the skeleton in the human hip and the second part of the thighs, calves up to the soles of the feet. The purpose of forming the skeleton of the body part is to make the user or user to observe the movement of the body (legs). The third stage is the detection of skeleton in the SLR movement. This stage is to get the SLR skeleton shape on the lying object. Following are RGB-Depth and Skeleton images in Fig. 5(a) and Fig. 5(b):



(a)



(b)

Fig. 5. (a). RGB-Depth, (b). Skeleton Images.

The skeleton drawing algorithm with the drawBone() method that accepts two joints is tracked and draws a line between the joints. Skeleton drawing between joints can be done as needed. Whenever it requires depicting a joint of the entire human body, sequences need to be carried out, starting from the head to the foot, both from the left and right side of the body. Kinect generates skeleton data in the form of SkeletonStream, which can be set as standard tracking mode (default) or seated tracking mode using SkeletonTrackingMode enumeration during skeleton stream initialization. The process of flow (stream) to capture data skeleton remains the same as used for data streams of color and depth.

B. Skeleton Tracking Coding in Library Kinect

1. Attach and call the *SkeletonFrameReady* event handler
2. Check whether the skeleton frame is active to receive skeleton data.
 - If *SkeletonFrame = null* then returns to the 'Kinect Device Initiation' algorithm
3. Is the skeleton frame active?
 - If *skeletonFrame <> null*
 - Determine the number of skeletons tracked by:
skeletons = new Skeleton
[skeletonFrame.SkeletonArrayLength];
4. Copy the skeleton frame data to local memory
skeletonFrame.CopySkeletonDataTo (skeletons);
5. Draw Bone / skeleton in JointCollection from Head to foot right that is Head – FootRight
SLR skeletonization algorithms in chart box in Fig. 6.

- Draw a Bone / skeleton in JointCollection from the head to the middle chest:
Drawbone skeleton joints. Head-ShoulderCenter,
Drawbon skeleton joints.ShoulderCenter- ShoulderLeft,
Drawbone skeleton joints.ShoulderCenter-ShoulderRight.
- Draw a Bone / skeleton in JointCollection from the left shoulder to the middle of the chest to the right elbow:
Drawbone skeleton joints. ShoulderCenter- ShoulderRight,
DrawBone skeleton joints. ShoulderLeft-ElbowLeft,
DrawBone skeleton joints.ShoulderRight-ElbowRight,
DrawBone skeleton joints ElbowLeft–WristLeft,
DrawBone skeleton joints.ShoulderRight-ElbowRight.
- Draw bone from the left elbow joint into the spine:
DrawBone skeleton joints ElbowLeft–WristLeft,
DrawBone skeleton joints ElbowRight – WristRight,
DrawBone skeleton joints WristLeft – HandLeft,
DrawBone skeleton joints WristRight – HandRight,
DrawBone skeleton joints ShoulderCenter– Spine.
- Draw bone from the spinal joint to the right knee:
DrawBone skeleton joints Spine - HipCenter.
DrawBone skeleton joints HipCenter- HipLeft,
DrawBone skeleton joints HipCenter – HipRight,
DrawBone skeleton joints HipLeft – KneeLeft,
DrawBone skeleton joints HipRight – KneeRight.
- Draw bone from the left knee joint to the right leg:
DrawBone skeleton joints KneeLeft – AnkleLeft,
DrawBone skeleton joints KneeRight – AnkleRight,
DrawBone skeleton joints AnkleLeft - FootLeft,
DrawBone skeleton joints AnklerRight - FootRight.

Fig. 6. SLR Skeletonization Algorithms.

IV. RESULT

After the calibration process is complete and the algorithm detects that the user is a candidate, then the skeleton detection process and accesses any joint information or draws something above the depth image. Test cases were carried out on 10 human objects with different people of different body sizes and height. The picture illustrates the movement of the foot up, and when needed to stop, when used by doctors, the

leg movement stops until the pain center. Images are generated by categorizing in depth images and RGB images and their skeleton in Fig. 7.



Fig. 7. Images of Depth, RGB and Skeleton.

V. CONCLUSION

The results of testing data on human objects that perform successful SLR movements can track the skeleton of both legs. The results of this skeletonization in addition to providing information about the potential of Microsoft Kinect as a tool that can determine the skeleton between the two legs can also later use the skeleton as a basis for calculating the degree of SLR ROM on the SLR test. In this study, there were 4 patients with spinal disorders, it is hoped that in the future more can be done in patients with conditions such as low back pain. In addition, it is also expected to be able to compare manual measurements and SLR ROM measurements in realtime with the help of other than Kinect as an application. Tracked joint position data can aid in the diagnosis of the cause of a lower lumbar disc herniation and can be used to aid psychotherapy training.

ACKNOWLEDGMENT

The authors would like to thank The Directorate of Research and Community Service - Ministry of Research, Technology and Higher Education (DRPM - Kemenristek / BRIN) for financial support of this experiment., Contract number 309/SP2H/LT/DRPM/2021, 18 March 2021.

REFERENCES

- [1] Kit, Chui Chun, Kinect in Physiotherapy. <http://i.cs.hku.hk/fyp/2012/fyp12013/application.html>.
- [2] Gaston O. Camino Willhuber, Nicolas S. Piuzei, "Straight Leg Raise Test", NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health. StatPearls Publishing; 2019.
- [3] Fritz, J., PhD, ATC, in Physical Rehabilitation of the Injured Athlete (Fourth Edition), 2012. <https://www.sciencedirect.com/topics/neuroscience/straight-leg-raise>.
- [4] Gregory, M., Khadir, S.A., Thomas, E., Sillo, O., 2016, Straight Leg Raise Test, http://www.physio-pedia.com/Straight_Leg_Raise_Test, tanggal akses 10 Maret 2016.
- [5] Karnath B., "Clinical Signs of Low Back Pain": pp.39-44,56, 2003.
- [6] A. Jana, "Kinect for Windows SDK Programming Guide", Packt Publishing, Birmingham, Mumbai, 2012.
- [7] Hong, S., Jung, H., Seo, S., 2016., Real-time AR Edutainment System Using Sensor Based Motion Recognition, International Journal of Software Engineering and Its Applications. Vol. 10, No. 1 (2016), pp. 271-278.
- [8] N. Kitsunezaki, E. Adachi, T. Masuda, J. Mizusawa, "KINECT Applications for The Physical Rehabilitation". 2013.
- [9] Pathak, D., Bhosale, V., K., "Fall Detection for Elderly People in Homes using Kinect Sensor". International Journal of Innovative Research in Computer and Communication Engineering. Com. Vol. 5, Issue 2, February 2017.
- [10] Galna, B. Barry, G. Rochester, L., "Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease." Published on Apr 1, 2014.
- [11] Wen-June Wang, Jun-Wei Chan, Shih-Fu Haung and Rong-Jyue Wang, "Human Posture Recognition Based on Images Captured by the Kinect Sensor", International Journal of Advanced Robotic Systems, 2015.
- [12] Nitescu, D. "Evaluation of Pointing Strategies For Microsoft Kinect Sensor Device," 2012.
- [13] Rob Miles. "Learn Microsoft Kinect API." United States of America: Programming/Microsoft Kinect, 2012.
- [14] Shotton, J. "Real-Time Human Pose Recognition in Parts from Single Depth Images", Conference Paper in Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition • June 2011.
- [15] Schafer, RC, "Lumbar Spine, Pelvic and Hip Injuries", Chiropractic Management of Sports and Recreational Injuries, ACA Press, 2013.

MAGITS: A Mobile-based Information Sharing Framework for Integrating Intelligent Transport System in Agro-Goods e-Commerce in Developing Countries

Stivin Aloyce Nchimbi¹, Mussa Ally Dida²
Janeth Marwa⁴, Michael Kisangiri⁵

School of Computational and Communication Science and
Engineering^{1, 2, 5}

School of Business Studies and Humanities⁴
Nelson Mandela African Institution of Science and
Technology, Arusha, Tanzania

Gerrit K. Janssens³

Research Group Logistics
Hasselt University
Diepenbeek Campus, Belgium

Abstract—The technological advancement in Intelligent Transport Systems and mobile phones enable massive collaborating devices to collect, process, and share information to support the sales and transportation of agricultural goods (agro-goods) from farmer to market within the Agriculture Supply Chain. Mobile devices, especially smartphones and intelligent Point of Sale (PoS), provide multiple features such as Global Positioning System (GPS) and accelerometer to complement infrastructure requirements. Despite the opportunity, the development and deployment of the innovative platforms integrating Agro-goods transport services with e-commerce and e-payment systems are still challenging in developing countries. Some noted challenges include the high cost of infrastructure, implementation complexities, technology, and policy issues. Therefore, this paper proposes a framework for integrating ITS services in agro-goods e-commerce, taking advantage of mobile device functionalities and their massive usage in developing countries. The framework components identified and discussed are Stakeholders and roles, User Services, Mobile Operations, Computing environment with Machine Learning support, Service goals and Information view, and Enabling Factors. A Design Science Research (DSR) method is applied to produce a framework as an artifact using a six-step model. Also, a case study of potato sales and transportation from the Njombe region to Dar es Salaam city in Tanzania is presented. The framework constructs the ability to improve information quality shared among stakeholders; provide a cost-effective and efficient approach for buying, selling, payment, and transportation of Agriculture goods.

Keywords—Intelligent transport system; stakeholders; mobile phone; agro-goods; information sharing; agriculture supply chain; machine learning

I. INTRODUCTION

The availability and sharing of quality information are of paramount importance to stakeholders involved in the whole process of transportation of agricultural goods (agro-goods) from farms to the markets or buyer locations. Traditional Transport Management Information System (TTMIS) can offer

such solution but has a limitation in achieving maximum efficiency such as the ability for users to access the location of the Transporter instantly, driving behavior of the driver, timely access to weather condition in the position of agro-goods, traffic congestion status with alternative routing in cities and precision in the estimation of the delivery time.

The Intelligent Transport System (ITS) introduction offers vast opportunities that address the limitations observed in TTMIS. ITS mainly utilizes Information and Communication Technologies and sensors to ensure cleaner, efficient and safer road transportation of freights and Passengers [1]– [3]. The ITS allows Information Systems to make decisions that will benefit stakeholders intelligently. However, the requirements for development and deployment vary from one country to another due to different factors such as geographical structure, the way transport infrastructure is planned, the availability of technologies, social and economic differences.

The adoption and use of ITS have increased tremendously, since its inception in Japan and later in the USA and Europe over three decades [4]. The importance of ITS cannot be undermined. The global ITS market counted to 26.68 billion USD in 2019, and it is predicted to reach a 5.8% annual growth rate from 2020 to 2027[5]; but, the COVID-19 pandemic affected the market growth predictions. Significant factors influencing such change include the growing number of road vehicles, lack of traffic data, increasing demand for real-time information exchange in logistics, and advancements in big data, Artificial Intelligence (AI), blockchain, and the Internet of Things (IoT).

The developments and impacts of ITS have been realized high in developed countries [6], especially in transportation services, such as autonomous vehicles and applications to smart cities. There is little focus on Freight Intelligent Transport Systems (FITS) when viewing local contexts in developing countries. Some of the main challenges associated with FITS are the lack of transparency in data sharing to keep a competitive advantage against business rivals [7], [8]. Even

though the adoption of blockchain technology promises to address this challenge as explained by [7], [8], high implementation cost and complexity of the nature of freights in terms of handling requirement parameters such as size, weight, type (hazardous vs. non-hazardous, perishable vs. non-perishable, etc.).

ITS technologies for handling freight transportation include ubiquitous computing, IoT, wireless sensor networks, RFID, GPS, GPRS, GIS, artificial intelligence, and mobile phones. The increasingly mobile phone usage is crediting ITS development. A recent report reveals that out of 7.8 billion of the total world population, 4.88 billion mobile phone users worldwide, having 3.8 billion are Smartphone users. Sub-Saharan Africa recorded 34% smartphone connection in 2017. It is estimated to reach 67% by 2025 [9]. Some of the fueling factors are lower cost of Internet, cheaper devices, and increasing realization of the service provided by smartphones over the regular phone. This increased penetration of mobile phones, especially smartphones as sensors, has changed the way businesses and services are conducted globally. The ubiquity, power, and versatility of smartphone platforms have become valuable data generators in commerce, transport [2], financial services, agriculture [9], and infotainment.

Despite the advantage offered by mobile phone users globally, there is little focus on applying Mobile phone services integrated with ITS and e-commerce platforms the Agriculture Supply Chain (ASC) to improve transport efficiency of agricultural goods between stakeholders. In developing countries like Tanzania, the application of ITS is at its infancy with limited ITS infrastructure, operational standards, and governing policies. Evidently, there are existing e-commerce platforms linking farmers to markets. For example, Tanzania Mercantile Exchange Market (<https://www.tmx.co.tz>), which is government-owned platform; mkilimo (<http://mkilimo.esrf.or.tz>), which uses USSD code; tigo kilimo using USSD, SMS and voice; WeFarm (<https://about.wefarm.com/platform>) focusing on livestock, and Connected-farmer, which aims to improve agri-business efficiency of smallholder farmers. Although the platforms efficiently operate in Tanzania, many available systems lack integration with transport services. Therefore, users still have limitations in terms of the ability to search, acquire, and manage transport information of agro-goods as.

The contribution of this paper is to propose a mobile phone-based framework integrating intelligent transport systems and e-commerce as a practical approach to assist stakeholders in sharing information during the planning and execution of agro-good sales and transportation. The framework uses mobile phones as a multifunctional tool for solving the challenge of scarce technical resources to facilitate information sharing. Also, the paper identifies and analyzes the stakeholders and needs based on their roles in the system. The study further addresses the technological challenges in the limited access and use of transport information for farmers identified in [10]; and removing missing information to link farmers and other stakeholders, information barriers caused by intermediaries in ASC, identified in [9], [11]. The framework will not cover the internal warehousing operations with

assumptions to be handled by Enterprise Resource Planning (ERP).

The rest of this paper compasses the case study as presented in Section II and related work in Section III. Section IV discusses the methodology used in this study, followed by a proposed framework as articulated in Section V. The illustration and evaluation of the framework are described in Sections VI and VII, respectively. Jointly, the conclusion and recommendation are given in Section VII.

II. CASE STUDY

A. Background

In Tanzania, agriculture is the backbone and the leading contributor of the national economy accounting for 28.2% of the GDP in 2018[12]. The total population estimates at 58 million [13], but [14] estimates at 61 million in 2021. About 65% of the entire labor force lives in rural areas and depends on agriculture for daily living. It is also estimated that 80% of the concentrated rural farmers are smallholder farmers. The survey report by [15] indicated that Tanzania produced 68,000 tons of Irish potato in 2019; the production increased by 59% to reach 108,000 tons in 2020. Recent studies show an increased demand for potato consumption in urban areas to process French fries, also known as chips [16], [17].

Road transportation is the primary means for delivering potatoes to the consumers. It is estimated that 75% of the total freights are transported through an estimated 86,472 km road network [18]. In the current practice, the transportation of Potato agro-goods flows from rural to urban. This study focuses on the effective and efficient use of the limited available resource on Potato agro-goods transportation from the Njombe region to the Dar es Salaam region. The selections of the areas rely on the fact that Njombe is the leading producer, and Dar es Salaam is the primary consumer of Irish potato.

In the Njombe region, as a rural area, the study is based on two districts of Makete and Wanging'ombe. Wherein Dar es Salaam, as the urban area, the study concentrated in Ilala, Ubungo, and Kinondoni districts because of the availability of concentrated major potato marketplaces where potatoes from Njombe are delivered. Potatoes will be transported from Njombe rural road network through the highway to Dar es Salaam City with urban roads.

The sales, purchases, and transportation of agro-goods from farmers (in Njombe) to buyers (in Dar es Salaam) require seamless information sharing between stakeholders. Mobile phone technologies and functionalities provide a unique chance to simplify the processes. For example, the Tanzania Communications Regulatory Authority (TCRA) quarterly subscription report [19] recorded 51.2 million mobile network subscribers, 32 million mobile money subscribers, and 28.4 million internet subscribers by December 2020. Such information reveals the stimulating environment for mobile phone use in accessing services. In addition, mobile phones require sharing information with other devices such as vehicle tracking systems, intelligent cameras, etc., to build IoTs. Tanzania has a machine-to-machine (M2M)/IoT management guideline to create a better environment for exchanging

information [20]. The guidelines are vital in using mobile phones to develop an Intelligent Transport System to facilitate agro-goods sales, purchase, and transport.

B. Challenges

In the ASC, the enabling environment and features provided by mobile phones assure services such as seamless linking of farmers to buyers or markets, the provisioning of inputs for farming, and the formal and informal exchange of agricultural information and recommendations [21], [9], [22], [10]. Many existing ICT initiatives focus on market, pest and disease, weather, soil, financial, insurance, and technical advisory information [11]. Despite usefulness, the existing solutions lack transport integration as the fundamental aspect to ensure the movement of agro-goods from one point to another within ASC stakeholders. The transport integration challenges have also been expressed in the paper [10]. The road transportation of agro-goods is currently inefficient due to a lack of information sharing among agro-goods stakeholders. Since Tanzania is characterized by smallholder farmers, lack of information provides middlemen the opportunity to set prices for farmers [23], blocking the market's bargaining power. Also, transporters on their business face a similar challenge of middlemen while linking with buyers. In particular, tracing what, when, why, where, how, and who has transported agro-goods is a severe challenge [11]. Smallholder farmers have to wait for a long time for the transporters to collect small agro-goods from other sources to meet truckload requirements. Hence, delivery time becomes uncertain, and transport charges are becoming higher due to a lack of information sourcing. Besides, data on transport and agro-goods that eventually illustrate the actual demand for goods is inadequate.

III. RELATED WORKS

A. Mobile Phone Usage in Agriculture Supply Chain

In developing countries, mobile phones have contributed to agriculture and rural development by providing access to information, knowledge, services, and market [24], [25], [26]. The mobile phone penetration has further streamlined agro-goods supply and demand. It reduces post-harvest loss through real-time information exchange between agriculture stakeholders, makes delivery more efficient, and forces closer links between farmers and consumers. In Malaysia, the study

by [7] proposed five linking participants and their roles engaging in the physical flow of the pepper agri-food supply chain. At each stage connecting the participants, the third-party logistics (3PL) provider ensures an exchange of agri-food information. Tanzania largely depends on the movement of agricultural goods from rural to urban and involves several stakeholders or Actors. A study by [27] presented a typical chain from Producer, Collection Distribution, Marketing, and Consumption as modified in Fig. 1. A clear description of the Transport Service Client (appear as Transport user) and Transport Service Provider is found in the paper [28]. In the process of developing a framework for accessing market information for farmers, the recommendation for improving ICT connectivity, electrifying rural areas, and increasing awareness of ICT issues were provided [27]. The author posits that minimizing communication costs will motivate rural farmers to purchase mobile phones as an effective tool for accessing marketing information. A survey performed in the Iringa region –Tanzania revealed that mobile phones are not strongly impacting the yields; the agriculture stakeholders positively perceive mobile phone use in agricultural tasks such as accessing information about fertilizer transportation [9]. Hence, call for policy intervention on promoting the use of the mobile phone. While mobile phones, especially smartphones, have become more popular and affordable among agriculture stakeholders in Tanzania [10], [19], their ability to access and disseminate information has positively impacted efficiency, reduced transaction costs, and increased income among ASC stakeholders.

B. Mobile Phone and Freight Intelligent Transport System

Before digging further, it is an excellent reminder to note that agricultural goods (Agro-goods) are agricultural yields for commercial purposes whose source details have come from the farmer and can be perishable or non-perishable. Transportation of perishable goods differs from non-perishable in handling during transport and monitoring parameters such as temperature and humidity, and delivery time. In such a situation, accessing real-time information impacts the value of agro-goods in transport. Mobile phones, especially smartphones, have been the best approach because they provide multiple functionalities using installed sensors to deliver the required information while also interacting with users using mobile applications.

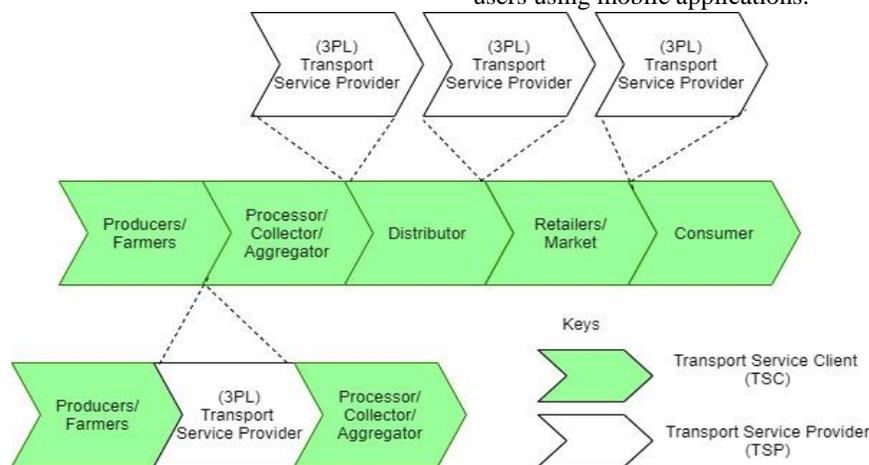


Fig. 1. Information Sharing between TSC and TSP in the Agro-goods Supply Chain (Modified from [7], [27]).

The argument is also supported by [29], posit that smartphone sensors such as Magnetometer, GPS, and accelerometer are critical in developing mobile applications in the Intelligent Transport System environment to capture the compass, location, and speed. Previous studies have complemented GPS-based traffic information exchange between vehicles and infrastructure specifically for road transport performance using GPS and other Intelligent Transport System sensors. In [30], a study demonstrated using GPS data collected from freight trucks for dynamic route planning and visualization using various traffic information such as travel time, route planning, and other congestion indices from the FITS-based application. The vehicle identification, location, and timestamp of record were captured. The study used the floating phone to observe long path trips. A web was developed to fit in more complex logistic business applications and scientific research usage.

In [31], a study identified low information utilization and flow, lack of timely and effective feedback of data, long time, low efficiency, high cost, inadequate security monitoring, and difficulties to meet product quality, food safety, and the market needs are the main challenges to the agricultural goods cold chain transport. The approach to employing an Intelligent Transport System based solution is proposed. The solution combines consumer and driver in the transportation of agricultural goods by designing a mobile phone-based system integrated with RFID and GPS for real-time information sharing to monitor the flow of perishable agricultural products. The author also designed an alarm for temperature and humidity during transportation as key quality variables for perishable agro-goods. A mobile payment Module to facilitate electronic commerce was developed using a two-dimension code to scan the code to transfer payment on goods delivery. With such an approach, tracing the reverse logistics chain is possible as all vendors are captured through the system during the forward chain.

Chapter six of the book [32], contributed by Charalambos A. Marentakis, provides a detailed view of telematics for transportation and agri-food product distribution. In the study, the author gives an ICT tool to offer monitoring and interaction capabilities between vehicle drivers and freight stakeholders from planning to monitoring and reporting need to ensure just-in-time (JIT) delivery, quality preservation, and efficiency. The author posits that the requirements for freight transportation vary depending on the nature of the operation. The author further added that e-commerce, e-business, and enterprise resource planning (ERP) directly influence the transport of agri-food in terms of performance and ability to handle large amounts of data generated through position sensing such as GPS and other sophisticated Information Systems applied in the transportation of agri-food. The author identified some of the requirements such as vehicle location, state of transported goods, including its temperature and humidity, time of arrival and time of departure for loading and unloading points, distance driven, length of the personal break, refueling stations, freight weight, transportation costs as well as vehicle, driver and trailer information. Advanced portable devices such as mobile phones provide features and functionalities for instant data capturing, such as barcode scanning, SMS, proof-of-

delivery, and reconciliation. Also, their ubiquities offer movable information-sharing services among participating stakeholders. The author proposed architecture for telematics applications integrating stakeholders, freight with their particular requirements, position sensing, tachograph operation (speed, stop, and idle), and shared transportation information portal. The author introduces Shipper, Customers, Recipients, Forwarders, Carrier, Vehicle Owners, and Fleet Owners as the users and stakeholders of the system.

A national large and medium agricultural-fresh products management system to monitor the transport of agricultural products considering the number, type, environmental-related temperature and humidity, vehicle speed, routes, locations, and driver information parameters was successfully designed, developed, and tested [33]. The main objective is to minimize the damages of agricultural products caused by transportation flowing from the original point to the destination/sales points/wholesale markets. The system comprises four subsystems focusing on testing environmental-related variables, mainly the temperature and humidity of agriculture products, collecting speed and line of the vehicle, wireless data transmission, and data control center. The study uses GPS for positioning services, GPRS, and the Internet for communication between the vehicle on-board unit (OBU) and the control center. ZigBee wireless network technology was applied for receiving temperature and humidity and communicates to the control center. Although the system met the requirements, the system required enhancement of planning functionalities, optimization functionalities, and emergence assistance services.

C. Existing Freight based Intelligent Transport System Frameworks

While looking for the possibility of the plan and manage co-model freight transport without a centralized system, [28] hypothesize that all transport services providers can publish their information on the Internet in a standardized format. The aim is to allow efficient sharing of accurate information between transport service providers and transport users. To achieve that, the authors argue on the efficient integration of all transport services and users in the transport chain and the lower cost of new connections and accessing the Internet. The ARKTRANS (Norwegian framework architecture for multimodal ITS) methodology to introduce abstraction levels ranging from transport policy, overall concept, logical aspect, and communication aspect together with Rapid Unified Process (RUP) were used. The RUP is used for an architecture-based iterative software development process. The four stakeholders identified were Transport Service Provider, Transport user, Road Network Manager, and Transport Regulator. The identified attributes were based on cost, status, from-to addresses, cargo details, and the reference document to show the Transport User and Transport Service Provider agreement.

In [34], freight brokering system architecture integrates web services and agents combining freight owners and transport services providers using web services integration gateway (WSIG) in JADE, two-layered architecture WS2 JADE RESTful web services were proposed. Users can register and post their products and offers in the public directory

through the virtual platform process. Potential customers can then browse, search and manually inspect the posted products that suit their needs. The available information contains real-time information on transportation and goods. Through employed intelligent services, the system included the provision of transport needs such as optimal transport routes focusing on requirements of the customers and the availability of free transport. The system can also negotiate with transport providers and customers to efficiently choose among the many delivery segments. Each element is covered, and no vehicles travel without cargo on the schedule. The study proposed the Customers, Transporters, and Brokers. Also include Payment, Statistics, and Email-notification service providers.

Also, a step-by-step approach to ITS system architecture for developing countries from scratch while considering limited resources in the countries established in [34]. The author emphasized adopting a framework from developed countries where affordability, compatibility and integration, geopolitics, and technical capabilities are the main identified criteria for building the ITS framework. It is proposed that the list of ITS user services be derived from existing services from developed countries; however, user services may differ in every country depending on the user's needs. Therefore, for countries like Tanzania, broad coverage of mobile networks and the massive availability of users amplify the effort toward implementing such services. The author further proposed to create a common data model, establishing the communication standard, using available communication technologies, and promoting standardization.

IV. METHODOLOGY

The road to the described above is complex and requires extensive multidisciplinary knowledge. Therefore, different documented approaches presented by [4], [35]– [40] were reviewed to establish an appropriate method for the desired framework. As a result, this study adopted Design Science Research. The selection of the Design Science Research Model (DSRM) [40], as inspired by [41], is based on its appreciation to produce new knowledge and seeking for innovations in solving problems. The six steps identified in DSRM are 1) problem identification and motivation, 2) defining the solution's objective, 3) creating the artifact, 4) demonstration, 5) evaluation, and 6) communication.

STEP1: The study used an observation approach, domain knowledge acquired, challenges identified from the case study, the need to use Intelligent Transport System in ASC, and literature reviews. Detailed information is available in sections I, II, and III.

STEP2: Through literature review, different solutions were proposed. For the case study presented in this paper, the objective is to develop a new framework using the mobile phone as a multifunctional tool for developing and managing stakeholders' information sharing systems. Also, the FREIGHTWISE Reference model [39] was extended to identify and analyze the stakeholders based on the roles and type of exchanged information while interacting in the ASC in developing countries. Sections II, III, and V describe more details.

STEP3: The artifact was developed as the main artifact of this study to foster innovation in the new products. The artifact, in this case, means the solution for the challenges, as detailed in Section V.

STEP4: The proposed framework was demonstrated using a scenario context mapping the existing potato agro-goods supply chain from Njombe rural, as the producers of potato, and Dar es Salaam, as main buyers. Section VI provides a detailed demonstration. A descriptive approach using case and scenario was used to illustrate the proposed framework [42], [41].

STEP5: The proposed framework evaluated in section VII. A total 12 experts used to provide opinion based on constructed framework (the artifact) based on challenges presented in the case study. The aim is to gain views on how it addresses the challenges.

STEP6: This paper is communicated to the audience the output of the study.

V. MAGITS FRAMEWORK

A proposed framework, MAGITS, is presented in this section. The MAGITS is abbreviated from a Mobile-based Agro-goods Intelligent Transport System. The framework integrates M-commerce, M-payment, and Intelligent Transport System. The following are primary objectives of the framework:

- 1) To address the existing gap of information sharing among stakeholders,
- 2) To simplifying the current complexity in the transportation of agricultural goods, consider differences in stakeholders' requirements in rural as well as urban transit,
- 3) To reduce the highly technological infrastructure demands through efficient utilization of mobile phones and additional functionalities such as GPS to share information among stakeholders.
- 4) To reduce the transport and intermediaries costs associated with routing through sharing of agro-goods information and easily deployable, readily available,
- 5) To achieve common language and standard information exchange patterns.

The MAGITS framework extends the Freightwise Framework (FWF) components [43]. The MAGITS framework, shown in Fig. 2, has seven components. These components are Main Services, Computing Center, Enabling factors, key performance goals (KPG), Mobile Operations, Operating environment, Stakeholders and their roles, and Information view. The enabling factors comprise the ITS Infrastructure, Policy, Management, Operating environment, Finance, communication viewpoints, and External services. In addition to the stated objectives, the MAGITS framework also uses mobile phone technologies to address the limitation of an on-board support, which was not part of the FWF Reference Model as stipulated in [28], [44], [43].

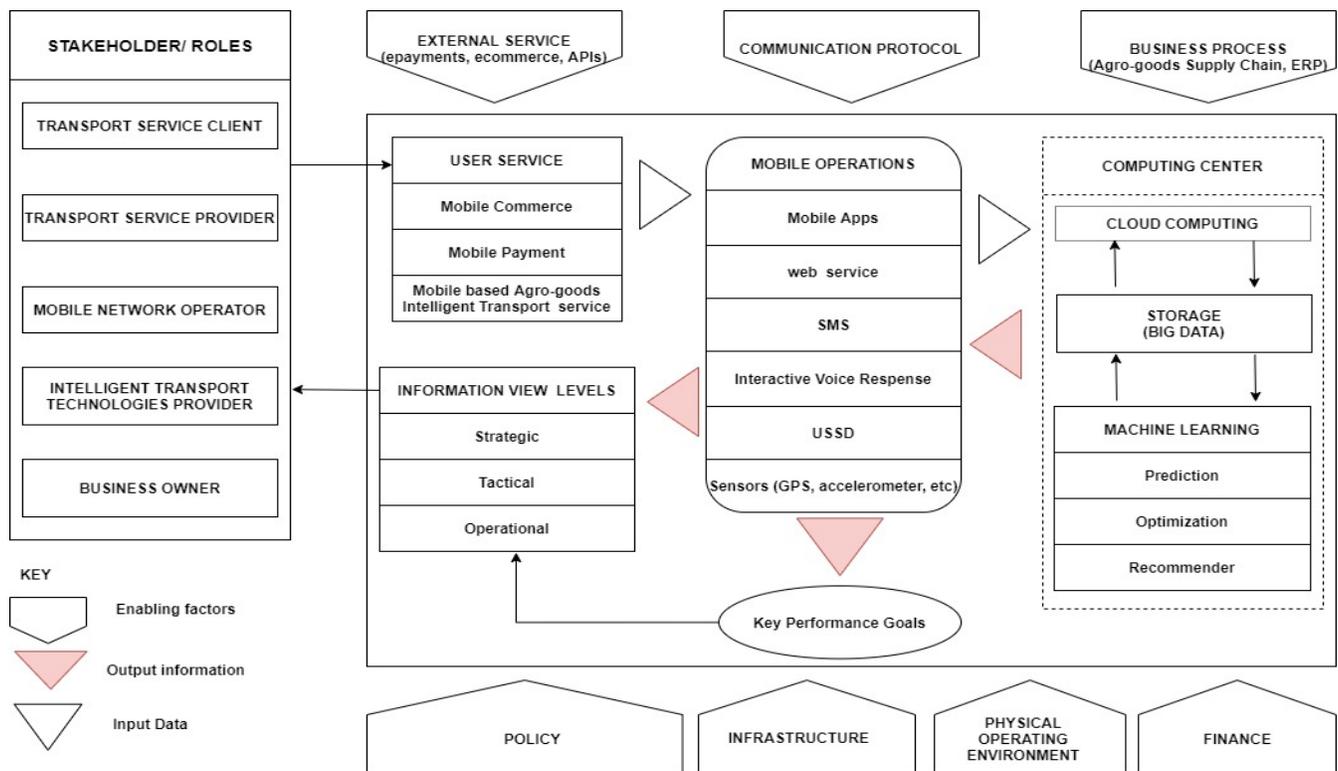


Fig. 2. A View of a Proposed MAGITS Framework.

A. Core Components of the Framework

1) *User services*: The core of the MAGITS framework is the services the stakeholders benefit from accessing and using the Systems developed under the framework. The MAGITS integrates three primary services, Mobile Commerce (M-Commerce), Mobile payment (M-Payment), and Mobile-based Agro-goods Transport Service, as central functions of the system. These services will use an auction-based approach to ensure fairness in the business operation. Based on [46], Researchers propose a most applicable first-price sealed bid as it creates privacy among competing bidders. The services are described in Table 1.

TABLE I. USER SERVICES

No	User Service	Description of service as per framework
1	M-commerce	Sell and buy agro-goods and services using mobile devices such as mobile phone or tablets
2	M-payment	Electronic money transaction for sell and buy of agro-goods or services using mobile devices
3	M-intelligent transport	Using mobile phone to hire and provide intelligent transport services for moving agro-goods between stakeholders in the Agriculture supply chain

2) *Computing center*: This framework component is responsible for capturing, analyzing, processing, storing, and sharing processed information between stakeholders. It ensures the security of information using technologies such as blockchain. The machine learning technology for optimizing and predicting services, big data, and cloud computing technologies provides the environment for accessing and using the platform as a service by reducing the high cost of procuring and maintaining the MAGITS hosting equipment. It also provides a secure mechanism for electronic data interchange (EDI) with external services components using Application Programmer Interface (API) and other decided tunneling approaches.

3) *Mobile operations*: Mobile phones, especially Smartphones, offer important integrated features such as GPS for location identification and WiFi and Bluetooth for wireless communication. These features are critical to ITS services due to the ubiquitous nature of mobile phones and enable stakeholders to access ITS services. For instance, the ITS-enabled sensors installed as the Onboard Unit (OBU) in vehicles wirelessly communicate with mobile phones to share information such as speed, fuel consumption data, mileage, and instant location information. The integration of such information from different ITS components resulting in improving timely delivery, accurate information about clients and providers of transport services, demand and supply optimization, prediction of loss and profit from transport business, tracking and tracing of cargo and vehicles, information of agro-goods to be delivered to the destination,

and transportation costs. Mobile phone-based information systems hosted in computer centers are accessed online through Unstructured Supplementary Service Data (USSD), web, Interactive Voice Response (IVR), and mobile apps to provide a unique environment to support such information lifecycle.

4) *Stakeholders and their roles*: These are institutional entities or persons responsible for operating, managing, using, or maintaining one or more framework services. Each stakeholder in the framework has a role to perform. Some stakeholders have more than one role. For example, a Customer can possess the role of TSC when in need of a transport service to deliver agro-goods to the destination. Table 2 describes the roles of each stakeholder, information to supply and lead to a demand from other stakeholders.

5) *Information view*: Focusing on stakeholders' needs for the MAGITS, the three categories of information flows are Electronic document information flow, payment information flow, and agro-goods information flow. The information flows between partners participating in the services and their processes. Understanding the information flow is critically important in completing the movement of goods from F to C, as defined in Table 2. Through the integration of information flow, the business processes and the depiction of the main objective of the design of any system are realized. It is the responsibility of stakeholders to feed timely and accurate information to the system to achieve actual results and efficient management and as part of sustainability strategy. The framework identified three levels of information: operational, tactical, and strategic. These levels of information assist stakeholders in deciding for their business.

TABLE II. STAKEHOLDER-ROLES AND INFORMATION REQUIREMENTS

Stakeholder	Roles	Information supply parameters	Information demand parameters
Farmers (F) or Producers (P)	Agro-goods producers, sellers, and supply chain initiators	Product Information(Farm-gate Selling Price, Location data, Time of harvest, etc	TSP, TSC, MNO, C
Customer (C)	The purchaser of agro-goods and requires goods to be transported to the desired location	Product information such as quantity in demand, type, delivery location, expected delivery time, Purchasing price, business permits, payment method, etc	F, MNO, RNM, TR, TSC and TSP
Transport Service Client (TSC)	Requires transport service from TSP to move agro-goods from F to C as shown in Fig. 1	Customer's information, product information such as location, quantity, type, payment method and expected delivery time, etc	F, MNO, RNM, TR, TSC and TSP
Transport Service provider (TSP)	Owner of commercial vehicle purpose(Individual/Company) and lease transport service to TSC or direct to C delivery of goods from F to C based on agreed terms	Carriage information such as type, quantity, and transport licensing. Also, vehicle location, current loading information such as FTL or LTL, routing information nearest to the loading point	ITTP, RNM, TR, TSC, MNO, P and C
Road Network Managers (RNM)	Plan, manage and maintain the road network	Route information, road toll charge, and payment models	ITTP, TSC, TSP, P, and C
Transport Regulators (TR)	Road rules enforcement and regulations checker	Law parameters, road charges, and payment approaches	ITTP, RNM, TSC, TSP, P and C
Mobile Network Operators (MNO)	Mobile network provision and mobile-based payment services	Internet charges, connection, and payment approaches	ITTP, RNM, TR, TSC, TSP, P and C
Intelligent Transport Technology Providers (ITTP)	Intelligent Transport System technology and service providers	Location (GPS), weather and communication, Application Programmer Interface (API), OBU, RSUs, etc	MNO, RNM, TR, TSP
Business Owner/System owner	A framework manager. A person or entity who legally owns the system. Responsible for setting the environment for other stakeholders to interact. Responsible for day to day internal management of a system	All information	All information

6) *Key performance goals*: The Key Performance Goals (KPGs) are the overall outputs of the framework. These KPGs directly affect the sustainability of the framework. The key goals of this framework are improving efficiency, increasing safety, reducing carbon emission, and introducing intelligence in agro-goods transport through the collection, store, model (prediction and optimization), and analyzing operational transportation data to improve future transportation investment planning. The lower the KPG rating, the shorter lifetime of a framework. Likewise, the higher the KPG ratings, the longer the sustainability of the framework. The KPGs are achieved due to stakeholder's satisfaction level with the system and services. Performance Indicators (KPIs) are applied to the MAGITS framework to measure its performance, including but not limited to a percentage of uptime of e-commerce per time, percentage of detected overloaded vehicles, average delivery time per trip, number of the shortest routes compared to available routes, rate of carbon emission, improved transport safety, number/amount of electronic toll collection and e-payments per vehicles/time, average transport cost or profit per trip, average loading/unloading time of agro-goods, number of pickup and delivery points per same truck, average available time of tracking and tracing services per vehicle, number of prediction modeling services, number of goods delivered/loss during transportation, and average available time of information among stakeholders.

7) *Enabling factors*

a) *Policy and Governance*: Policies dictate the direction of the business. Therefore, policies in place must harness the integration of Intelligent Transport System with e-commerce and e-payment services to promote the development and deployment of the framework. The policies include how data will be shared between participating stakeholders at lower cost and timely, enabling the usage of services to impact the individual and national economies, strengthening sharing of business information by reducing unnecessary barriers from competing rivals. In addition, the presence of policy support validates the business. It assures participants of recognizing data generated from the system, i.e., electronic payments, contracts, and other e-documents, as legal evidence for business contracts.

b) *Communication Viewpoint*: The two communication types, namely short-range and long-range communication, are identified for this framework, as studied by [47]. The short-range communication protocol, commonly known as a Dedicated Short-Range Communication (DSRC) by IEEE, is a highly reliable, fast network acquisition, high interoperability, and low latency. They are, therefore, suitable to be applied for vehicular data exchange between the Vehicle-to-Infrastructure (V2I), especially roadside units (RSUs) and antennas, Vehicle-to-Vehicle (V2V), and Vehicle's OBUs to loaded cargo using wireless technologies such as ZigBee.

The long-range communication protocol is Cellular-Vehicle-to-Everything (C-V2X) PC5 technology, the cellular-based communication protocol. The detailed study by [48] pointed out two applicable communication protocols for ITS

are 802.11p-based, which contain ITS-G5. The 5G technology offers adaptive application-focused, increased speed, reduced latency (i.e., from 20 milliseconds to approximately one millisecond), and more service-oriented with 99.999 percent reliability. Still, many developing countries such as Tanzania still require more time to overtake the dominance of 3G and 4G in cellular long-range communication technologies.

c) *Business Processes*: The primary business processes are planning, execution and completion. For instance, in FREIGHTWISE[28][43], the planning phase is complete when the Transport Execution plan(TEP) is marked ready for execution. When proof of delivery has been issued, the execution phase is completed. In most cases, the business processes are integrated into the operating business's ERP. The Oasis Universal Business Language (UBL) – 2.3 [49]and ISO TC204 work groups[50] have established the standard documents that facilitate the development of systems related to freight transportation. It is easier for system developers to adopt such agreed design documents as baseline designs to conform to such recognized international standards. The UBL uses the XML standard for exchanging business data.

d) *Finance*: In-depth consideration of the investment cost, revenues, and expenditure must be made during the system's design, deployment, and operationalization. The factors such as payback period, minimizing expenses while increasing revenue, service charges such as internet-service-subscription charges, cost of equipment, labor charges, etc., are essential at every stage of the system lifecycle.

e) *Infrastructure*: In this context, infrastructure contains all physical assets that support the transportation of agro-goods to the consumer. These include roads, OBU, RSU, cameras, sensors, vehicles, agro-goods, Mobile phones, weighbridges, and road toll gates.

f) *Physical Operating environment*: The quality of products and services depends on the operating environment. It involves handling mechanisms in handling agro-goods physical flow to preserve expected quality. The forward physical flow of goods channels from Farmer as Producer/Seller to Collection Center Manager, Distribution Center Manager, and Consumer. Transport Service Provider is involved at each stage to pick up and deliver agro-goods successfully. Each stakeholder in the chain is responsible for preserving the quality of the product and feeding the system with accurate data such as GPS locations to meet the other stakeholders' expectations.

B. *High-Level Layered MAGITS Framework Architecture*

All information will be accessed centrally from the Computing Center using a client-server architecture. The idea is to achieve controlled data flow and reduce the data storage size in the stakeholders' smartphones or tablets and allow coordinated and controlled integration of information whenever needed by stakeholders' or other systems. The general concept of the MAGITS framework considers the use of micro-service architecture using a lightweight approach like application programming interface (API) to communicate different designed services in solving the management of information as the goods move from one point to another. The

advantages of using microservice architecture over typical monolithic architecture. The designed architecture has proven scalable, adaptable, and easy to maintain in the microservice approach, as shown in Fig. 3.

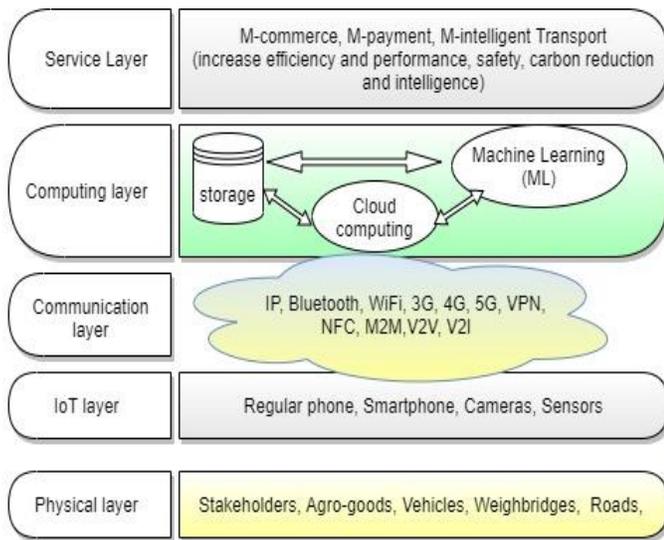


Fig. 3. High-Level Layered MAGITS Framework Architecture have been Defined by [45].

VI. ILLUSTRATION OF THE FRAMEWORK

As described in the methodology, this section illustrates the proposed framework using a descriptive scenario by combining the sales, purchases, payments, and transportations of potatoes in Tanzania. The Transport Service Clients (TSC) and Transport Service Providers (TSP) were selected for this evaluation phase. Fig. 4 illustrates the applicability of the framework. In the process, all stakeholders interact with each other using mobile phone services.

The mobile phone with GPS and mobile app services, such as Android and IOS, qualifies to be installed as vehicle OBU specifically for location serves as an input to the system. Each stakeholder will register in the system. Except for TSP whose locations are changing during transportation, all other stakeholders will register their GPS location. It is crucial in calculating the distance as the parameter for the vehicles. The majority of farmers are smallholders who can afford regular phones. The straightforward approach is for the collection center to register all the farmers located near their collection center using a Smartphone to capture the GPS coordinates of farms and other important information for later use. The essential criteria are that Collection Centers and Distribution Centers must directly connect to the main road to access all transporters. The framework also allows mobile phone information sharing from geographically dispersed smallholder farmers to aggregate potato agro-goods to meet the full truckload (FTL) concept.

The approach is essential for cost-sharing, providing access to markets, and increasing farmers' bargaining power in the markets. From Fig. 4, farmers as an initiator of the chain, feed information including quantity and price about potatoes using the mobile phone.

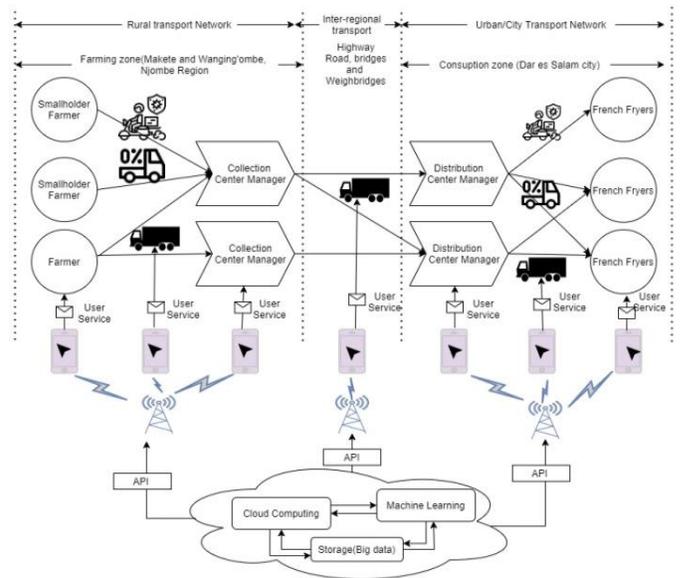


Fig. 4. Information Sharing Scenario of Potato Mobile-based Agro-goods Intelligent Transport System.

Since the location information was previously captured, the Collection Center will automatically receive notifications about farmers ready to sell using the mobile phone. The Collection Center Manager will confirm the request and advertise to French fryers (Fried Irish potato chips sellers) and other buyers. All buyers will confirm the order through a mobile payment system, where the system will send feedback to the collection center and farmer. The system will automatically publish the successful order to transporters showing parameters such as quantity, pickup location, and delivery location. For a specifically given time interval, each transporter will make a sealed bid for the service by manually filling the cost and expected delivery time.

The system will evaluate submitted offers and send a notification to the buyer to confirm the payment. Using the same payment system for paying transport costs, all parties will receive confirmation about the confirmation of the order. At each payment stage, the system will hold the amount until the buyer confirms the receipt of the potatoes consignment. The system keeps the amount on each payment while generating a secret code and sending it to the buyer's phone. The code will be entered into the mobile phone upon receipt of the order.

The application of Machine Learning (ML) in the proposed framework adds unique achievement to MAGITS. The applications of ML, namely optimization, prediction, and rating, are identified. For instance, each time stakeholders use the system will provide feedback about the service. The combination of rating, questions, comments, demographics, technographic, behavior, and geographic features will be processed using ML recommender modeling techniques to create patterns for decision-making purposes. These patterns aim to improve the system's services, logic, and decision-making. The model, using ML, will use GPS coordinates from smallholder farmers and the location of the nearest road; it is possible to optimize the nearest collection center accessible to the road and share transport to the buyer or markets. The

massive data generated from the use of the framework plays a critical role in predicting the demand and supply of potatoes in Tanzania.

VII. EVALUATION OF THE FRAMEWORK

A total of 12 experts were interviewed to provide expert opinions on the proposed framework, as shown in Table 3. One challenge encountered was the possibility of identifying an interviewee with all the required skills to evaluate the method.

All interviewees were given the paper using the walkthrough approach and establish views on the acceptability of the framework using the challenges as identified in Table 4. The comments were then discussed through a convened online meeting to agree or disagree with the proposed framework. From the discussion, all experts agree on the proposed framework approach to solve the identified challenges and suggest the need for developing web and mobile applications to integrate the proposed service for higher impact to the community, especially marginalized smallholder farmers. Experts also commended the simplicity of the ASC process using a system developed from the framework and suggested more research on the users' acceptability approach. Finally, the experts comment that many more researchers will benchmark this study.

All experts concluded that the framework works best in the presence of smartphones than regular phones.

TABLE III. EXPERTS BACKGROUND

Expert	Area of expertise	Experience in the industry (in years)
1	Agriculture	25
2	Electronic Business	13
3	Information Communication Technology	18
4	Information Communication Technology	12
5	Agri-business	18
6	Agricultural products Transportation	30
7	Information Communication Technology	16
8	Freight Intelligent Transport Systems	4
9	Agri-business	22
10	Smart Cities	5
11	Agriculture Extension Services	16
12	Internet Service Provision	20

TABLE IV. EVALUATE CHALLENGES

Challenge	How framework addresses the challenge?
Stakeholders' involvement	The framework considered all participants involved in the sales, purchase, and transportation of agro-goods.
Context-awareness coverage	Since agro-goods are produced in rural areas then transported to urban areas, the framework has incorporated the environment. The framework has also considered minimum enabling factors required during the design, deployment, and operations.
Resources scarcity	The additional use of mobile phone sensors acting as OBU and Point of Sales (PoS) minimizes the demands for deploying many devices where each performs a single task
Real-time information sharing capability	It is possible for stakeholders accessing and sharing sales, purchases, and payment information in real-time through installed mobile phones. Tracking and tracing of agro-goods loaded vehicles anywhere at any time.
Service demands	The Machine Learning integration in this framework is essential for travel route optimization, sharing of transport and transport cost for smallholder farmers, predicting agro-goods demand, and supply factoring location and time. The framework achieves the KPG indicated in this paper.

VIII. CONCLUSION AND FUTURE WORK

This study provides a unique guide for ASC practitioners linking technology to agro-goods business. It addresses the need for both Information Systems developers during the designers, government authorities in deliver service while seeking income generation, and business managers during investments and operations using limited resources. It also serves as a benchmark for further researchers focusing on developing new products supporting communities based on e-commerce and transportation using mobile phones to monitor and manage the supply chain.

The developed framework for integrating ITS in agro-goods e-commerce in developing countries is inevitable due to the high dependence on agriculture for economic and social survival. The requirements for the framework differ from one place to another, but the proposed framework unifies the high-level needs for successful and sustainable systems. A framework enables the design, implementation, deployment, and ITS in the Agriculture Supply chain in many developing countries from limited available infrastructure, especially mobile phones. For instance, it is possible to obtain traffic and weather information about particular route agro-goods transportation using data collected from other mobile phones located on in-transit vehicles using mobile phone sensors. On the TSP side, there are information limitation challenges that may arise when using the mobile phone as an OBU sensor, such as fuel consumption and break information will not be captured. Nevertheless, some recent vehicle OBU can transmit such data to mobile phones wirelessly using short-range communication such as Bluetooth and WiFi. Further research is required to use mobile phones when installed as OBU, such as power consumption, availability due to changes in environment, and durability.

In terms of Internet service, there is a need for harmonization between different MNOs (defined in Table 2). Therefore, the study establishes a service-focused internet subscription to achieve maximum service availability despite the vehicle in transit changes MNOs. One limitation of the proposed framework is the inability to handle multi-products in the same truck; therefore, the study recommends investigating and suggesting an appropriate ML modeling approach for managing more than one product within a single truck based on type, size, and weight. There is no evidence of Tanzania's framework to researchers' best knowledge. Still, researchers believe that implementing the framework will solve transport challenges of limited-resource many e-commerce platforms globally. This framework also provides the foundation for further research in big data in the transportation of agro-goods in developing countries.

REFERENCES

- [1] GSMA, "Intelligent Transportation Systems Report for Mobile," GSMA Connect. Living Program., p. 88, 2015.
- [2] S. TAO, "Mobile Phone-based Vehicle Positioning and Tracking and Its Application in Urban Traffic State Estimation," KTH Royal Institute of Technology, 2012.
- [3] ESCAP, "Development of Model Intelligent Transport Systems Deployments for the Asian Highway Network," Bangkok, no. December, 2017.
- [4] T. Yokota and R. J. Weiland, "ITS Technical Note For Developing Countries - Note 5 - System Architectures," World Bank Tech. Notes, pp. 1–16, 2004.
- [5] GVR, "Intelligent Transportation System Market Size, Share & Trends by Type (ATIS, ATMS, ATPS, APTS, EMS), by Application (Traffic Management, Public Transport), by Region, and Segment Forecasts, 2020 - 2027," San Francisco, 2020.
- [6] G. Gebresenbet and T. Bosona, "Logistics and Supply Chains in Agriculture and Food," Pathways to Supply Chain Excel., no. March 2012, p. 218, 2012, doi: 10.5772/25907.
- [7] K. Y. Chan, J. Abdullah, and A. S. Khan, "A framework for traceable and transparent supply chain management for the agri-food sector in Malaysia using blockchain technology," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 11, pp. 149–156, 2019, doi: 10.14569/IJACSA.2019.0101120.
- [8] P. Helo and Y. Hao, "Blockchains in operations and supply chains: A model and reference implementation," Comput. Ind. Eng., vol. 136, no. July, pp. 242–251, 2019, doi: 10.1016/j.cie.2019.07.023.
- [9] A. Quandt et al., "Mobile phone use is associated with higher smallholder agricultural productivity in Tanzania, East Africa," PLoS One, vol. 15, no. 8, p. e0237337, 2020, doi: 10.1371/journal.pone.0237337.
- [10] N. Newman et al., "Designing and Evolving an Electronic Agricultural Marketplace in Uganda," COMPASS '18 ACM SIGCAS Conf. Computing Sustain. Soc., p. 11, 2018, doi: <https://doi.org/10.1145/3209811.3209862>.
- [11] M. M. Mageza, "Linking Rural Farmers to Markets Using ICTs," Tech. Cent. Agric. Rural Coop., vol. 15/12, no. CTA Working Paper, 2015.
- [12] NBS, "National accounts statistics of Tanzania Mainland," 2019.
- [13] World Bank Group, "No Title," World population, Tanzania, 2019. [Online]. Available: <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=TZ>. [Accessed: 29-Apr-2021].
- [14] WPR, "World Population Review," World Population Review. [Online]. Available: <https://worldpopulationreview.com/countries/tanzania-population>. [Accessed: 29-Apr-2021].
- [15] NBS, "Highlights On The Third Quarter Gross Domestic Product (July – September) 2020, Base Year 2015," Dodoma, 2020.
- [16] D. H. Mende, K. A. Kayunze, and M. W. Mwatawala, "Contribution of Round Potato Production to Household Income in," J. Biol. Healthc., vol. 4, no. 18, pp. 1–11, 2014.
- [17] M. C. A. Wegerif, Feeding Dar es Salaam: a symbiotic food system perspective. 2017.
- [18] Tanzaniainvest.com, "Tanzania Roads," Report, 2016. [Online]. Available: <https://www.tanzaniainvest.com/transport/tanzania-transport-sector-report>. [Accessed: 30-Apr-2021].
- [19] TCRA, "Quarterly Communication Statistics," Dar es Salaam, 2020.
- [20] TCRA, "Guidelines on the Management of Applications of Machine To Machine Communication and Internet of Things Numbering Resources," 2019.
- [21] K. Masuki et al., "Mobile phones in agricultural information delivery for rural development in Eastern Africa: Lessons from Western Uganda," ACM, pp. 1–11, 2010.
- [22] A. R. Chhachhar and M. S. H. Md Salleh Hassan, "The Use of Mobile Phone Among Farmers for Agriculture Development," Int. J. Sci. Res., vol. 2, no. 6, pp. 95–98, 2012, doi: 10.15373/22778179/june2013/31.
- [23] N. Mramba, M. Apiola, E. A. Kolog, and E. Sutinen, "Technology for street traders in Tanzania: A design science research approach," African J. Sci. Technol. Innov. Dev., vol. 8, no. 1, pp. 121–133, 2016, doi: 10.1080/20421338.2016.1147208.
- [24] D. Tesh W and O. Chidi, "ICTs in Agricultural Production and Potential Deployment in Operationalising Geographical Indications in Uganda Tesh W Dagne and Chidi Oguamanam," no. October, 2019, doi: 10.2139/ssrn.3241169.
- [25] Omri Van Zyl, "ICTs for agriculture in Africa," eTransform Africa, 2012.
- [26] R. Syiem and S. Raj, "Access and Usage of ICTs for Agriculture and Rural Development by the tribal farmers in Meghalaya State of North-East India," J. Agric. Informatics, vol. 6, no. 3, 2015, doi: 10.17700/jai.2015.6.3.190.
- [27] M. M. Mageza, K. Michael, and J. Ko, "Towards a framework for accessing agricultural market information," Electron. J. Inf. Syst. Dev. Ctries., vol. 66, no. 1, pp. 1–16, 2015, doi: 10.1002/j.1681-4835.2015.tb00473.x.
- [28] T. Zunder, H. Westerheim, R. Jorna, and J. T. Pedersen, "Is it Possible to Manage and Plan Co-Modal Freight Transport Without a Centralised System?," Int. J. Appl. Logist., vol. 3, no. 2, pp. 25–39, 2012, doi: 10.4018/jal.2012040103.
- [29] A. Dharani, "Mobile as a Sensor in Intelligent Transportation System for Street Route," 2018 Int. Conf. Comput. Power Commun. Technol., pp. 138–141, 2019.
- [30] C. Yang, X. Ma, and Y. Ban, "Demonstration of intelligent transport applications using freight transport GPS data," Present. 95th Annu. Meet. Transp. Res. Board, 10-14 January 2016, Washingt. DC, USA, no. July 2015, pp. 1–16, 2016.
- [31] C. Chen, T. Chen, C. Zhang, and G. Xie, "Research on Agricultural Products Cold-Chain Logistics of Mobile Services Application," pp. 247–254, 2014, doi: 10.1007/978-3-642-54341-8_26.
- [32] M. Bourlakis, I. Vlachos, and V. Zaimpekis, Intelligent Agrifood Chains and Networks, 1st ed. Oxford: Blackwell Publishing Ltd., 2011.
- [33] L. Jianting and Y. Zhang, "Design and Accomplishment of the Real-time Tracking System of Agricultural Products Logistics Process," International Conference on E-Product E-Service and E-Entertainment, p. 4, 2010.
- [34] F. Leon and C. Badica, "A Freight Brokering System Architecture Based on Web Services and Agents," in 7th International Conference IESS 2016, 2016, no. May, pp. 537–546, doi: 10.1007/978-3-319-32689-4.
- [35] T. Alameri, S. Abudulmajeed, A. S. Shibghatullah, and M. M. Jaber, "Towards proposing an electronic information sharing model for the intelligence sector: A methodological framework," J. Eng. Sci. Technol., vol. 14, no. 3, pp. 1687–1702, 2019.
- [36] A. Di Febraro, N. Sacco, and M. Saeednia, "An agent-based framework for cooperative planning of intermodal freight transport chains," Transp. Res. Part C Emerg. Technol., vol. 64, pp. 72–85, 2016, doi: 10.1016/j.trc.2015.12.014.
- [37] H. G. C. Góngora, T. Gaudré, and S. Tucci-Piergiorganni, "Towards an Architectural Design Framework for Automotive Systems," Researchgate, no. December, 2012, doi: 10.1007/978-3-642-34404-6.

- [38] ADB, "Conceptual Design of the Intelligent Transport Systems Project—Case in Gui'an New District," Asia Dev. Bank, 2019, doi: 10.22617/TCS190561-2.
- [39] M. Hagaseth, H. Westerheim, C. Antoniou, G. Tsoukos, and Christina Paschalidou, "Management Framework for Intelligent Intermodal Transport," Architecture, no. Project no.TREN/06/FP6TR/S07.60148, p. 72, 2008.
- [40] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A Design Science Research Methodology for Information Systems Research," J. Manag. Inf. Syst. Vol., vol. 24, no. 3, pp. 4201–4204, 2007.
- [41] Z. Rehena and M. Janssen, "Towards a Framework for Context-Aware Intelligent Traffic Management System in Smart Cities," Web Conf. 2018 - Companion World Wide Web Conf. WWW 2018, pp. 893–898, 2018, doi: 10.1145/3184558.3191514.
- [42] A. Dennis, R. Jones, D. Kildare, and C. Barclay, "Design science approach to developing and evaluating a national cybersecurity framework for Jamaica," Electron. J. Inf. Syst. Dev. Ctries., vol. 62, no. 1, pp. 1–18, 2014, doi: 10.1002/j.1681-4835.2014.tb00444.x.
- [43] K. E. Fjørtoft, H. Westerheim, A. Vennesland, and Marianne Hagaseth, "FREIGHTWISE Framework Architecture," MARINTEK, vol. 44, no. 2, p. 164, 2006.
- [44] M. Natvig, H. Westerheim, T. K. Moseng, and A. Vennesland, "ARKTRANS The Norwegian framework architecture for multimodal ITS," 2009.
- [45] I. Shabani, E. Mëziu, B. Berisha, and T. Biba, "Design of Modern Distributed Systems based on Microservices Architecture," vol. 12, no. 2, pp. 153–159, 2021.
- [46] C. A. Marentakis, "The Expansion of E-Marketplace to M-Marketplace by Integrating Mobility and Auctions in an Environment : Services,".
- [47] K. N. Qureshi and H. Abdullah, "A Survey on Intelligent Transportation Systems Market-oriented grid computing View project Automated Remote Healthcare Monitoring View project," Middle-East J. Sci. Res., vol. 15, no. 5, pp. 629–642, 2013, doi: 10.5829/idosi.mejsr.2013.15.5.11215.
- [48] D. Naudts et al., "Vehicular Communication Management Framework: A Flexible Hybrid Connectivity Platform for CCAM Services," Futur. Internet, vol. 13, no. 3, p. 81, 2021, doi: 10.3390/fi13030081.
- [49] UBL-2.3, "Universal Business Language Version 2.3," UBL-2.3, 2021. [Online]. Available: <https://docs.oasis-open.org/ubl/cs01-UBL-2.3/UBL-2.3.html>. [Accessed: 01-May-2021].
- [50] ISO/TC 204, "ISO TC 204 Intelligent Transport Systems — Data dictionary and message set to facilitate the movement of freight and its intermodal transfer — Road transport information exchanges for supply chain freight time-sensitive delivery (Road-Air Freight-Road)," 2005.

Software Project Estimation with Machine Learning

Noor Azura Zakaria¹, Amelia Ritahani Ismail², Afrujaan Yakath Ali³
Nur Hidayah Mohd Khalid⁴, Nadzurah Zainal Abidin⁵

Department of Computer Science
Kulliyah of Information and Communication Technology
International Islamic University Malaysia, Kuala Lumpur, Malaysia

Abstract—This project involves research about software effort estimation using machine learning algorithms. Software cost and effort estimation are crucial parts of software project development. It determines the budget, time and resources needed to develop a software project. One of the well-established software project estimation models is Constructive Cost Model (COCOMO) which was developed in the 1980s. Even though such a model is being used, COCOMO has some weaknesses and software developers still facing the problem of lack of accuracy of the effort and cost estimation. Inaccuracy in the estimated effort will affect the schedule and cost of the whole project as well. The objective of this research is to use several algorithms of machine learning to estimate the effort of software project development. The best machine learning model is chosen to compare with the COCOMO.

Keywords—Software effort estimation; project estimation; constructive cost model; COCOMO; machine learning

I. INTRODUCTION

Problems are created for software professionals, their clients, and stakeholders from the impractical project strategy and budget overruns. Despite many studies and numerous attempts to learn from experience, the problem of inaccurate often happen and has not been solved yet [1]. Software cost, effort, and resources are estimated at the beginning of the development. That information will be used by developers and clients to estimate the budget and time needed to finish develop an application or a system. Techniques and models were invented to assist the developers in estimating budget and effort. However, the problem of inaccuracy in estimation still becomes one of the problems for the developers and stakeholders. Even the emergence of one of a well-established project estimation model in the 1980s, COCOMO model, does not solve the problem of inaccuracy in software project estimation. Therefore, in this research, machine learning algorithms are used to estimate the effort of a software project that is more accurate compared to the COCOMO model. COCOMO model datasets are used to build machine learning models.

Although the COCOMO model has many advantages, it has some weaknesses too. One of it is its estimation varies as time progress [2]. Furthermore, the COCOMO model works depends on historical project data which are not available at all times [3]. COCOMO model cannot be used to estimate in all Software Development Life Cycle phases [3]. A large amount of data required for the COCOMO model to work [4]. A user has to insert input of 15 effort multipliers in order to get output from the COCOMO model. Thus, it will consume a lot of time

for industry that has to estimate a large number of projects. COCOMO has difficulty in learn and identify data patterns [4] which is an important element in the regression model such as COCOMO.

Our main three objectives are to pre-process and analyze the COCOMO dataset. Second, is to apply several algorithms and to predict the output based on the COCOMO dataset and to evaluate the performances of the selected algorithms with the COCOMO method.

An application called SOFREST Estimator is developed to demonstrate how the estimation work and what are the inputs that needed to produce the outcome. The application will require user to insert five inputs about a project, which are number of Lines of Code (LOC), Database Size (DATA), Required Software Reliability (RELY), Execution Time Constraint (TIME) and Main Storage Constraint (STOR). The output of the application will be estimation of effort that needed for that particular project in person-months unit.

This project is significant because it concerns the accuracy in predicting the budget and time needed to develop a whole project. By doing the project cost and time estimation using machine learning, higher accuracy of cost and effort of a software project estimation will be produced since, in machine learning, we build a prediction model by train and test the dataset. By having machine learning as the project cost and effort estimator, money and time can be saved as it will need less human effort. This project will be useful for every software development company for them to estimate the cost and effort of a project. The best algorithm to be used in this project cost and time estimation can be determined based on the highest classification accuracy in machine learning.

II. RELATED WORK

A. Software Project Estimation

Project estimation is an essential part of completing a project. Projects are planned in terms of cost, effort, and budget at the beginning phase of development. Precise effort estimation of software development plays a main task to predict how much workforce should be prepared during the works of a software project so that it can be completed on time and with the budget that planned without ignoring the quality of a software [5]. Accuracy of development cost estimation is a key factor in the success of a construction project and influenced the decision-making by the stakeholders of a software project [6] and to bid a contract with them [7].

The capacity of a budget estimating model is determined by calculating its bias, stability, and precision. Measures of bias, stability, and precision are concerned with the difference in the average between the actual costs and the estimated costs, considering both the degree of variation around the average and the combination with bias and consistency. By far, the most popular evaluation criteria used involve statistics such as mean, standard deviation, and coefficient of variation [6].

Identifying and calculating software metrics are important for various reasons, including estimating programming execution, measuring the effectiveness of software processes, estimating required efforts for processes, reduction of defects during software development, and monitoring and controlling software project executions [8]. An example of the wrong cost estimation that happened recently was in estimating the budget of international arrivals facility that being built at Seattle-Tacoma International Airport in Seattle, Washington, USA. Initially, in 2013 the budget was estimated at US\$ 300 million but then the budget increased up to US\$ 968 million in September 2018 [9]. Research shows that usually projects seem to be unclear at the beginning and become less vague as they progress [10].

One of the software metrics that used to estimate the cost and effort is called lines of code (LOC) metric and is considered basic software metric [11] as it is used in most software project estimation techniques.

B. Project Estimation Techniques

It is hard to quickly and accurately predict the development budget at the planning stage because the documentation is generally incomplete. For this reason, various procedures have been created to accurately predict construction costs with the limited project data available in the early phase [6]. There are three known models that have been used to estimate the project effort, cost and resources which are the Constructive Cost Model (COCOMO), Analogy-based Model, Use Case Points model.

1) *COCOMO Model*: COCOMO (Constructive Cost Model) is a screen-oriented, interactive software package that assists in budgetary planning and schedule estimation of a software project [12]. The intermediate COCOMO model used 15 drivers to estimate the cost of a project. The drivers are classified into four attributes; Product attributes, Hardware attributes, Personnel attributes and Project attributes [7].

Table I shows the intermediate driver of the COCOMO model. Each driver has its own multipliers (refers to Table II) that divided into six categories which are Very low (VL), Low (L), Neutral (N), High (H), Very High (VH), Extra High (XH) [7].

2) *Analogy-based model*: The core of the Analogy-based model is to differentiate the projects that will be estimated with all the software project's former data. Dataset will be from the company itself or that available publicly. The comparison will be carried out to identify which former projects are similar to the current project that its cost and effort will be estimated. Similar projects will be chosen to be

reworked so that the estimated effort of the new project can be identified. Similarity measures, how near the distance between project, will be measured on each type of attribute and using three measurement methods; Euclidean, Manhattan and Minkowski distance [5].

3) *Use-case points model*: Use Case Points (UCP) is a notable size estimate designed mainly for object-oriented projects that use the use case diagram to estimate the size of projects at the beginning development phase. Other software sizing methods that depend on functional requirements, called Function Point, was what encouraged the idea of UCP [13].

TABLE I. INTERMEDIATE COCOMO DRIVERS

Category	Drivers
Product Attributes	Required Software Reliability (RELY)
	Database Size (DATA)
	Product Complexity (CPLX)
Hardware Attributes	Execution Time Constraint (TIME)
	Main Storage Constraint (STOR)
	Virtual Machine Volatility (VIRT)
	Computer Turnaround Time (TURN)
Personnel Attributes	Analyst Capability (ACAP)
	Application Experience (AEXP)
	Programmer Capability (PCAP)
	Virtual Machine Experience (VEXP)
	Programming Language Experience (LEXP)
Project Attributes	Modern Programming Practices (MODP)
	Use of Software Tools (TOOLS)
	Required Development Schedule (SCED)

TABLE II. INTERMEDIATE COCOMO MULTIPLIERS

	VL	L	N	H	VH	XH
RELY	0.75	0.88	1.00	1.15	1.40	-
DATA	0.94	1.00	1.08	1.16	-1.23	-
CMPL	0.70	0.85	1.00	1.15	1.30	1.65
TIME	1.00	1.11	1.30	1.66	-1.30	1.66
STOR	1.00	1.06	1.21	1.56	-1.21	1.56
VIRT	0.87	1.00	1.15	1.30	-1.49	-
TURN	0.87	1.00	1.07	1.15	-1.32	-
ACAP	1.46	1.19	1.00	0.86	0.71	-
AEXP	1.29	1.13	1.00	0.91	0.82	-
PCAP	1.42	1.17	1.00	0.86	0.70	-
VEXP	1.21	1.10	1.00	0.90	1.34	-
LEXP	1.14	1.07	1.00	0.95	1.20	-
MODP	1.24	1.10	1.00	0.91	0.82	-
TOOL	1.24	1.10	1.00	0.91	0.83	-
SCED	1.23	1.08	1.00	1.04	1.10	-

C. Machine Learning

A computer becomes much more intelligent with their ability that can think by using Artificial Intelligence (AI). One of the subfields of AI is Machine Learning (ML). The computer intelligence is developed through various methods of learning. Thus, there are many types of Machine Learning which are Supervised Learning, Unsupervised Learning, Semi-Supervised Learning, Reinforcement, Evolutionary Learning and Deep Learning [14]. Machine Learning models are build based on learning the dataset using algorithms such as Regression Tree, Linear Regression, Neural Network, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Random Forest, etc. The training and testing will be carried out to the dataset to build the ML models.

Previously, ML has offered self-driving cars, speech recognition, systematic web explores, and improved realization of the human generation. Today machine learning is available everywhere that one can possibly use it many times a day without knowing it. A lot of researchers consider it as an excellent way of moving towards human-level as machine learning are advanced to the extent that it can recognize speech like human [15].

Machine learning is a subfield of computer science. It allows machines such as computers to build analytical models of data and find hidden perceptions by learning the data itself. It has been applied to a variety of aspects in modern society, ranging from Deoxyribonucleic Acid sequences classification, credit card fraud detection, robot locomotion, to natural language processing. It can be used to solve many types of tasks such as classification and prediction. Software project estimation is one of the tasks that machine learning is capable of [16].

Machine Learning is important as it is always up to date with the current environment and the model keeps improving its performance itself by learning data or experience. Human efforts and mistakes can be reduced by using Machine Learning.

The traditional software effort performance criterion is not accurate and not satisfying. There were many metrics and a number of techniques in cost estimation have been proposed. Unfortunately, most of them have lacked one or both of two characteristics which are sound conceptual, theoretical bases and statistically significant experimental validation.

Most performance criterion metrics have been defined by an individual and then tested in a very limited environment [17]. So, there is a need design optimization algorithm for correct, precise and reliable effort estimation [18].

Data mining is about acquiring perception in the data in order to detect useful patterns that imply information. Data mining has a record of success in business and more recently in scientific applications. Data mining is usually carried out using process models and employs tens of techniques that span a wide spectrum of interdisciplinary fields including statistics, machine learning, and pattern recognition. The use of data mining in software project prediction has recently gained remarkable popularity inspired by a large amount of error in traditional estimation methods and the continuous

improvements of machine learning algorithms which could help to provide more accurate prediction [7].

Machine Learning has been used to predict the software project cost and effort estimation since late 1980 [6]. Measuring the performance of estimation of machine learning models is accomplished by calculating the metrics including Sum Squared Errors (SSE), Root Mean Square Error (RMSE), Mean Magnitude of Relative Error (MMRE), Mean Absolute Error (MAE), etc.

They are the well-known parameters that are used for the performance evaluation of methods [19]. The evaluation consists of comparing the accuracy of the estimated effort with the actual effort. There are many evaluation criteria for software effort estimation and among them, the most frequent one is the Magnitude of Relative Error (MRE) [20]. Linear regression and Multi-perceptron are the most popular machine algorithms for software development effort estimation [21].

In this research, four Machine Learning algorithms that are used are Random Forest (RF), Linear Regression (LR), Regression Tree (RT) and Support Vector Machine (SVM). Each model's performance was measured by the Mean Magnitude Relatives Error (MMRE). Among the four algorithms, the best one is chosen to build a Machine Learning model that can predict the cost and time of a software project.

D. Literature Analysis

This section contributes to the knowledge of previous studies on software project estimation by using the state-of-art machine learning techniques available. Alongside with the limitation mentioned in the subsection C, most of previous literature studies addressed to measure the project and cost estimation using a single machine learning algorithm which reveals numerous limitations of the particular algorithm. Besides, most previous related work have inadequately present an extensive comparison and evaluation towards their proposed solution. Therefore, this paper aims to extend the evaluation with any other similar machine learning algorithm with four different datasets to further investigate the performance of the most sophisticated algorithm compared to COCOMO model.

III. METHODOLOGY

This research uses an experiment as a methodology to develop the prediction model for software project cost and effort estimation using selected machine learning algorithms. The experiment procedure is illustrated in Fig. 1. There are four selected machine learning algorithms which are Support Vector Machine, Linear Regression, Regression Tree and Random Forest.

A. Data Collection

In this experiment, software project measurement datasets use for developing the prediction model using machine learning. All datasets can be accessed publicly from <http://promise.site.uottawa.ca/serepository/datasets-page.html> and a study conducted by Kaushik et al, 2012. Sources of the datasets are from COCOMO NASA 1, COCOMO NASA 2, COCOMO81 and Kaushik et al. Details of the datasets used in this experiment are tabulated in Table III.

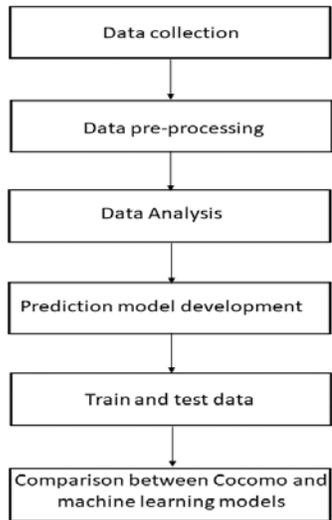


Fig. 1. Methodology.

TABLE III. DATASET USED IN THE EXPERIMENT

Dataset	Attributes	Number of projects
COCOMO NASA 1	17	60
COCOMO NASA 2	24	93
COCOMO81	15	63
Kaushik et al, 2012	16	15

B. Data Pre-processing

The data is pre-processed in order to calculate the effort estimation. In this experiment, the data is imported in r studio. Mice package is used to check the missing values, the datasets contain no missing values. The value of the drivers is in numerical weight converted to numerical values due to avoid bias during constructing the machine learning model. The mode constant is assigned based on the COCOMO predefined values.

C. Prediction Model Development

Four regression machine learning algorithms are used for this experiment.

1) *Linear regression model*: The linear regression model summarizes a relationship between two variables, independent and dependent variables. The practical use of linear regression in this experiment is to find the approximate prediction as a predictive model. The relationship of the prediction and the actuals data is then observed from the best fit line. The best fit line is where the total error prediction is as small as possible.

2) *Support vector machine*: Support vector machine model is a linear model for classification and regression problems. Linear and non-linear problems can be solved by support vector machine model. The aim of this model is to create a hyperplane and separate the data into classes. In support vector machine model, between the data points and the hyperplane we can find maximum margin to reduce misclassifications. Also, it can be used to solve unbalanced data problem.

3) *Regression tree algorithm*: Regression tree is a type of decision tree and it is a method that can create and visualize prediction models from the data. The output of this model is numeric output, and the average value is assigned to the leaves of tree. The decision making in regression tree is easier compared to other method because the undesired data will be filtered and reduces the work on the data as it goes deeper in the tree. The regression tree is used due its ability to reduce ambiguity in decision-making.

4) *Random forest algorithm*: Random Forest model is model made up of many decision trees that each tree depends random vectors values. This model called random because during building trees it uses random sampling for training data points and during splitting nodes it uses random subsets of features considered. Each tree in random forest learns from a random sample of the data points. The random forest is used due to it can produce high accurate classifier.

5) *Metrics*: Mean Squared Error (MSE), the mean squared error or mean squared deviation of an estimator measures the average of the squares of the errors that is, the average squared difference between the estimated values and what is estimated. The lower the value of MSE, the better accuracy.

$$MSE = \frac{1}{N} \sum |Effort_{estimated} - Effort_{actual}| \quad (1)$$

Root Means Squared Error (RMSE). It represents the sample standard deviation of the differences between predicted values and observed values (called residuals).

$$RMSE = \sqrt{\frac{\sum (Effort_{estimated} - Effort_{actual})^2}{N}} \quad (2)$$

Mean Absolute Error (MAE) is the average of the absolute difference between the predicted values and observed value. The MAE is a linear score which means that all the individual differences are weighted equally in the average.

$$MAE = \frac{1}{N} \sum \left| \frac{Effort_{estimated} - Effort_{actual}}{Effort_{actual}} \right| \quad (3)$$

Mean Absolute Percentage Error (MAPE) to measure prediction accuracy as percentage, or also known as the average absolute percent error for each predicted minus actual value and then divided by actuals values. The lower the lower of MAPE, the better the accuracy.

$$MAPE = \frac{100\%}{N} \sum \left| \frac{Effort_{estimated} - Effort_{actual}}{Effort_{actual}} \right| \quad (4)$$

The accuracy of the cost estimation models is evaluated by Magnitude of Relative Error (MRE) and the Mean Magnitude of Relative Error (MMRE). The optimum value of MRE and MMRE is closest to zero.

$$MRE = \left| \frac{Effort_{estimated} - Effort_{actual}}{Effort_{actual}} \right| \quad (5)$$

Mean Magnitude of Relative Error (MMRE) is the measure of predicted effort and actual effort value relative to the actual effort value.

$$MMRE = \frac{\sum_{i=1}^N \frac{|\text{Effort}_{\text{estimated}} - \text{Effort}_{\text{actual}}|}{\text{Effort}_{\text{actual}}}}{N} \quad (6)$$

Min-Max Accuracy is a good metric to see how much they are close that considers the average between the minimum and the maximum prediction. The higher the value of Min-max accuracy the better the accuracy.

Correlation Accuracy is the correlation between predicted and actuals used as an accuracy measure. The Pearson product-moment correlation coefficient is used to measure the strength of predicted and actuals value of the experiment. The predicted and actuals value has similar directional movements when the correlation accuracy is high.

P-Value also known as calculated probability is to determine the significance of the experiments results. The P-Value is lower than 0.05 shows strong prove against the null hypothesis, thus the null hypothesis is rejected. The smaller the P-Value, the stronger the evidence to reject the null hypothesis.

Null hypothesis of this project is, the population correlation coefficient is not significantly different from zero. There is no significant linear correlation between control and experimental values in the population.

Alternative hypothesis of this project is, the population correlation coefficient is significantly different from zero. There is a significant linear relationship between control and experimental values in the population [22].

Vargha and Delaney A (VDA) measure is one of the examples to measure effect size, differentiate between two samples of observations, control and experimental sample. The range of VDA is from 0 to 1. VDA is calculated there is no effect result in a value of 0.50 [23]. Interpretation of A measure:

- A value around 0.56 = small effect;
- A value around 0.64 = medium effect;
- A value around 0.71 = big effect;

Wilcoxon Rank Sum Test is nonparametric test is used to compare two related samples on a single sample to see if their population ranks differ. The null hypothesis is difference between the two samples has equal medians. The alternative hypothesis is there is no difference between the two samples. If the p-value is larger than 0.05, we must accept the null hypothesis because there is enough evidence to conclude. The null hypothesis is rejected, there is sufficient evidence to conclude the sample has no identical distributions [24].

D. Data Training and Testing

Training dataset are 80% and testing dataset are 20% for COCOMO81, COCOMO NASA 1. Training dataset are 70% and testing dataset are 30% for COCOMO NASA 2, and Kaushik et al.

IV. DATA ANALYSIS

In this project, correlation matrix uses to evaluate the correlation of the two variables. The dependent variable is actual effort attribute, while the 15 cost drivers and the line code are independent variable. From Fig. 2, 3, and 4, there are

five attributes that high positive correlation towards actual effort attribute, LOC, DATA, TIME, TOOL and STOR for COCOMO81, COCOMO NASA 1, COCOMO NASA 2. While, other attributes show negative correlations towards actual effort attribute.

Then, the project builds predictive machine learning models with COCOMO81, COCOMO NASA 1, COCOMO NASA 2 and Kaushik et al datasets, using all attributes. The predictive machine learning models are Support Vector Machine, Linear Regression, Regression Tree and Random forest. The result record is in the Table IV. The project evaluates the result and records in a table.

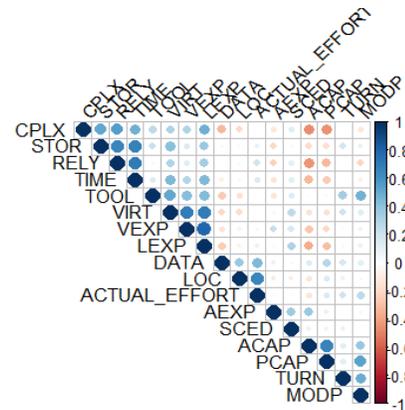


Fig. 2. Correlation of COCOMO81 Attributes.

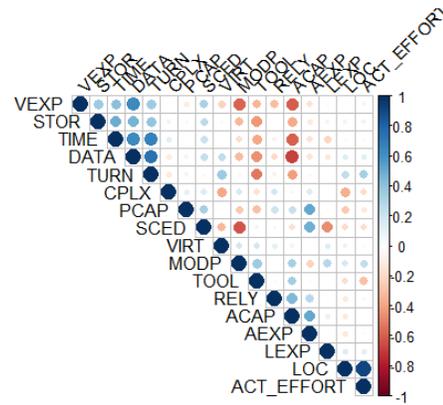


Fig. 3. Correlation of COCOMO NASA 1 Attributes.

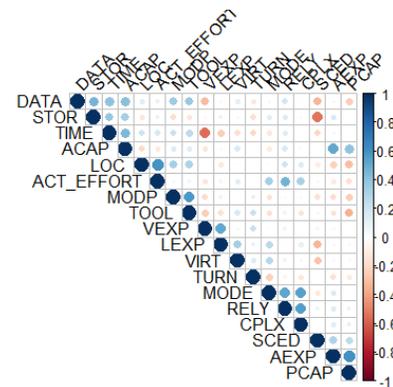


Fig. 4. Correlation of COCOMO NASA 2 Attributes.

TABLE IV. COMPARISON OF MACHINE LEARNING MODELS WITH COCOMO81 DATASET

Algorithm	Support Vector Machine	Linear Regression	Regression Tree	Random Forest
MAE	1002.844	1006.062	943.3075	928.3318
MSE	7675315	6939266	5193249	5769025
MAPE	4.721069	5.151246	4.803107	7.575327
RMSE	2770.436	2634.249	2278.87	2401.880
Min Max Accuracy	0.2742493	0.2647441	0.2758899	0.2895891
Correlation Accuracy	0.8218062	0.5933549	0.7427998	0.8671952
P-value	0.00568	0.03254	0.003628	0.001236
Significant Value	Significant	Not significant	Significant	Significant

Table IV explains the experiment is made on COCOMO81 with training dataset of 80 percent and the testing dataset is 20 percent. Based on the Table IV, Support Vector Machine, Regression Tree and Random Forest are significant due to the p-value which is less than 0.005 except Linear Regression model which p-value is 0.03254.

Table V explains the experiment that made on COCOMO NASA 1 with training dataset of 80 percent and the testing dataset is 20 percent. Based on the Table V, Support Vector Machine, Linear Regression, Random Forest are significant due to the p-value which is less than 0.005. The best MAE value among the four models is Support Vector Machine model which is the lowest, 31.5403163. The second-best MAE goes to Random Forest model with 36.8215429, followed up by Regression Tree with MAE 44.7018519. The worst MAE which has the highest value goes to Linear Regression with 47.7723733.

For the experiment that is made on COCOMO NASA 2 with training dataset of 80 percent and the testing dataset is 20 percent. Based on the Table VI, Support Vector Machine, Linear Regression, Random Forest and Regression Tree models are significant due to the p-value which is less than 0.005.

The experiment is made on Kaushik et al with training dataset of 70 percent and the testing dataset is 30 percent. Based on the Table VII, Support Vector Machine, Linear Regression, Random Forest are significant due to sufficient evidence to reject null hypothesis as the p-value which is less than 0.005; except Regression Tree model which p-value is 0.18 higher than 0.005, fail to reject null hypothesis.

From the evaluation above, Support Vector Machine and Random Forest show consistent and statistically significant between the two variables, predicted effort value and actual effort values; with the four datasets. On the other hand, Linear Regression and Regression Tree show inconsistent result and fail to reject null hypothesis. However, we cannot strongly prove that the statistically significant result of the research hypothesis is correct (100% certainty) and we need to calculate the effect size of the control vs experimental values.

TABLE V. COMPARISON OF MACHINE LEARNING MODELS WITH COCOMO NASA 1 DATASET

Algorithm	Support Vector Machine	Linear Regression	Regression Tree	Random Forest
MAE	31.5403	47.77237	44.701852	36.82154
MSE	2755.54	5078.1734	2719.832	2421.112
MAPE	0.290164	0.4632032	0.5614929	0.4032879
RMSE	52.49332	71.2613	52.15201	49.20480
Min Max Accuracy	0.8019366	0.5631169	0.645906	0.7328254
Correlation Accuracy	0.956159	0.95610	0.933626	0.9400819
P-value	3.67e-05	1.193e-06	9.069e-06	5.498e-06
Significant Value	Significant	Significant	Significant	Significant

TABLE VI. COMPARISON OF MACHINE LEARNING MODELS WITH COCOMO NASA 2 DATASET

Algorithm	Support Vector Machine	Linear Regression	Regression Tree	Random Forest
MAE	262.6466	412.9564	305.56240	281.0156
MSE	158576.63	318435.7	360594.68	187571.8
MAPE	2.101547	3.674328	1.61244	1.363585
RMSE	398.2165	564.3010	600.49536	433.0956
Min Max Accuracy	0.452203	0.332110	0.530223	0.5227043
Correlation Accuracy	0.712561	0.697992	0.568656	0.727761
P-value	0.0006161	0.008902	0.01108	0.004126
Significant Value	Significant	Significant	Significant	Significant

TABLE VII. COMPARISON OF MACHINE LEARNING MODELS WITH KAUSHIK ET.AL, 2012

Algorithm	Support Vector Machine	Linear Regression	Regression Tree	Random Forest
MAE	22.28604	14.77559	54.408	36.4
MSE	683.8073	468.77943	5247.853	2366.2582
MAPE	0.1721194	0.09580776	0.5615453	0.2335138
RMSE	26.14971	26.65131	72.4420	48.64420
Min Max Accuracy	0.82788	0.90449	0.64094	0.766486
Correlation Accuracy	0.9902	0.98828	0.7098	0.99224
P-value	3.67e-05	0.00152	0.18	0.008192
Significant Value	Significant	Significant	Not significant	Significant

In this research further experiment is to analysis the regression models with selected attributes COCOMO dataset using A measure and Wilcoxon Rank Sum test.

From Table VIII, the project calculates the accuracy metrics of COCOMO NASA 1, train and test Support Vector Machine and Random Forest with all attributes of COCOMO NASA 1, also with 5 selected attributes and one attribute only, LOC. From the result, COCOMO NASA 1, has the highest correlation accuracy compared to other experiments with machine models. There is sufficient evidence to reject the null hypothesis as the p-value of COCOMO NASA 1 is smaller than 5 percent, thus COCOMO NASA 1 is statistically significant. The MMRE of COCOMO NASA 1 is smaller compared to machine learning models' MMRE. The smaller the MMRE indicates the more accurate of the estimation [25].

As for Vargha and Delaney A measure there is no effect of difference between actual effort and the predict value of COCOMO NASA 1 model, to support the statement rank sum p-value is used to measure the distribution between the control and experimental sample. It shows that the p-value of Wilcoxon rank sum test is higher than 5 percent, the null hypothesis is fail to reject, the two samples has identical distributions.

To compare the trained machine learning models with all attributes and machines learning models only with 5 selected attributes, the five selected attributes performance better in producing more accurate results. The correlation accuracy of the five selected attributes have higher relationship between the actual effort and the predicted effort values compared to the all attributes. MMRE of five selected attributes has lower values compared to all attributes, this show that five selected attributes has more accurate estimations between the actual and predicted effort values. For Vargha and Delaney A measure, there are only slight differences between the all attributes and the five selected attributes, Support Vector Machine with all attributes has no effect differences compared to Random Forest model. The rank sum of all attributes and five selected attributes are statistically significant, there are enough evidence to support null hypothesis and to reject the alternative hypothesis.

Then experiment is using only one attribute, LOC. The correlation accuracy of machine models are increased compared to five attributes and all attributes. The p-value of Pearson's correlation also show the models are statistically

significant. The MMRE of Random Forest is lower than Support Vector Machine, 67.6706. The Vargha and Delaney A measure of Support Vector Machine and Random Forest has no effect in difference, this mean the distribution of actual and predicted effort values are identical. The Wilcoxon rank sum test of Support Vector Machine and Random Forest are statistically significant where, there are enough evidences to support null hypothesis and reject the alternative hypothesis since the p-value of both machine learning models is higher than 5 percent.

To conclude, the machine learning models learn and prove that not all attributes are needed to trained and needed. From this experiment, using one attribute, LOC can have closer MMRE towards the COCOMO prediction model, higher correlation accuracy and identical distribution of actual and predicted effort values.

From the Table IX, the project calculates the accuracy metrics of COCOMO NASA 2, train and test Support Vector Machine and Random Forest with all attributes of COCOMO NASA 2, also with 5 selected attributes and one attribute only, LOC. COCOMO NASA 2 has 93 projects and highest number projects compared to other COCOMO datasets. From the result, COCOMO NASA 2 has the lower correlation accuracy compared to experiments with machine learning models that used all attributes. There is no sufficient evidence to reject the null hypothesis as the p-value of COCOMO NASA 1 is larger than 5 percent, thus COCOMO NASA 2 is not statistically significant. The MMRE of COCOMO NASA 2 is smaller compared to machine learning models' MMRE.

The smaller the MMRE indicates the more accurate of the estimation (Malhotra, 2014). As for Vargha and Delaney A measure there is intermediate differences between actual effort and the predict value of COCOMO NASA 2 model, however the Wilcoxon rank sum p-value is used to measure the distribution between the control and experimental sample. It shows that the p-value of Wilcoxon rank sum test is higher than 5 percent, the null hypothesis is fail to reject, the two samples has identical distributions. The COCOMO NASA 2 is still statistically significant according to rank sum test.

TABLE VIII. COMPARISON OF COCOMO NASA 1 DATASET WITH DIFFERENT ATTRIBUTES

Algorithm	COCOMO NASA 1	All Attributes		5 Attributes		LOC Only	
		Support Vector Machine	Random Forest	Support Vector Machine	Random Forest	Support Vector Machine	Random Forest
Correlation Accuracy	0.97578	0.8157	0.6623	0.7636	0.8642	0.7772	0.85521
P-value	6.295e-08	0.00122	0.01895	0.003842	0.002882	0.002933	0.000391
Significant Value	Significant	Significant	Significant	Significant	Significant	Significant	Significant
MMRE	29.61	127.5024	110.7032	109.684	108.4127	66.98639	67.6706
A Measure	0.5 (No effect)	0.4514 (No effect)	0.4097 (Small)	0.42361 (Small)	0.42361 (Small)	0.45833 (No effect)	0.4861 (No effect)
Rank Sum	1	0.7074	0.4704	0.5443	0.5443	0.7508	0.931
Significant Value of Rank Sum	Significant	Significant	Significant	Significant	Significant	Significant	Significant

TABLE IX. COMPARISON OF COCOMO NASA 2 WITH DIFFERENT ATTRIBUTES

Algorithm	COCOMO NASA 1	All Attributes		5 Attributes		LOC Only	
		Support Vector Machine	Random Forest	Support Vector Machine	Random Forest	Support Vector Machine	Random Forest
Correlation Accuracy	0.4388394	0.8142797	0.7077862	0.6552592	0.7265169	0.4203878	0.4567935
P-value	0.06016	2.202e-05	0.0006982	0.002324	0.004269	0.07311	0.04929
Significant Value	Not Significant	Significant	Significant	Significant	Significant	Not Significant	Significant
MMRE	150	281.5356	201.6638	226.7969	334.0268	205.99	133.2845
A Measure	0.1952663 (Large)	0.318559 (Medium)	0.3490305 (Small)	0.3268698 (Medium)	0.2908587 (Medium)	0.3434903 (Small)	0.3822715 (Small)
Rank Sum	0.1611	0.05771	0.1149	0.07025	0.02854	0.102	0.22
Significant Value of Rank Sum	Significant	Significant	Significant	Significant	Not Significant	Significant	Significant

In the experiment of machine learning models are using all attributes, there are large differences result obtained from Support Vector Machine models and Random Forest model. The correlation accuracy of the Random Forest is lower than Support Vector Machine however Random Forest has higher MMRE value compared to Support Vector Machine. Moreover, Vargha and Delaney A measure, Support Vector Machine and large difference between the predicted and actual effort values, while Random Forest has closer accuracy of control and experimental sample. The Wilcoxon rank sum test for Support Vector Machine and Random Forest model are statistically significant where there are enough evidence to support null hypothesis and reject the alternative hypothesis, as the p-value of both machine learning is higher than 5 percent.

Then, the experiments are evaluated between all attributes and five selected attributes. The correlation accuracy of the five selected attributes have higher relationship between the actual effort and the predicted effort values compared to the all attributes. However the MMRE of Random Forest for the five selected attributes has higher value compared to Random Forest for the all attributes, while for Support Vector Machine is vice versa.

For Vargha and Delaney A measure, the selected five attributes for both machine learning show medium differences between the actual and predicted effort value. The rank sum of all attributes and five selected attributes are statistically significant, there are enough evidence to support null hypothesis and to reject the alternative hypothesis.

Then experiment is using only one attribute, LOC. The correlation accuracy of machine models drop compared to five attributes and all attributes. The p-value of Pearson's correlation only show the Random Forest model is statistically significant. The MMRE of Random Forest is lower than Support Vector Machine compared to all attributes and COCOMO NASA 2 model. The Vargha and Delaney A measure of Support Vector Machine and Random Forest have small in difference, this mean the distribution of actual and predicted effort values are closely identical.

The Wilcoxon rank sum test of Support Vector Machine and Random Forest are statistically significant where, there are

enough evidences to support null hypothesis and reject the alternative hypothesis since the p-value of both machine learning models is higher than 5 percent.

To conclude, the machine learning models learn and prove that not all attributes are needed to trained and needed. From this experiment, using one attribute, LOC can have closer MMRE towards the COCOMO prediction model, higher correlation accuracy and identical distribution of actual and predicted effort values.

From the Fig. 5, the machine learning model support vector machine and random forest show similar pattern towards to the actual effort, while the calculation of COCOMO deviates at 100, 14, 302, 113, 350 and 339 lines of codes. Support Vector Machine and Random Forest has proximity predicted effort values compared to effort prediction of COCOMO models. Refer to appendices for development of machine learning model.

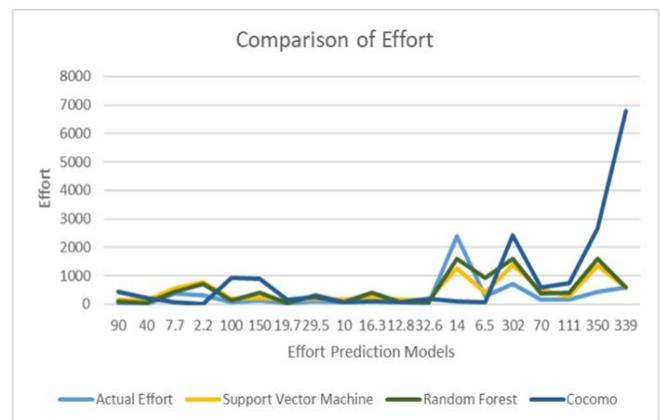


Fig. 5. Comparison of Actual Effort and Estimated Effort.

V. DISCUSSION

The results of the experiments found clear support that Support Vector Machine and Random Forest algorithms impressively give consistent results with the COCOMO datasets regardless on the number of effort attribute used.

However, the planned comparison in this paper reveals the good performance and increase in the accuracy of Support Vector Machine for estimation of software project effort. Support Vector Machine also able to delivers significantly better results with five important attributes compared to all attributes used to estimate project effort and cost, as some of the attributes are irrelevant to estimate.

VI. CONCLUSION AND FUTURE WORK

To conclude, many existing machine learning algorithms can train predictive models, however, the right and suitable machine learning model are needed to give an accurate estimation. In this research, the five selected attributes with high positive correlation toward actual effort attribute are obtained from the correlation matrix, DATA, STOR, LOC, TIME and TOOL. The five important attributes give better result compared from using all the attributes in COCOMO dataset. Hence, not all attributes in the dataset are relevant to be used to measure the project estimation. Further improvement, during training and testing data, the data is advised to split three parts, 70 percent of training data, 20 percent of testing data, and 10 percent of validation data.

Further improvement of this research on machine learning models is to perform ensemble stacking also known as blending, to ensemble the four machine learning models in order to optimize the predictive model. In development of machine learning model, percentage of accuracy should be included to show the difference percentage between the predicted value and the actual effort value.

ACKNOWLEDGMENT

This research is funded by IIUM Research Acculturation Grant Scheme with the Ref: IRAGS18-012-0013.

REFERENCES

- [1] M. Jorgensen, "What We Do and Dont Know about Software Development Effort Estimation," *IEEE Softw.*, vol. 31, no. 2, pp. 37–40, 2014.
- [2] J. Kaur, T., & Singh, "A Review on Cost Estimation Models for Effort Estimation," *Int. J. Sci. Eng. Res.*, vol. 6, no. 5, pp. 179–183, 2015.
- [3] M. Aljohani and R. Qureshi, "Comparative Study of Software Estimation Techniques," *Int. J. Softw. Eng. Appl.*, vol. 8, no. 6, pp. 39–53, 2017.
- [4] U. Shekhar, S., & Kumar, "Review of Various Software Cost Estimation Techniques," *Int. J. Comput. Appl.*, vol. 141, no. 11, pp. 31–34, 2016.
- [5] S. Ardiansyah, A., Mardhia, M. M., & Handayaningsih, "Analogy-based model for software project effort estimation," *Int. J. Adv. Intell. Informatics*, vol. 4, no. 3, pp. 251–260, 2018.
- [6] G.-H. Cho, H.-G., Kim, K.-G., Kim, J.-Y., & Kim, "A Comparison of Construction Cost Estimation Using Multiple Regression Analysis and Neural Network in Elementary School Project," *J. Korea Inst. Build. Constr.*, vol. 13, no. 1, pp. 66–74, 2013.
- [7] A. Banimustafa, "Predicting Software Effort Estimation Using Machine Learning Techniques," *8th Int. Conf. Comput. Sci. Inf. Technol. CSIT 2018*, (October), pp. 249–256, 2018.
- [8] R. Kalaivani, N., & Beena, "Overview of Software Defect Prediction Using Machine Learning Algorithms," *Int. J. Pure Appl. Math.*, vol. 118, pp. 3863–3873, 2018.
- [9] J. Thomas, "Blown Budgets and Destroyed Schedules. Sometimes, It's Weak Project Estimation That's to Blame," *Sci. Uncertain.*, pp. 56–61, 2019.
- [10] N. Ghatasheh, H. Faris, I. Aljarah, and R. M. H. Al-Sayyed, "Optimizing Software Effort Estimation Models Using Firefly Algorithm," *J. Softw. Eng. Appl.*, vol. 8, no. 3, pp. 133–142, 2015.
- [11] M. Z. Alsaeedi, A., & Khan, "Software Defect Prediction Using Supervised Machine Learning and Ensemble Techniques: A Comparative Study," *J. Softw. Eng. Appl.*, vol. 12, no. 5, pp. 85–100, 2019.
- [12] B. W. Boehm et al., *Software Cost Estimation with COCOMO II*. Upper Saddle River, NJ: Prentice Hall, 2000.
- [13] A. B. Azzeh, M., & Nassif, "A hybrid model for estimating software project effort from Use Case Points," *Appl. Soft Comput.*, vol. 49, pp. 981–989, 2016.
- [14] M. Fatima, M., & Pasha, "Survey of Machine Learning Algorithms for Disease Diagnostic," *J. Intell. Learn. Syst. Appl.*, vol. 9, no. 1, pp. 1–16, 2017.
- [15] N. Rambhajani, M., Deepanker, W., & Pathak, "A Survey on Implementation of Machine Learning Techniques for Dermatology Diseases Classification," *Int. J. Adv. Eng. Technol.*, vol. 8, pp. 194–195, 2015.
- [16] N. Wang, "Bankruptcy Prediction Using Machine Learning," *J. Math. Financ.*, vol. 7, no. 4, p. 908–918.
- [17] P. K. Tripathi, R., & Rai, "Machine Learning Methods of Effort Estimation and It's Performance Evaluation Criteria," *Int. J. Comput. Sci. Mob. Comput.*, vol. 6, no. 1, pp. 61–67, 2017.
- [18] C. Wadhwa, A., Jain, S., & Gupta, "An Effective Precision Enhancement Approach to Estimate Software Development Cost: Nature Inspired Way," *J. Telecommun. Electron. Comput. Eng.*, vol. 9, no. 3, 2017.
- [19] S. Malathi, S., & Sridhar, "Analysis of size matrices and effort performance criterion in software cost estimation," *Indian J. Comput. Sci. Eng.*, vol. 3, no. 1, 2012.
- [20] S. Kaushik, A., Chauhan, A., Mittal, D., & Gupta, "COCOMO Estimates Using Neural Networks," *Int. J. Intell. Syst. Appl.*, vol. 4, no. 9, pp. 22–28, 2012.
- [21] V. K. Bhatia, S., & Attri, "Machine Learning Techniques in Software Effort Estimation Using COCOMO Dataset," vol. 2, no. 6, pp. 101–106, 2015.
- [22] B. Illowsky, & S. Dean, *Testing the Significance of The Correlation Coefficient*. Introductory Statistics, Rice University, Houston, USA, 2013.
- [23] A. R. Ismail, "Immune-Inspired Self-Healing Swarm Robotic Systems," *York*, 2011.
- [24] "Texas Gateway," 2019. [Online] <https://www.texasgateway.org/resource/95-additional-information-and-full-hypothesis-test-examples>
- [25] R. Malhotra, "A systematic review of machine learning techniques for software fault prediction," 2014. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S1568494614005857>.

Do Learning Styles Enhance the Academic Performance of University Students? A Case Study

Jorge Muñoz-Mederos¹, Elizabeth Acosta-Gonzaga², Elena Fabiola Ruiz-Ledesma³, Aldo Ramírez-Arellano⁴
Instituto Politécnico Nacional, UPIICSA, CDMX, México^{1,2,4}
Instituto Politécnico Nacional, ESCOM, CDMX, México³

Abstract—Higher education models appear to be not entirely designed to support students in facing severe challenges, such as failure in exams and dropping out of school. To solve these challenges, several models of learning styles have been proposed, under the premise that these studies contribute to improving the student's learning experience. This research aims at quantifying the impact of learning styles (learning preferences/dimensions) on students' academic performance from a higher education institution. Ninety-six undergraduate students were surveyed during the 2018-2019 school year and randomly divided into two groups: control (CG) and experimental (EG). The learning preferences of the students were identified using the Unified Learning Style Model (ULSM) instrument. Subsequently, the level of students' knowledge concerning the course was determined employing a pre-test exam. As a following step, the students of the EG consulted the learning objects designed considering different learning styles. The CG attend their lessons in a face-to-face environment; both groups answered a post-test exam to assess their learning. The learning styles' effect on learning objects were designed to cover several learning styles- on academic performance is quantified employing an ANOVA analysis. The results differ from those postulated in previous researches based on the ULSM since there is no statistical evidence that learning styles influence students' academic performance. Therefore, it is necessary to explore other cognitive and affective factors that make the student's learning experience efficient and effective.

Keywords—Learning styles; learning objects; unified learning style model (ULSM); academic performance; higher education

I. INTRODUCTION

According to the National Association of Universities and Institutions of Higher Education of Mexico, in the 2016-2017 school year, there were 2,655,711 students enrolled in public schools [1]. Of these students, the National Polytechnic Institute (IPN) in 2019 [2], in any of its 56 engineering and bachelor's degrees, met the demands of 108,296 students. Of these 66,139 were male students and 42,157 were female students. According to the IPN strategic management program, 13,751 students were enrolled during the 2016-2017 school year in the UPIICSA public institution [3].

However, higher education faces great challenges, and one of these problems is related to school dropout due to various factors, such as: the high failure rate that is mainly observed in subjects considered "difficult", that demand a high level of knowledge, and a lot of dedication on the part of the student. However, the educational models in public institutions do not seem to be designed entirely to help students meet this challenge.

One of the strategies to solve this challenge has been the proposal of various models of learning styles, which classify the way in which a student learns and with it, develop educational materials that serve to facilitate the learning process.

This leads the researchers to think that if students knew how they can learn better, they would notice a significant change in their learning. Moreover, most learning style models capture characteristics in the face-to-face learning context, which made it difficult to identify the learning styles of online learning students. For this reason, The Unified Learning Style Model (ULSM) [4] was proposed to unify the most relevant learning style models, and consequently, to be able to facilitate students learning in the era of web 4.0.

The identification of the way in which students learn can influence their school performance, however, this has been a matter of controversy due to the fact that there have been several studies in favor of the hypothesis that the academic performance of students is given through educational materials designed to suit learning styles, and other studies that contradict this idea.

Due to the aforementioned, this main objective of this research is to determine if the learning objects, which were designed considering the student's learning preferences, are able to impact on their academic performance compared to those students who reviewed the same topics of the learning objects, but in a face-to-face class. For this reason, the students learning style has been classified by means of the ULSM instrument and the Learning Objects (LOs) has been adapted to the detected style, then its influence on academic performance was quantified.

The design and development of personalized LOs that facilitate the retention and understanding of knowledge is considered a fundamental part of solving this problem [5].

On the other hand, there are many criticisms regarding the use of learning styles or preferences due to the fact that some learning styles or preferences are not fully compatible with the use of e-learning systems or based on teaching through the web [4]. In the same way, it warns about the risk that exists when the student is mistakenly pigeonholed in some style [6]. Additionally, most learning style models place students into mutually exclusive groups, for example, a student cannot be verbal and visual at the same time [7]. This leads to ask the question whether or not this nominal classification of having a learning preference is correct or should be considered as a gradual and not dichotomous scale.

II. RELATED WORKS

Reference [8] studied the platforms that use learning styles as personalization criteria, which are also known as adaptive educational systems based on learning styles (Web-based Educational system with Learning Style Adaptation (WELSA)). This study involved 64 graduate students, who were surveyed using the ULSM instrument. Each student was characterized, and adaptation rules were created according to the preferences detected by the ULSM. The results showed that the students whose learning preferences were considered to recommend the courses, achieved a favorable result in their learning, while the students who did not use the recommended course, had an adverse effect.

Similarly, the research by [9] identified the learning styles of 136 engineering students, under the premise that not all students are interested in online learning, in particular, those students belonging to the engineering areas, and that this phenomenon is related to the learning style and their perception of online learning. For this effect, the questionnaire, inventory of learning styles by [10] and the questionnaire, student perception towards online learning adapted from [11] were used. The results revealed that learning styles do not influence engineering students' perception of online learning.

The study by [12] analyzed the effects that arise when equating the teacher's learning style with the student's learning style, and thus, seek a better academic performance of the students. For this purpose, the authors used the learning styles scale of [13], which was applied to instructors and students. The results contradict what was expected; it was already desired that by matching the teacher's learning style and the student's learning style, there would be favorable effects for the students [12].

There have been studies whose objective has been showed to identify the dominant learning style of students, for example, the study directed by [14] which conducted an experiment in a high school applying a survey to 584 students in order to determine the dominant learning style based on the inventory of learning styles of [10]. The results showed that 50.5% of the students were classified as assimilators, which means that the students fluctuate between reflective observation and abstract conceptualization. Additionally, the study concludes that there is no significant difference in terms of learning style and gender, which supports the idea that the learning styles of both men and women have similar preferences.

Likewise, [5] identified the learning preferences of 58 first-year university students using the ULSM instrument. The authors also carried out evaluations on the variables, satisfaction with learning, motivation, time invested in the study, and effort while studying, and identifying the students' preferences in six key aspects: perception, processing, dependence or independence of the field, reasoning, organization and social preferences.

The results showed that after a semester in which students were made aware of their learning preferences, no significant changes were found in the variables evaluated, learning

satisfaction, motivation, time spent studying or effort invested in the study. Furthermore, it is suggested to make changes in study techniques and in the way, students incorporate their learning preferences into their study habits. Finally, the authors concluded that learning preferences are a state and not a trait in students, that is, a trait is associated with stable characteristics, while states are temporary feelings or behaviors of the student at a given time.

Similarly, the study by [15] developed a management system aimed at merging LO, and the consequent adaptation to the student's learning preference. In this study 56 university students participated who were divided into two groups: control and experimental. The ULSM instrument was used to classify students according to their learning style. The proposed system searched the didactic contents for a repository of LOs and then merged LOs according to the learning preferences detected. Subsequently, the LOs fused and adapted to the students of the experimental group were shown.

The authors showed that there are favorable results for the experimental group, when by presenting the merged LOs to the students, it was possible to make the students focus on the subject of study, and avoid the distractions underlying the location of other materials that may or may not be related to their learning preference or of the study subject. A more detailed study by the same authors showed that the fusion of LOs was responsible for improving learning performance that may be related to cognitive load [16].

III. THEORETICAL FRAMEWORK

Learning is a process that has had an innate complexity, without showing a direct and unique solution since there are many factors that affect it, and that cannot be easily controlled [17]. Therefore, multiple ways of studying learning have been proposed, and thus, being able to propose alternatives to make more efficient results. One of the proposed alternatives is to determine the style in which a student learns.

A learning style is defined as "characteristic cognitive, affective, and psychosocial behaviors that serve as relatively stable indicators of how students perceive, interact with, and respond to the learning environment" [18]. Consequently, it is a set of ways by which a person learns as a result of various intrinsic and extrinsic factors, among which the affective and cognitive factors stand out, which may at some point influence the student's performance.

According to literature, 71 different models of learning styles can be identified, however, thirteen models have been recognized as the best results [6] and these are: Allinson's and Hayes Cognitive Style Index (CSI) by [19], Myers-Briggs Type Indicator (MBTI) by [20], Learning Style Inventory (LSI) by [21], Apter's Motivational Style Profile (MSP) by [22], Dunn and Dunn's Model by [23], Entwistle's Approaches to Studying Inventory (ASI) [24], Gregorc's Mind Styles Model and Style Delineator (GSD) [25], Herrmann Brain Dominance Instrument (HBDI) by [26], Learning Styles Questionnaire (LSQ) by [27], Learning Styles Profiler (LSP) by [28], among others.

The huge number of models indicates that being able to identify students' learning styles is not a simple task, because each model considers some particular characteristic of their learning, leaving aside other important features. This led educational researchers to propose an instrument that could unify the cited models to identify the learning styles for each student more accurately.

A. Unified Learning Styles Model (ULSM)

The unified model of learning styles synthesizes the main characteristics of the models mentioned in literature, known as integrative taxonomy [8]. This model integrates learning preferences related to the modality of perception, form of information processing and organization, as well as social and motivational aspects.

Some of the ULSM preferences have a direct corresponding with a dimension of a learning style model, while others represent only a trait that determines a certain style [8].

One of the corresponding dimensions is field dependence or independence, which was based on the learning styles model of [29], including the name and the definition. The opposite case can be observed in the preference regarding attention to details, since it does not have direct correspondence with any of the dimensions of the existing learning styles models [8].

The ULSM includes the following dimensions and preferences: Perception modality and its preferences: visual vs verbal; Information processing and its preferences: abstract vs concrete, serial vs holistic; Experimentation and its preferences: active vs reflective observation, careful vs not careful with details; Field and its preferences: dependency / independence; Reasoning and its preferences: deductive vs inductive; Organization of the information and its preferences: synthesis vs analysis; Motivation and its preferences: intrinsic vs extrinsic, deep vs superficial vs strategic vs resistant approach; Persistence and its preferences: high vs low; Pace and its preferences: focus on one task at a time vs alternating tasks and topics; Social aspects and its preferences: individual work vs team work, introversion vs extroversion, competitive vs collaborative; Coordinating body and its preferences: affectivity vs thought.

The preferences of the perception modality dimension are included from several models of traditional learning styles such as: the Felder-Silverman Learning Style Model (FSLSM) (visual vs verbal dimension) [30], VARK (Visual, Aural, Reading/writing, Kinesthetic) [31], VAK (Visual, Auditory, Kinesthetic), of the Dunn and Dunn model [23] (visual, auditory, kinesthetic, tactile) and of the Riding model (verbalizer / image generator) [32]. Moreover, only visual vs. verbal preference was considered, because kinesthetic or tactile preference is difficult to perceive in an online environment [8]. One of the characteristics that was preserved to identify this dimension was the definition of the Felder-Silverman Learning Style model (FSLSM) which says: "visual students remember better what they see (images, diagrams, graphs, etc.), on the other hand, verbal students rely more on either spoken or written words" [8].

The information processing dimension includes preferences: concepts and generalizations, abstract vs. concrete, practical examples. These preferences were based on Kolb's learning cycle (abstract conceptualization vs concrete experimentation) [21] and Gregorc's model (abstract vs concrete) [25].

The serial vs. holistic preference was inspired by the FSLSM (sequential or global) [30] and by Pask's model (serial or holistic) [33].

From Kolb's learning cycle, the preference of active experimentation vs. reflective observation was taken, which is also described in the FSLSM (active or reflective) [30] and in the Honey and Mumford model (activist or reflector) [27].

According to [8], students who have abstract preference tend to trust contextual interpretation, while students who demonstrate concrete preference rely on immediate experience to capture learning. Similarly, students with a sequential preference tend to understand knowledge linearly, while students who have a global learning preference tend to learn in disorder. Some consider that they perform their learning in large leaps, and that they have the ability to make fast connections between different topics.

The field-dependent vs. field-independent dimension has been based on the model of [29]. This dimension refers to the student's environment and how the environment affects or does not affect their interpretation and the ability to locate information. The meaning is that students who depend on the field have difficulty locating the precise information they need. If this information is hidden or if there is other information superimposed that prevents it from being easily accessible, so they are more oriented towards people. On the other hand, students classified as independent of the field find it easier to recognize and select what is important in their environment, with a kind of abstraction and additionally have an impersonal orientation [8].

The reasoning dimension and its inductive vs. deductive preferences were taken from the first version of the FSLSM. Inductive students prefer to reason from particular facts or situations, and thus, be able to reach a general conclusion. That is, they go from the specific to the general. They also respond better to problem-based learning (fictional or real), as well as inquiry learning. Deductive students prefer to reason from the general to the specific, which is why they prefer the course to begin with the theoretical foundations and with the basic principles and then continue with the applications corresponding to the course [8].

The information organization dimension and its preferences, synthesis vs analysis, had no basis in any previous learning style model. Similar concepts can be found in the Allinson and Hayes model (intuitive or analytical) [19] and in the Riding model (holistic or analytical). A student with a preference for synthesis is the one who has a general image of the topic and who tends to combine different elements to understand something completely. Furthermore, a student with a preference for analysis focuses on each of the parts of a whole, as well as on the basic principles, and therefore be able to build what remains [8].

The motivation dimension and the deep vs. strategic vs. superficial vs. reticent preferences was based on the Entwistle model [24]. From the Grasha-Riechmann model the reticent preference was taken which is similar to the Grasha-Riechmann avoidance concept [34].

The intrinsic vs extrinsic preference does not have a direct correspondence with a learning style model, but it is related to the Entwistle model, as well as the telic (serious) and paratelic (playful) dimensions of [22].

According to [8], the preferences of the motivation dimension have the following characteristics: students with the preference of profound are oriented towards meaning. They try to understand ideas by themselves or by their own means. They generally show interest in the full content of the course. Students with a preference for strategic focus, tend to be achievement-oriented; generally, in a group of students they are the ones who want to obtain the highest grades and they are constantly attentive to the requirements and evaluation criteria. Likewise, they try to guide the work to perceived preferences in teachers.

Students with a superficial preference are oriented towards reproduction, that is, their intention is to pass the different evaluations that are presented to them, by memorizing facts, data and whatever else they need. They generally do not find meaning in the new ideas that are presented to them; they tend not to reflect on the purpose or strategy of a certain action, and they feel excessive pressure as they worry about the work.

Students with the resistant preference show a total disinterest in the course and generally refuse to participate in the learning activities suggested by the teacher. Most of the time they are apathetic students; they do not pay attention to their classes and they are disobedient. Students who are intrinsically motivated learn by the mere fact that their attention will generate a new experience in their life, while those who are extrinsically motivated learn to obtain an external reward which serves as motivation itself [8].

For the persistence dimension, its preferences were based on the Dunn and Dunn model [23] (persistent or non-persistent). In this dimension, students who demonstrate a high persistence preference are those students who are inclined to complete all the tasks that are entrusted to them, sometimes regardless of the cost. They tend to spend a lot of time studying and sometimes, if necessary, they review the study material over and over again. Students who have a low persistence preference tend to take consecutive breaks and rarely return to study material [8].

The pacing dimension includes two preferences, concentrating on one task at a time or alternating between multiple tasks and topics. This dimension was not based on any previous learning style model. Students who demonstrate a preference to concentrate on one task at a time are considered as having linear learning. This means that they will not continue with the learning if they do not complete the current task. Sometimes, they have few changes, or they have no changes and jumps of activity. Students who have the preference to alternate subjects or topics are considered to a certain extent inconstant because they jump from topic to

topic and from subject to subject and this includes jumps from one course to another [8].

The social aspects dimension and their preference to learn alone vs. learning among peers, has been based on the Dunn and Dunn model, and has also been related to other learning styles such as the active or reflective of the FSLSM.

Introversion vs. extraversion preference was based on the Myers-Briggs classification (MBTI). From the Grasha and Riechman-Hruska model [13], the competitive vs. collaborative preferences were taken, which is also related to the autistic domain preference of [22].

Students who have a preference for introversion are those students who tend to avoid social contact and are constantly preoccupied with their thoughts and inner feelings. Extroverted students are constantly involved with social and practical realities of that same nature, instead of worrying about their thoughts and feelings [8]. Students who demonstrate a competitive preference tend to participate in any activity as long as they have competition with someone else. While students who have a preference for collaboration, feel better helping a common goal for all, constantly considering the "win-win" [8].

The coordinating instance dimension of the learning process and its affectivity vs. thought preferences is related to the feeling vs. thought preferences of the Myers-Briggs model (MBTI). Students who demonstrate a preference for affectivity tend to complete tasks based on intuition and their way of feeling, while those students who have a preference for learning thinking, make decisions based on analysis, logic, and reasoning [8].

As mentioned above, the ULSM has only integrated learning style preferences related to the web-based educational context, rather than the context for face-to-face learning. Accordingly, only those preferences related to the context of education 4.0 have been included.

Educational environments for education 4.0 have been enriched by the incorporation of Learning Objects (LOs), which are defined as a set of educational experiences obtained through reflection and experimentation in practice aimed at solving real educational problems, and that are represented in Multimedia Teaching Materials (MTM), in order to avoid memorization of information [35]. LOs are all those digital or non-digital entities that can be reused or referenced during learning with technological support [36]. For this research LOs have been considered as the set of reusable or not reusable multimedia educational materials, and that are intended to facilitate learning for students.

Currently, LOs are used in interactive learning environments, distance learning systems, and collaborative learning environments; and are required by multiple companies to offer online training, such as Oracle company that offers its training courses virtually through what is known as Oracle University, which uses learning objects (texts, videos, illustrations, etc.) to facilitate the learning of the personnel concerned.

IV. METHOD

The study was carried out in a public higher education school with 96 students who had enrolled on the web technologies course of the Industrial Administration career of 2019 scholar year. The age of the students ranged from 20 to 23 years old. It was explained to teachers and students of six groups the objective of the investigation. They both agreed to participate. In each class, the students' total number was randomly divided into two groups: the Control Group (CG) and the Experimental Group (EG).

LOs were designed and developed, including the following dimensions of the ULSM and its respective preferences (learning styles), perception, information processing, field independence, reasoning and organization. The LOs were upload to Moodle platform to be available for the students.

The ULSM instrument was employed to gather the students' preferences. It has been validated through the electronic learning platform, WELSA [8]. Each student's prior knowledge about the course was determined by an evaluation (pre-test) related to the course topic. No statistical difference was found; the details are presented next.

The GC students took all their classes in person, attending their classroom with their teacher. The GE students attended the computer room to take all their course classes using the Moodle platform.

Then, a post-evaluation (post-test) was administrated to the students. This evaluation was carried out to determine the knowledge acquired from the face-to-face or online lessons. The GC students carried out their evaluation on paper, and the EG students through the platform. The questions on both tests were the same. The normality and homoscedasticity of both pre-test and post-test were evaluated. Then, two-way ANOVA (group x dimension) were performed to determine the effect of the dimensions of the ULMS on the students' performance in both groups (CG, EG).

V. RESULTS

The dominant learning style of each student was determined through the ULSM instrument. The results are shown in Table 1.

TABLE I. STUDENTS' LEARNING STYLE ACCORDING TO ULMS DIMENSIONS

Perception dimension	
Kinesthetic	65%
Aural	21%
Visual	6%
Reading/writing	8%
Information processing dimension	
Concrete	86%
Abstract	14%
Holistic	64%
Sequential	36%
Active	50%
Reflective	50%

Field dimension	
Dependent	33%
Independence	67%
Reasoning dimension	
Deductive	26%
Inductive	74%
Information dimension	
Synthesis	66%
Analyze	32%
Motivation dimension	
Deep	44%
Superficial	31%
Strategic	21%
Disinterest	4%
Persistence dimension	
High persistence	51%
Low persistence	49%
Pacing dimension	
Completed one task	70%
Jumped between tasks	30%
Social dimension	
Work in a team	25%
To do it individually	75%
Collaborative	84%
Competitive	16%
Extrovert	48%
Introvert	52%

A. Influence of Learning Styles on Students Performance

The Shapiro-Wilk test showed that the grade obtained in the pre-test of the CG y GE has normal distribution, see Table 2.

TABLE II. THE SHAPIRO-WILK TEST FOR PRE-TEST EVALUATION

Pre-test			
	Statistic	df	Sig
CG	0.972	43	0.376
EC	0.975	53	0.326

($p=0.05$)

Levene's test showed that the grade of the pre-test is homoscedastic, $F(94) = 0.576, p = 0.450$.

The student's t-test showed no difference between the grade of pre-test of EG students and CG students, see Table 3.

Thus, the students in both groups have the same level of prior knowledge.

TABLE III. STUDENT T-TEST OF THE GRADE FOR THE PRE-TEST EVALUATION

Pre-test					
	M	DS	t-value	df	p-value
CG	6.638	1.1479	-1.18	94	0.241
EC	6.337	1.3553			

The post-test evaluation showed that the data are normal, this is shown in Table 4.

TABLE IV. THE SHAPIRO-WILK TEST FOR POST-TEST EVALUATION

Post-test			
	Statistic	df	Sig
CG	0.953	43	0.074
EC	0.976	53	0.347

($p=0.05$)

Also, the grade was homoscedastic, $F(94) = 3.721, p = .057$.

The student's t-test shown that the students who used the LOs (EG) and those who attended the lesson face to face (CG) have not statistical difference in post-evaluation, see Table 5.

TABLE V. STUDENT T-TEST OF THE GRADE FOR THE POST-TEST EVALUATION

Post-test					
	M	DS	t-value	df	p-value
CG	7.2563	1.4840	-1.156	94	0.25
EC	7.5447	1.2148			

Therefore, both groups had the same level of knowledge after reviewing the LOs or attending face-to-face lessons. Thus, the LOs designed based on several learning styles of the ULSM do not influence academic performance.

B. Influence of the Dimensions of Learning Styles on Students Performance

A two-way ANOVA test (group x preference) were carried out to analyze each dimension of the students' learning styles in their performance. The ANOVA showed no significant difference between the interaction of the LOs and preferences: kinesthetic, visual, aural, reading/writing, $F(2.89) = 1.293, p = .279$. Thus, the students with kinesthetic, visual, aural and reading/writing preferences obtained the same performance by consulting the LOs or attending the teacher's lecture.

Likewise, the results have shown no significant difference between the interaction of the LOs and the information processing dimension (abstract vs concrete), $F(1.92) = .254, p = .615$. Therefore, students with abstract, concrete preferences obtained the same performance by consulting the LOs or attending class with the teacher. Also, no statistically significant difference has been found between the interaction of the LOs and the information processing dimension (active vs reflective), $F(1.92) = 1.548, p = .217$. Consequently, students from both groups with active or reflective preferences obtained the same performance. The statistical analysis indicated no difference between the interaction of LOs and the information processing dimension (sequential vs global), $F(1.92) = 2.006, p = .160$. Therefore, students with sequential or global preferences obtained the same grade.

Furthermore, no statistically difference between the interaction of the LOs and the field dependence dimension (dependent or independent), $F(1.92) = .047, p = .829$ was found. A similar result was realized for deductive or inductive

learning, $F(1.92) = .166, p = .685$, for synthesis vs analysis, $F(1.92) = .455, p = .502$; and for individual vs team, $F(1.92) = 1.787, p = .185$.

VI. DISCUSSION

The results have shown that there is no evidence that the individual learning styles and the dimensions of the ULSM influence students' academic performance. Therefore, there was no statistical difference between face-to-face and blended learning which shows that blended learning can be just as effective as face-to-face learning. This is certainly a good choice to meet a challenge of the growing student population demanding quality education [37].

A previous study [14], whose objective was to identify the dominant learning style of students, shows that the majority of students are kinesthetic (65%), with inductive reasoning (74%), with a preference for synthesis (66%), deeply motivated (44%), highly persistent (51%), and who prefer to work individually (75%).

The study of [8] showed that the courses recommended based on students learning styles improve academic achievements. In contrast, our results show that learning styles —included in the LOs — do not affect the students learning performance.

Our results agree with those found in [9], who showed that learning styles do not influence students' perception of online learning. It is also consistent with the study [5], which found no significant difference in learning satisfaction changes, motivation, time spent studying, and effort invested in the study after considering the students' learning preferences. Thus, a learning preference is a state—temporary feelings or behaviors—and not a trait —associated with stable characteristics—in students [5].

VII. CONCLUSIONS

In conclusion, research supports the conjecture that learning styles positively influence students' academic performance. On the other hand, our results argue that student performance does not depend on learning style. Consequently, for future work, it is necessary to explore other psychological, affective, and cognitive factors that could influence the students' learning performance to depict “the big picture” and develop strategies to mitigate their negative impact. In developing countries, online learning can be a solution for many aspirants that are not admitted to public universities in a face-to-face career. However, this strategy will be successful if the challenges such as the failure in exams and the high rate of students drop out are solved.

ACKNOWLEDGMENT

This work was supported by Instituto Politécnico Nacional [grant numbers SIP20201101, SIP20200832].

REFERENCES

- [1] ANUIES, “Anuarios Estadísticos de Educación Superior”, Información Estadística de Educación Superior, 2019. [Online]. Available at: <http://www.anui.es/informacion-y-servicios/informacion-estadistica-de-educacion-superior/analisis-estadistico-de-educacion-superior>. [Consulted: 27-feb-2020].
- [2] IPN, “Agenda Estadística. Enero-Diciembre 2018”, 2018.

- [3] IPN, "Estadística Básica. Inicio de periodo escolar 2018-2019/1 y Fin de periodo escolar 2018-2019/2", 2019.
- [4] E. Popescu, "Diagnosing students' learning style in an educational hypermedia system", *Cogn. Emot. Process. Web-Based Educ. Integr. Hum. Factors Pers.*, pp. 187–208, 2009.
- [5] C. A. Giuliano, L. R. Moser, V. Poremba, J. Jones, E. T. Martin, y R. L. Slaughter, "Use of a unified learning style model in pharmacy curricula", *Curr. Pharm. Teach. Learn.*, vol. 6, núm. 1, pp. 41–57, 2014.
- [6] F. Coffield, D. Moseley, E. Hall, y K. Ecclestone, *Learning styles and pedagogy in post-16 learning. A systematic and critical review*, Primera., núm. January. Londres: Learning and Skills Research Centre, 2004.
- [7] P. A. Kirschner, "Stop propagating the learning styles myth", *Comput. Educ.*, vol. 106, pp. 166–171, 2017.
- [8] E. Popescu, "A unified learning style model for technology-enhanced learning: What, why and how?", *Int. J. Distance Educ. Technol.*, vol. 8, núm. 3, pp. 65–81, 2010.
- [9] M. S. A. Mansor y A. Ismail, "Learning Styles and Perception of Engineering Students Towards Online Learning", *Procedia - Soc. Behav. Sci.*, vol. 69, núm. Icepsey, pp. 669–674, 2012.
- [10] D. Kolb, "The Process of Experiential Learning", en *Experiential Learning: Experience As The Source Of Learning And Development*, New Jersey: Prentice Hall, Inc., 1984, pp. 20–38.
- [11] J. O'Malley y H. McCraw, "Students Perceptions of Distance Learning, Online Learning and the Traditional Classroom", *Online J. Distance Learn. Adm.*, vol. 2, núm. 4, pp. 1–10, 1999.
- [12] S. Dinçol, S. Temel, Ö. Ö. Oskay, Ü. I. Erdoğan, y A. Yılmaz, "The effect of matching learning styles with teaching styles on success", *Procedia - Soc. Behav. Sci.*, vol. 15, pp. 854–858, 2011.
- [13] A. Grasha y S. Riechman-Hruska, "Teaching style survey (Encuesta de estilos de enseñanza)", *Teaching style survey*, 1996. [En línea]. Disponible en: <http://www.longleaf.net/teachingstyle.html>. [Consultado: 28-feb-2019].
- [14] Ö. Karakiş, "Dominant Learning Styles of Preparatory Class Students", *Procedia - Soc. Behav. Sci.*, vol. 55, pp. 1079–1088, 2012.
- [15] A. Ramírez-Arellano, J. Bory-Reyes, y L. M. Hernández-Simón, "Learning Object Assembly Based on Learning Styles", en *Smart Education and e-Learning 2016*, 2016, vol. 59, pp. 447–462.
- [16] A. Ramírez-Arellano, J. Bory-Reyes, y L. M. Hernández-Simón, *Learning Object Retrieval and Aggregation Based on Learning Styles*, vol. 55, núm. 6, 2016.
- [17] E. Brown, T. Fisher, y T. Brailsford, "Real users, real results: Examining the limitations of learning styles within AEH", *Hypertext 2007 Proc. Eighteenth ACM Conf. Hypertext Hypermedia, HT'07*, pp. 57–66, 2007.
- [18] F. Romanelli, E. Bird, y M. Ryan, "Learning styles: A review of theory, application, and best practices", *Am. J. Pharm. Educ.*, vol. 73, núm. 1, 2009.
- [19] C. W. Allinson y J. Hayes, "The cognitive style index: A measure of intuition-analysis for organizational research", *J. Manag. Stud.*, vol. 33, núm. 1, pp. 119–135, 1996.
- [20] I. Briggs-Myers, M. McCaulley, N. Quenk, y A. Hammer, *MBTI Manual. A Guide to the Development and Use of the Myers-Briggs Type Indicator*, Tercera. Palo Alto, California: Consulting Psychologists Press, Inc., 1993.
- [21] A. Kolb y D. Kolb, "The Kolb Learning Style Inventory—Version 3.1 2005 Technical Specifications", HayGroup, núm. January 2005, 2015.
- [22] M. J. Apter, *Motivational Styles in Everyday Life. A Guide to Reversal Theory*, Primera. American Psychological Association, 2001.
- [23] R. Dunn y S. Griggs, *Synthesis of the Dunn and Dunn learning-style model research: Who, what, when, where, and so what?* St. John's University Press, 2007.
- [24] N. Entwistle, "Improving Teaching through Research on Student Learning", en *University teaching: International perspectives*, Primera., J. Forest, Ed. Garland Publishing, Inc, 1998, pp. 73–112.
- [25] A. F. Gregorc, *Gregorc Style Delineator: A Self-assessment Instrument for Adults*, 3a ed. Gregorc Associate, 2006, 1985.
- [26] N. Herrmann, *The Whole Brain business book: Unlocking the power of whole brain thinking in organizations, teams, and individuals*. McGraw Hill, 1996.
- [27] P. Honey y A. Mumford, *The Learning Styles Helper's Guide*. Maidenhead, UK: Peter Honey Publications (2000), 2000.
- [28] C. Jackson, *Manual of the Learning Styles Profiler*. Sydney: Cymeon Research, 2002.
- [29] H. A. Witkin, R. B. Dyk, H. F. Fattuson, D. R. Goodenough, y S. A. Karp, *Psychological differentiation: Studies of development*. Oxford, England: Wiley, 1962.
- [30] R. Felder y L. Silverman, "Learning and teaching styles in engineering education", *Eng. Educ.*, vol. 78, núm. 7, pp. 674–681, 1988.
- [31] N. Fleming, "I'm different; not dumb. Modes of presentation (V.A.R.K.) in the tertiary classroom", en *Research and Development in Higher Education, Proceedings of the 1995 Annual Conference of the Higher Education and Research Development Society of Australasia (HERDSA)*, 1995, pp. 308–313.
- [32] R. Riding y S. Rayner, *Cognitive Styles and Learning Strategies. Understanding Style Differences in Learning and Behaviour*. Novena. New York: Routledge. Taylor & Francis Group, 2007.
- [33] G. Pask, "Learning Strategies, teaching strategies, and conceptual or learning style", en *Learning Strategies and Learning Styles*, R. R. Schmeck, Ed. New York: Plenum Press, 1988, pp. 83–100.
- [34] A. Grasha, "Teaching with Style: The integration of teaching and learning styles in the classroom.", *Prof. Organ. Dev. Netw. High. Educ.*, vol. 7, núm. 5, 1995.
- [35] S. Horton, *Web Teaching Guide: A Practical Approach to Creating Course Web Sites*, Primera. Yale University Press, 2000.
- [36] IEEE Learning Technology Standards Committee, "The Learning Object Metadata standard", Working groups and study groups, 2007. [En línea]. Disponible en: <https://www.ieeeltsc.org/working-groups/wg12LOM/lomDescription/>. [Consulted: 29-feb-2020].
- [37] La Jornada, "Entre 2000 y 2021, la matrícula de educación superior creció 2 millones", 2021. [On line]. Available at: <https://www.jornada.com.mx/notas/2021/03/17/sociedad/entre-2000-y-2021-la-matricula-de-educacion-superior-crecio-2-millones/> [Consulted: 27-jun-2021].

Application of Expert System with Smartphone-based Certainty Factor Method for Hypertension Risk Detection

Dodon Yendri¹, Derisma², Desta Yolanda³, Sabrun Jamil⁴
Computer Engineering Department
Andalas University, Padang, Indonesia

Abstract—Hypertension is a factor health problem with a high prevalence in Indonesia. The total amount of hypertension nationwide is 34.11%. There are only 1/2 were diagnosed, and the remaining 1/2 are undiagnosed. Certainty factor was needed to prove the certain or not of a fact in a metric form, and it commonly was used in expert systems. This method was perfect for diagnosing something uncertain. This research aimed to build a smartphone-based hypertension risk detection system. This system was built by MPX5500DP pressure sensor components that served for blood pressure measurement, Bluetooth module HC-05 as microcontroller data transmission, Arduino Uno as sensor data processing, and Smartphone as hypertension risk detection interface and database access. Measuring blood pressure on the user's right or left arm was used to test the system. Then the data was sent to the smartphone to be classified and processed by using the certainty factor method implanted in the android smartphone to detect the risk of hypertension and then followed by the selection of symptoms and risk factors of hypertension according to previous experience. The results showed that the success rate of the system in detecting the risk of hypertension was 100%, the mean error of the systolic value of the right arm was 2.092%, and the average error of the diastolic value of the right arm was 2.977%, while the average error of the systolic value of the left arm. 1.944% and the mean error of the left arm diastolic value was 2.800%.

Keywords—Hypertension; certainty factor; blood pressure; mpx5500dp

I. INTRODUCTION

Based on the 2018 Basic Health Research, the national hypertension prevalence is 34.11%. In the percentage of 34.11%, half of the people who have hypertension have been diagnosed, and most of them undiagnosed. Data shows that only 54.4% of people diagnosed with high blood pressure take hypertension medication [1]. This explanation tells that most hypertension sufferers are unaware that they are suffering from hypertension [2],[3],[4] and are not getting treatment [4], [5]. Hypertension is one of the big problems in the public health field [6], especially in Indonesia. That expressed because hypertension is a health problem with a high prevalence [3], [7], [8]. The Estimated one in three people will suffer from hypertension in 2025 [4].

Hypertension is a condition where the systolic blood pressure value is ≥ 140 mmHg, or the diastolic value is ≥ 90 mmHg [9]. The increase of blood pressure in a long time can cause kidneys, heart, and brain damage until stroke if it is not

detected early [10] and does not receive adequate treatment [11], [12]. A sphygmomanometer, a tool in the medical world, is used to check blood pressure [13] in humans and, it also a medium that can detect hypertension. The Sphygmomanometer consists of an analog sphygmomanometer and a digital Sphygmomanometer [14]. Analog Sphygmomanometer works manually means to be able to know blood pressure precisely depend on the user's expertise [15]. It works using Korotkoff methods in determining the patient's systolic and diastolic through coronary sound from the stethoscope. While digital A sphygmomanometer works based on oscillometry methods in which to determine systolic and diastolic patients use pressure sensors as transducers that will detect blood pressure and oscillation signal changes due to heart rate [4], [16].

There are many mobile health apps (mHealth apps) available in the market currently. This app is designed to facilitate a wide range of health issues and concern that is intended for use outside of the clinic. The review study also found that researchers and publishers of *mHealth* focused more on individual preventative diseases such as self-management for diabetes [17], [18], [19], gout [20], chronic obstructive pulmonary disease [21], bipolar disorder [22] and bowel disease [23], stroke [24], [25], heart disease [26] and flu [27].

Fitriani et 'al research [28] developed a mobile application for early prediction of diabetes type 2 and hypertension based on the factor data individual risk by integrating iForest, SMOTETomek, and ensemble learning. iForest was used to detect and remove outlier data from the dataset. Meanwhile SMOTETomek was used to balance unbalanced datasets and ensemble learning was applied to predict disease. This study collected risk factor data and sent it to a remote server for diagnosis with the proposed prediction model. The prediction results were sent back to the car application. The weakness of this research was the process of sending data back and forth to get action to prevent individual risk.

Furthermore, Egejuru et 'al [29] developed a cellular-based hypertension risk monitoring system. BP_HRM System developed for users can access the system from distant locations easily. The mobile application was designed using an integrated modeling language and implemented using the Adaptive Neuro-Fuzzy Inference System (ANFIS) and Extensible Mark-Up Language methods with the Java

programming language for layout and mobile content. JavaScript Object Notation was used to implement the system's data storage and retrieval mechanisms. This study had not provided information about the symptoms and factors that caused hypertension.

Based on previous research, we have carried out developmental research in addition to classifying hypertension as well as knowing the symptoms and factors that cause hypertension and the level of risk that will occur.

II. COMMON FACTORS AND SYMPTOMS OF HYPERTENSION

Hypertension is the pressure that pushes a wall of the arteries when the blood pumps into the arteries [30]. Systolic pressure or high blood pressure is given to the walls of the arteries when the heart contracts, whereas diastolic pressure or low blood pressure, the pressure given to the arteries when the heart relaxes. Blood pressure is generally measured in millimeters of mercury (mmHg) [31].

Hypertension referred to as "the silent killer" [32], [33]. Generally, hypertension is asymptomatic [34]. Most people don't feel anything, even though their blood pressure is not normal [18]. It can continue for many years until the sufferer (who does not feel suffering) falls into an emergency and even has heart disease, sexual dysfunction, stroke, or damaged kidneys [35]. Based on the case, some hypertension sufferers do not specific hypertension symptoms, they only get headaches, impaired vision, restlessness, dizziness, fatigue, pain in the chest, and heavy feeling on the nape [36].

Hypertension generally causes the sufferer's unhealthy behavior. Some factors unhealthy behaviors can cause hypertension [8], [37]:

1) *Smoking*: Smoking increases the heart rate [37]. Therefore the muscles - heart muscles need a lot of oxygen. Smoking habits of hypertension sufferers will increase the risk of damage to arterial blood vessels.

2) *Lack of physical activity*: Exercise can help stabilize blood pressure[38], and it is beneficial for sufferer's mild hypertension. Moreover, aerobic exercise can drop blood pressure, although the weight has not dropped.

3) *Excessive salt consumption*: Salt causes a build-up of fluid in the body because it collects fluid inside the cell. It can cause the volume of blood pressure to increase [39]. The people who consume 3.5 grams or less salt had lower blood pressure than the community that consumes about 9-12 grams [40].

4) *Excessive alcohol consumption*: The effect of alcohol can increase blood pressure. Several studies showed a relationship between blood pressure and alcohol intake. The American Heart Association recommends limiting alcohol intake in order to avoid developing hypertension. Women should have no more than 1 drink per day and no more than 2 drinks per day for men [41].

5) *Psychosocial and stress*: Stress or mental tension (depressed, moody, angry, revenge, fear, guilt) can stimulate the kidney glands to release adrenaline hormones and spur the beat heart is faster and stronger than usual. It can cause blood

pressure to increase [42]. If people get stress for a long time, their body will occur pathological changes and organic abnormalities. The Symptoms can be hypertension or stomach ulcers. Table 1 shows the common factors and symptoms of hypertension.

TABLE I. SYMPTOMS AND RISK FACTORS FOR HYPERTENSION [18]

No.	Symptom Name	Value (CF)
Physical Complaints		
1.	Headache	0.6
2.	Dizzy	
3.	Tired easily	
4.	Restless	
5.	Pain in the chest	
6.	Interrupted vision	
7.	Heart palpitations	
8.	Nape weight	
Personal History of Hypertension		
9.	In the Treatment Period for Hypertension (anti-hypertensive drugs)	0.6
History of Hard Disease		
10.	Heart	0.8
11.	Brain	
12.	Eyes	
13.	Diabetes Mellitus	
14.	Kidney	
Social History		
15.	Doing Strenuous Activities	0.4
16.	Moker	
17.	Consuming alcohol	
18.	Consumption of Excessive Salt	
19.	consuming fatty foods	
20.	consuming coffee	
21.	Stress	
22.	Lack of exercise	
Family History of Hypertension		0.2

III. CLASSIFICATION OF HYPERTENSION

Hypertension has two types Based on the cause. They are Essential hypertension (primary) and Secondary hypertension [43], [44]. Essential hypertension (primary) is a type of hypertension that often occurs in hypertension cases about 95 percent. This type is related to lifestyle factors such as lack of movement (inactivity) and diet [45]. Meanwhile, secondary hypertension, type hypertension seldom in hypertension cases. It is related to medical conditions, such as heart disease or reactions medications.

According to the Sevent's Report of the Joint National Committee on Detection, Evaluation and Treatment of High Blood Pressure (JNC7) 2003 [41], [46], hypertension classified into four categories in Table 2.

TABLE II. CLASSIFICATION OF HYPERTENSION ACCORDING TO JNC7 2003 [41]

Category	TDS (mmHg)		TDD (mmHg)
Normal	<120	and	< 80
Prehypertension	120-139	or	80-89
Hypertension stage 1	140-159	or	90-99
Hypertension stage 2	≥ 160	or	≥ 100

IV. CERTAINTY FACTOR

The certainty factor method is used in facing a problem that the answer is uncertain. This uncertainty is a probability. This method was introduced in the 1970s by Short-life Buchanan. Buchanan used this method when he diagnosed and therapy meningitis and blood infections (Wikipedia). The development team of this method noted that doctors often analyzed existing information with phrases such as almost certain.

This method was almost similar to fuzzy logic because uncertainty is represented by the degree of trust. The difference is in the fuzzy logic's rules [47]. The fuzzy logic does not have a confidence value if it calculates the premise more than one. The fuzzy logic's calculation can see the lower value in AND operator and the highest value in the OR operator. It is different with a certainty factor [48]. It has trust value on each premise and shows the measure of certainty of a fact or rule.

$$CF[h,e] = MB[h,e] - MD[h,e] \quad (1)$$

Description:

$CF[h,e]$ = Certainty factor

$MB[h,e]$ = Measure of belief, a measure of belief or confidence level of the hypothesis (h), if given evidence (e) between 0 and 1.

$MD[h,e]$ = Measure of disbelief, a measure of disbelief, or confidence level of the hypothesis (h), if given evidence (e) between 0 and 1.

There are several combinations of certainty factor against certain premises, namely:

1) Certainty factor with one premise.

$$CF[h,e] = CF[e]*CF[rule] = CF[user]*CF[pakar] \quad (2)$$

2) Certainty factor with more than one premise.

$$CF[A \wedge B] = \text{Min}(CF[a],CF[b]) * CF[rule] \quad (3)$$

$$CF[A \vee B] = \text{Max}(CF[a],CF[b]) * CF[rule] \quad (4)$$

3) Certainty factor with a similar conclusion.

$$CF_{\text{combined}}[CF1,CF2] = CF1 + CF2 * (1-CF1) \quad (5)$$

This method has advantages such as suitable for measuring something certain or uncertain in diagnosing a disease. In addition, this method only applied one calculation, and it reached out two data therefore its accuracy was maintained.

The planning of the test was needed by getting the best result of the test, both hardware and software testing. Table 3 shows the hardware test design, and Table 4 shows the software design.

TABLE III. HARDWARE TEST DESIGN

No	Hardware	Test Design	Destination
1	blood pressure measuring device	Get error values for measuring tools and digital sphygmomanometers.	Get a comparison of the measurement results of tools and digital sphygmomanometers.

TABLE IV. SOFTWARE TEST DESIGN

No	Hardware	Test Design	Destination
1	Mobile Application	Testing the detection of hypertension risk was successfully obtained from the data obtained	To find out the results of detecting the user's hypertension risk successfully obtained
2	Certainty factor method	Testing the obtained hypertension risk value	To determine the risk value of hypertension obtained

V. ANALYSIS OF RESULT

The expert system implementation shows in Fig. 1 using a blood pressure gauge that is Arduino Uno, MPX5500DP Pressure Sensor, Bluetooth HC-05 module, Arm cuff is an airbag that circles on the user's arm, and Android Smartphone.

1) *MPX5500DP Pressure sensor testing*: This test found out the extended function of the MPX5500DP sensor measuring the user's blood pressure. Testing was conducted by connecting sensors, cuffs, and analog Sphygmomanometers to a hose that has three ends. Once the three components had connected, the air pumped into the cuff until it gets a pressure value to be tested. It consists of the systolic value(s) and the diastolic value (d) read by the embedded software. Once the pressure value is obtained, then it is sent to the app via a serial communication function between Arduino and the android app via Bluetooth. The results of pressure measurements using the system shown in Fig. 2:

Furthermore based on the pressure value obtained, then the value is classified and displayed on the smartphone as seen in Fig. 3.

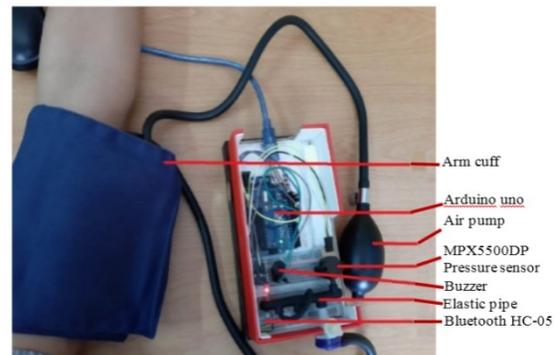


Fig. 1. Blood Pressure Gauge.

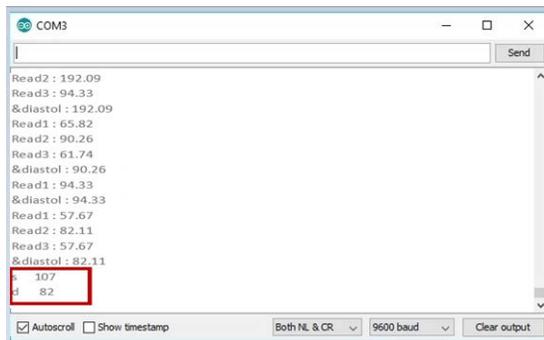


Fig. 2. Measurement of Blood Pressure by the System.

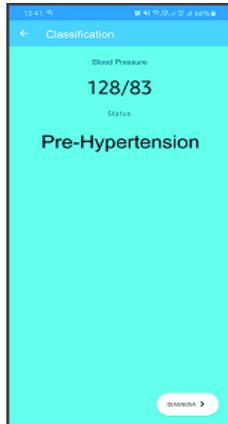


Fig. 3. Pressure Classification Page.

The determining of the system accuracy was conducted by comparing medical devices that had the same function as analog Sphygmomanometers. First, the cuffs were attached to the user's arm, and then performed air pumping through an elastic balloon into the cuff until the air pressure on the cuff reached ≥ 180 mmHg, then the air was released slowly by the mechanic's tool. It took systolic and diastolic values when the pressure drops continuously. Systolic value took if the tool would reach out the first spike in the pressure value, and Diastolic value took the last spike in pressure. After the data obtained, it is sent to the app via Bluetooth, and it is transferred to the smartphone to be classified and processed using the certainty factor method. Furthermore, the same way did test with analog Sphygmomanometers.

Fig. 4 shows the results of a blood pressure measurement from 20 tests using the MPX5500DP pressure sensor and analog Sphygmomanometer.

From the graph in Fig. 7, it can be seen that the measurement results using the MPX5500DP pressure sensor are almost the same as the measurements using an analog Sphygmomanometer.

2) *Blood pressure testing on right arm and left arm:* This test is to find out the system extent implemented to measure the user's blood pressure. This test is performed on the right arm and left arm user. Fig. 5 and Fig. 6 show the comparison of measurement test results of blood pressure on the right arm and left arm from 10 different users using mpx5500DP sensors and analog Sphygmomanometers.

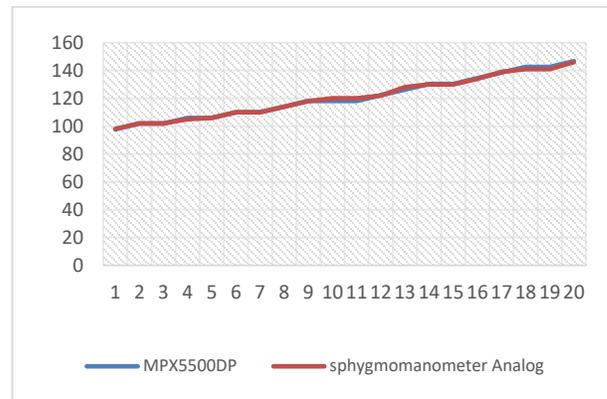


Fig. 4. Blood Pressure Comparison.

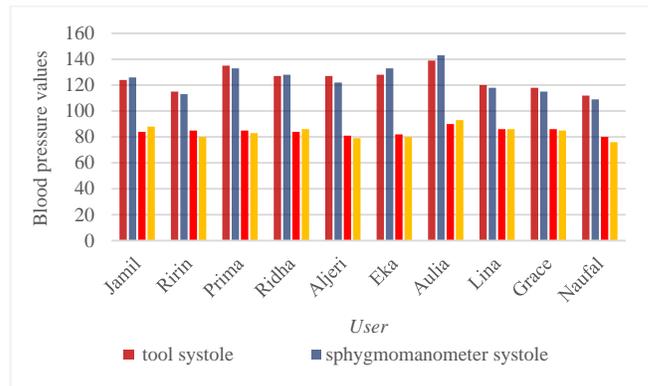


Fig. 5. Blood Pressure Measurement Testing Graph on Right Arm.

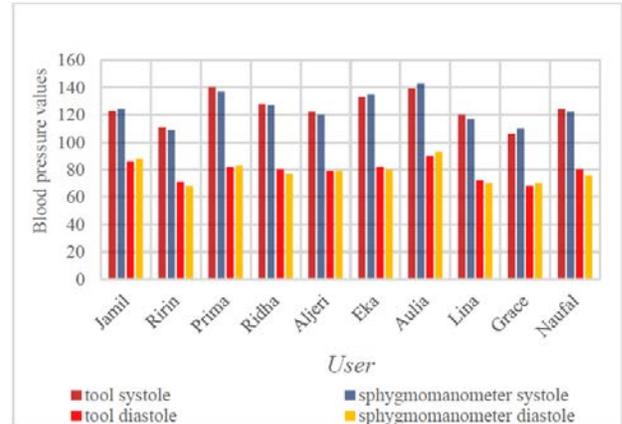


Fig. 6. Blood Pressure Measurement Testing Graph on Left Arm.

Based on a graph test performed shows in Fig. 5 and Fig. 6 states that testing using the MPX5500DP sensor and analog Sphygmomanometer provides almost the same results. From the data of blood pressure measurement tests on both arms, the calculation of systolic and diastolic error values of both of them was calculated using the formula:

- $$\text{Average Right Arm Systolic Value Error} = \frac{\text{Number of Right Arm Systolic Error Values}}{\text{Number of Experiments}} \times 100\% = \frac{20.927}{10} \times 100\% = 2.097\%$$

- Average Right Arm Diastolic Value Error =
$$\frac{\text{Number of Right Arm Diastolic Error Values}}{\text{Number of Experiments}} = \frac{29.778}{10} \times 100\% = 2.977\%$$
- Average Left Arm Systolic Value Error =
$$\frac{\text{Number of Left Arm Systolic Error Values}}{\text{Number of Experiments}} = \frac{19.441}{10} \times 100\% = 1.944\%$$
- Average Left Arm Diastolic Value Error =
$$\frac{\text{Number of Diastolic Tool Error Values}}{\text{Number of Experiments}} = \frac{28.009}{10} \times 100\% = 2.800\%$$

From the calculation of the equation above, there was a significant difference in the average value of systolic and diastolic errors in both arms. The difference between them could be caused by several vascular abnormalities in both of a person's arms. The greatest possibility was a blockage of the walls of arterial blood vessels, either due to fat or other plaque, on arms with higher blood pressure. The difference in the average error value of both arms shows the graph in Fig. 7.

3) *Accuracy testing*: Accuracy testing is performed to determine the capabilities of expert systems embedded in the software using the Certainty factor method. This test shows that the application of the certainty factor method could calculate the hypertension risk for the users based on the symptoms and risk factors of hypertension chose by them.

The Table 5 shows the results of accuracy tests conducted based on one of the system tests as a whole.

Based on Table 3 of the test results above, the percentage value of Hypertension risk was derived from the equation of the Certainty factor method embedded in hypertension risk detection software. The equation of the certainty factor method would process the certainty factor (CF) value according to the symptoms and hypertension factors selected by the user. This is a manual calculation of certainty factor equations to see the accuracy of the methods embedded in the software:

$$CFA = CF1 + CF2 * (1 - CF1)$$

$$CFA = 0,6 + 0,4 * (1 - 0,6)$$

$$CFA = 0,6 + 0,4 (0,4)$$

$$CFA = 0,6 + 0,16$$

$$CFA = 0,76$$

$$\text{Percentage CF} = 0,76 * 100\% = 76\%$$

CF1 and CF2 values refer to Table 2. The results of the calculation of hypertension risk with certainty factor method manually according to the results of detection using the certainty factor method embedded into hypertension risk detection software.

4) *Early detection testing of overall hypertension risk*: This test was conducted to determine whether an Android-based expert system application has an embedded knowledge base on detecting hypertension risk that supports the system using the Certainty Factor method. The certainty factor method could show a measure of certainty based on facts or rules. That factor method showed the results in the numbers form. The use of blood pressure measurement data had classified an application implanted in a smartphone, then the system processes it by implementing equations (5). The results of the process were stored in Firebase and displayed on android smartphones in real-time. Table 6 shows the results of blood pressure classification testing and early detection of Hypertension Risk levels based on symptoms and factors displayed on smartphones from 10 test cases using the Certainty Factor method.

From Table 4, it can be explained that the system can receive blood pressure values from blood pressure measuring devices and classify them based on the received blood pressure according to the knowledge that has been instilled (based on Table 2).

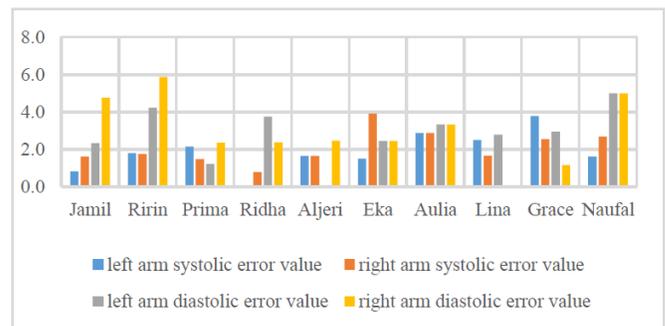


Fig. 7. Graph of difference in Average Blood Pressure Testing Error on the Right Arm and Left Arm.

TABLE V. HYPERTENSION RISK TEST RESULTS

Name	Blood pressure received		Classification	Symptoms and factors	Risk of Hypertension with Certainty factor
	S	D			
Jamil	123	86	Pre-Hypertension	<ul style="list-style-type: none"> • Headache • Nape Weight • Heart palpitations • Coffee consumption • Stress • Lack of exercise 	76.0 % (True)

TABLE VI. RESULTS OF BLOOD PRESSURE CLASSIFICATION TESTING AND THE EARLY DETECTION OF HYPERTENSION RISK LEVELS

No	Name	Blood pressure received		Classification	Symptoms and factors	Risk of Hypertension with Certainty factor (True/False)
		S	D			
1	Jamil	123	86	Pre-Hypertension	<ul style="list-style-type: none">• Headache• Nape Weight• Heart palpitations• Coffee consumption• Stress• Lack of exercise	76.0 % (True)
2	Ririn	111	71	Normal	<ul style="list-style-type: none">• Headach• Impaired vision• Consuming excess salt• Consuming coffee• Stress• Lake of Exercise	76.0% (True)
3	Prima	140	82	Pre-Hypertension	<ul style="list-style-type: none">• Consuming fatty foods• Consuming coffee• Stress• Lack of exercise• Have a family history of hypertension	52.0% (True)
4	Ridha	128	80	Pre-Hypertension	<ul style="list-style-type: none">• Tired easily• Take Hypertension Drugs• Diabetes Mellitus Lack of exercise	96.46% (True)
5	Aljeri	122	79	Normal	<ul style="list-style-type: none">• Tired easily• Consumption of Hypertension Drugs• Diabetes mellitus• Lack of exercise	80.8% (True)
6	Eka	133	82	Pre-Hypertension	<ul style="list-style-type: none">• Headache• Dizzy• Consumption of Hypertension Drugs• Diabetes mellitus• Consumption of fatty foods• Lack of exercise• Have a family history of hypertension	98.46% (True)
7	Aulia	139	90	Pre-Hypertension	<ul style="list-style-type: none">• Headache• Undergo strenuous activity• Consumption of fatty foods• Excessive Salt Consumption• Stress• Lack of exercise• Have a family history of hypertension	80.8% (True)
8	Lina	120	72	Normal	<ul style="list-style-type: none">• Headache• Dizzy• Excessive Salt Consumption• Lack of exercise	76.0% (True)
9	Grace	106	68	Normal	<ul style="list-style-type: none">• Stress• Lack of exercise	40.0% (True)
10	Naufal	124	80	Normal	<ul style="list-style-type: none">• Dizzy• Impaired vision• Smoker• Lack of exercise	60.0% (True)

And the process of detecting the risk of hypertension continued with the selection of symptoms and risk factors of hypertension according to the user's experience. Meanwhile, the results of the detection were stored in a database and displayed on a smartphone (see Table 1). The percentage value of Hypertension Risk was obtained from the equation using the Certainty Factor method which was calculated based on the hypertension risk or the symptoms of hypertension by examined users. The results obtained from calculating the risk of hypertension with the certainty factor method, were manual system followed the results obtained from applications implanted on Android, and the success rate was 100%. Therefore the built system using certainty factor methods could work properly and provide results as expected.

VI. CONCLUSION

The system built for the blood pressure measurement process using MPX5500DP sensor was done by reading pressure changes on the arm cuff that wore by the user using the certainty factor method can give the following conclusions:

1) The certainty factor method is able to detect the risk of hypertension with the equation $CF(n) = CF(x) + CF(y) * (1 - CF(x))$.

2) The Data hypertension risk detection contains in the Firebase Database Real-time displayed successfully on the smartphone.

3) The blood pressure testing gauges conducted on the right arm of 10 users produced an average error value of 2.092% on systolic value readings and 2.977% on diastolic values.

4) The blood pressure testing gauges conducted on the left arm of 10 users produced an average error value of 1,94% on systolic value readings and 2.8% on diastolic values.

5) The application of calculation hypertension risk implanted in android using the certainty factor method can give success results rate 100%.

6) In the subsequent development of hypertension risk detection software, use different methods such as Machine Learning or other Decision Support System methods.

REFERENCES

- [1] Badan Penelitian dan Pengembangan Kesehatan., Riset Kesehatan Dasar 2018. Jakarta: "Kementerian Kesehatan Republik Indonesia", http://labdata.litbang.kemkes.go.id/images/download/laporan/RKD/2018/Laporan_Nasional_RKD2018_FINAL.pdf [accessed on June, 19, 2021, 07.40 AM].
- [2] Shikha Singh, Ravi Shankar, and Gyan Prakash Singh., "Prevalence and Associated Risk Factors of Hypertension: A Cross-Sectional Study in Urban Varanasi", *International Journal of Hypertension*, 2017; 2017: 5491838, Published online 2017 Dec 3. doi: 10.1155/2017/5491838.
- [3] N. M. Banki, S. L. Chan, V. A. Rao, R. B. Melles, and D. L. Bhatt, "Effect of Systolic and Diastolic Blood Pressure on Cardiovascular Outcomes," pp. 243–251, 2019, doi: 10.1056/NEJMoa1803180.
- [4] M. M. M. P. H. C-phi, J. A. Osheroff, and B. R. Mph, "Improving Identification and Diagnosis of Hypertensive Patients Hiding in Plain Sight (HIPS) in Health Centers," *Jt. Comm. J. Qual. Patient Saf.*, vol. 44, no. 3, pp. 117–129, 2018, doi: 10.1016/j.jcjq.2017.09.003.
- [5] Mahnaz Ashoorkhani, Reza Majdzadeh, Jaleh Gholami, Hassan Eftekhari, and Ali Bozorgi, "Understanding Non-Adherence to Treatment in Hypertension: A Qualitative Study," *Int J Community Based Nurs Midwifery*. 2018 Oct; 6(4): 314–323., <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6226612/>.
- [6] J. Kitt, R. Fox, K. L. Tucker, R. J. Mcmanus, and K. L. Tucker, "New Approaches in Hypertension Management : a Review of Current and Developing Technologies and Their Potential Impact on Hypertension Care," 2019.
- [7] Shikha Singh, Ravi Shankar, and Gyan Prakash Singh., "Prevalence and Associated Risk Factors of Hypertension: A Cross-Sectional Study in Urban Varanasi", *International Journal of Hypertension*, 2017; 2017: 5491838, Published online 2017 Dec 3. doi: 10.1155/2017/5491838.
- [8] W. K. Bosu and D. K. Bosu, "Prevalence , awareness and control of hypertension in Ghana : A systematic review and meta-analysis". 2021.
- [9] NICE Guideline, "Hypertension in adults: diagnosis and management", London: National Institute for Health and Care Excellence (UK); 2019 Aug. ISBN-13: 978-1-4731-3503-1.
- [10] C. Update, "The Transition From Hypertension to Heart Failure," vol. 5, no. 8, 2017, doi: 10.1016/j.jchf.2017.04.012.
- [11] E. Ku, B. J. Lee, J. Wei, and M. R. Weir, "Hypertension in CKD : Core Curriculum 2019," *Am. J. Kidney Dis.*, vol. 74, no. 1, pp. 120–131, 2019, doi: 10.1053/j.ajkd.2018.12.044.
- [12] Y. W. Kusumaningtyas, T. B. Indrato, M. P. A. T. P, and B. Utomo, "Digital Sphygmomanometer Based on Arduino Using TFT LCD Display," vol. 1, no. 1, pp. 34–38, 2019, doi: 10.35882/ijeemi.v1i1.6.
- [13] S. Definition, "Sphygmomanometer . Definition , How to Use Sphygmomanometer Definition Practice How To Use a Sphygmomanometer Procedures," pp. 1–3, 2021.
- [14] O. Article, B. Shahbabu, A. Dasgupta, K. Sarkar, and S. K. Sahoo, "Which is More Accurate in Measuring the Blood Pressure ? A Digital or an Aneroid Sphygmomanometer," 2016, doi: 10.7860/JCDR/2016/14351.7458.
- [15] F. Written, "Does High Blood Pressure Cause Headaches ?," pp. 1–11, 2021.
- [16] W. H. O., "Who Technical Specifications for Automated Non-Invasive Blood Pressure Measuring Devices With Cuff Who Medical Device Technical Series", 2020, ISBN 978-92-4-000265-4 (electronic version), ISBN 978-92-4-000266-1 (print version).
- [17] K. Santo et.al, "The Potential of mHealth Applications in Improving Resistant Hypertension Self-Assessment, Treatment and Control, National Library of Medicine", *Curr Hypertens Rep .* 2019 Oct 9;21(10):81. doi: 10.1007/s11906-019-0986-z.
- [18] V. N. Shah and S. K. Garg, "Managing diabetes in the digital age," *Clin. Diabetes Endocrinol.*, no. January, 2016, doi: 10.1186/s40842-015-0016-2.
- [19] J. R. White, "Evaluation and Evolution of Diabetes Mobile Applications : Key Factors for Health Care Professionals Seeking to Guide Patients," pp. 211–215, 2011.
- [20] Nguyen, A. D. et al. "Mobile applications to enhance self-management of gout". *International journal of medical informatics*, 94, 2016, pp. 67-74.
- [21] M. Hardinge et al., "Using a mobile health application to support self-management in chronic obstructive pulmonary disease : a six-month cohort study," *BMC Med. Inform. Decis. Mak.*, pp. 1–10, 2015, doi: 10.1186/s12911-015-0171-5.
- [22] Matthews, M. et al. "Development and evaluation of a smartphone-based measure of social rhythms for bipolar disorder". *Assessment*, 23(4), 2016, pp. 472–483. <https://doi.org/10.1177/1073191116656794>.
- [23] D. Con and P. De Cruz, "Mobile Phone Apps for Inflammatory Bowel Disease Self-Management : A Systematic Assessment of Content and Tools," no. July 2019, 2016, doi: 10.2196/mhealth.4874.
- [24] Feigin, V. et al., "Geomagnetic Storms Can Trigger Stroke: Evidence From 6 Large Population-Based Studies in Europe and Australasia, Stroke", 45(6), 2014, pp. 1639-1645 <https://doi.org/10.1161/STROKEAHA.113.004577>.
- [25] N. Kasabov et al., "Evolving spiking neural networks for personalised modelling , classification and prediction of spatio-temporal patterns with a case study on stroke," pp. 1–5, 2021.

- [26] Schuurin, M. J. et al. "Mobile health in adults with congenital heart disease: current use and future needs". *Netherlands Heart Journal*, 24(11), 2016 pp. 647-652. <https://doi.org/10.1007/s12471-016-0901-z>.
- [27] Li, N. et al. "HHeal: A Personalized Health App for Flu Tracking and Prevention". In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, April 2015, pp. 1415-1420. <https://doi.org/10.1145/2702613.2732804>.
- [28] N. C. Egejuru and P. A. Idowu, "Development of a Mobile-Based Hypertension Risk Monitoring System," no. July, pp. 11–23, 2019, doi: 10.5815/ijieeb.2019.04.02.
- [29] N. L. Fitriyani and M. Syafrudin, "Development of Disease Prediction Model Based on Ensemble Learning Approach for Diabetes and Hypertension," *IEEE Access*, vol. 7, pp. 144777–144789, 2019, doi: 10.1109/ACCESS.2019.2945129.
- [30] Health Encyclopedia, "High Blood Pressure / Hypertension What is high blood pressure?," pp. 1–9, 2021.
- [31] J. S. Shahoud, T. Sanvictores, and N. R. Aeddula, "Physiology , Arterial Pressure Regulation," no. Cn Ix, pp. 4–7, 2021.
- [32] R. Posts, "High blood pressure : The silent killer," pp. 19–21, 2021.
- [33] M. C. Marketplace, "High blood pressure (hypertension)."
- [34] W. H. O., "Hypertension 17," no. May, pp. 4–7, 2021.
- [35] T. R. Pressure, "Prehypertension : A Little Too Much Pressure , A Lot of Trouble What Do the Blood Pressure Numbers Mean ? Can You Help Control Your Blood," pp. 1–3, 2021.
- [36] H. Blood and P. Guide, "Symptoms of Hypertension," pp. 1–18, 2021.
- [37] A. E. Taylor et al., "Effect of Smoking on Blood Pressure and Resting Heart Rate," pp. 832–841, 2015, doi: 10.1161/CIRCGENETICS.115.001225.
- [38] Sheila M. Hegde and Scott D. Solomon, Influence of Physical Activity on Hypertension and Cardiac Structure and Function, *Curr Hypertens Rep*. 2015 October ; 17(10): 77. doi:10.1007/s11906-015-0588-3.
- [39] N. Rakova et al., "and decreases fluid intake Increased salt consumption induces body water conservation and decreases fluid intake," vol. 127, no. 5, pp. 1932–1943, 2017.
- [40] W. H. O., "Salt reduction 29," no. April 2020, pp. 1–6, 2021.
- [41] J. T. Section, "High Blood Pressure from Alcohol Consumption Need Help Now?," pp. 1–13, 2021.
- [42] S. Ranabir and K. Reetu, "Stress and hormones Gonadal dysfunction," pp. 4–9, 2021.
- [43] Prodia, Hypertension: The Oft-Ignored Cause of a Heart Attack. <http://m.prodia.co.id/en/content/viewcontentsdetails/hipertensi-penyebab-serangan-jantung-yang-sering>. [accessed on June 20, 2021, 23:25 PM].
- [44] Matthew R Alexander, et al, "What is the difference between primary (essential) and secondary hypertension (high blood pressure)?," p. 241381, 2021.
- [45] E. M. Amoah, D. E. Okai, A. Manu, A. Laar, J. Akamah, and K. Torpey, "The Role of Lifestyle Factors in Controlling Blood Pressure among Hypertensive Patients in Two Health Facilities in Urban Ghana : A Cross-Sectional Study," vol. 2020, 2020.
- [46] P. Verdecchia and F. Angeli, "The Seventh Report of the Joint National Committee on the Prevention , Detection , Evaluation , and Treatment of High Blood Pressure : The Weapons Are Ready," pp. 843–847, 2003.
- [47] A. Agusta, F. Y. Arini, and R. Arifudin, "Implementation of Fuzzy Logic Method and Certainty Factor for Diagnosis Expert System of Chronic Kidney Disease," vol. 2, no. April, pp. 61–68, 2020.
- [48] A. S. Sembiring, O. Manahan, and M. Helentina, "Implementation of Certainty Factor Method for Expert System," 2019, doi: 10.1088/1742-6596/1255/1/012065.

Using API with Logistic Regression Model to Predict Hotel Reservation Cancellation by Detecting the Cancellation Factors

Sultan Almotiri¹, Nouf Alosaimi², Bayan Abdullah³,
College of Computer and Information Systems
Umm Al-Qura Univeristy

Abstract—The aim of establishing hotels is to provide a service activity to its customers with the aim of making a profit. So, for that the cancellations are a key perspective of inn income administration since their effectiveness on room reservation systems. Cancelling the reservation eliminates the outcome. Many expected factors affect this problem. By knowing these factors, the hotel management can make a suitable cancellation policy. This project aims to create an API that can provide a function to predict if a reservation is most likely to cancel or not. That API can integrate with the hotel management systems to evaluate each reservation process with the same parameters. To do this, the study starts by defining the factors using Chi test, correlation to find the effective variables, and coefficient of the variables in the linear regression. And the results that have been found for the factors are: `is_repeated_guest`, `previous_cancellations`, `previous_bookings_not_cancelled`, `required_car_parking_spaces`, and `deposit_type`. For API function, the intercept and coefficients have been used from the logistics regression model to create a scoring function. Scoring function can be calculated by the sum of the factors multiplied by their coefficients in addition to the intercept. This score is to be evaluated as a probability later using the logistic function.

Keywords—Prediction; API; factors; logistics regression

I. INTRODUCTION

The tourism sector in any region is based on a several of constituents and pillars, among them are the availability of tourist facilities and number and quality of services provided by those facilities. Perhaps the most prominent of these facilities, and the most important, is the hotel sector. In numerous tourism goals, the inn industry is regularly the major financial sector [1]. Hotels are defined as those that provide places for accommodation, meals, and other services for visitors in general. Hence, the presence of hotels in countries has become an indispensable thing. Committing room numbers to the diverse reservations could be a day by day action in inns [2]. In hotels, the revenue Management employments data-driven modelling and optimization strategies to choose What, when, whom to offer, and for which cost, in arrange to extend revenue and benefit [3].

The hotel industry is considered as a basic pillar of tourism because it provides services in the field of economy, because of the means of transportation it pumps and because it is a means of obtaining aid for comprehensive development plans in the world. The aim of establishing hotels is to provide a service activity to its customers with the aim of making a profit. Consequently, the cancellations are a key perspective of

inn income administration since their effectiveness on room reservation systems. By means of the e-commerce increases the chance of cancellation since client can effortlessly compare conditions among lodgings and can cancel his/her reservation [4].

Moreover, Lodging cancellation can cause numerous issues on the hospitality industry in numerous perspectives such as affecting hotel's reputation, since clients can be influenced by others' criticism about the inn [5], [6]. A lodging booking cancellation arrangement can depend on a few variables, such as the rate of the booking and the date of check-in. Cancelled by certain date approach, this type of cancellation approach gives travelers the choice to cancel inn reservations free of charge up until a certain date. Once this date passes, inn cancellation approaches can either: Charge a standard cancellation fee require the complete installment for the reservation. One-night penalty policy, for this sort of lodging room cancellation arrangement, inn charges a cancellation expense proportionate to one night's remain at the inn. The prepaid, nonrefundable hotel reservation policy, this sort of inn booking requires installment for the aggregate of the reservation at the time of booking and is nonrefundable. In truth, exceptionally small is known almost the reasons that lead clients to cancel, or how it can be maintained a strategic distance from.

It is more than common in these days to employ computer programs and services to support management and decision taking. The contemporary algorithms and artificial intelligence (AI) techniques encouraged the involvement of automated decision support systems in every business-oriented information system, especially ERP programs, to analyze problems, detect factors, propose solutions, and help in improving or building new strategies using data-driven approaches.

One of the applications of these services is hotel management application. It started as invoices application which developed to include keeping the reservation information. Later, it added automated reservation services. After many reservations, the data can be analyzed to detect reservation problems' factors and expect results.

Cancelling reservation is one of the most common problems in hotel reservation services. Many expected factors affect this problem like the reservation location, the hotel rate, the price offer, and most importantly, the history of the client and his solemnity. The modern techniques can elicit the question of what these factors are based on data-driven analysis and data mining in the reservation records.

The current state of hotel management and reservation systems is interrelated and cooperative. Almost all reservation systems provide API for the third-party reservation systems which promote the offers and spread the popularity of the hotel. This means we have three main stakeholders in the reservation process: the hotel management, the external reservation promotional, and the client. Although the process is automated, it can be improved by implementing data mining and AI techniques to solve the major problems. We still do not have a concrete API that can evaluate each reservation process to control that process accordingly in the matter of price, payments, and priority.

This study aim to provide a technique that can detect the reasons behind reservation cancelling. These reasons are to be evaluated and weighted where they can be used later to evaluate reservations with similar parameter values. This technique is to be shaped in a ready to implement model that specify the inputs, processes, and output of factor detecting processes. The outputs of this technique is a weighted factor rule that will be added to the factors knowledge base to be checked in later reservations.

A. The Motivation

The importance of the study is derived from its function. Reservation imply the fact that the asset will not be usable until that reservation is over which prevent any exploit of that asset. This is acceptable as the expected outcome of that reservation is granted. However, cancelling the reservation prevents the usability of that asset and eliminate the outcome. In hotels, cancelling reservation have a higher impact as it eliminates the main purpose of the business which is renting the rooms. While there are many consequences for reservation canceling, it is better to avoid that cancel in the first place. We can avoid that cancel, or at least measure the risk of cancellation, by knowing the effective factors that results in cancellation. As knowing hotel cancellation factors can be imperative to make strides overbooking and cancellation approaches, which are two exceptionally critical themes in RM investigate [7].

B. The Contribution

This study aim to create an API that can provide a function to predict if a reservation is most likely to cancel or not. To do this, the study starts to define the factors. And the results that have been found of the factors are is repeated guest, previous cancellations, previous bookings not cancelled, required car parking spaces, and deposit type. The model using these factors show precision of 0.86 while the one that uses all the variables shows an average precision of 0.82.

C. The Paper Structure

First, the paper starts with introduction that introduce the problem and the background. Then, related work that present the current work. After that, the methodology followed by the results discussion. Finally, the conclusion for the paper.

II. BACKGROUND AND RELATED WORK

A. Hotel Cancellation Policies

As is known, all the hotels establish a price and conditions or policies for booking. Some of lodgings make the cancel-

lation access tough by limiting the free cancellation windows and/or by setting higher cancellation punishments. Some hotels put some cancellation fees and consider it as an other revenue [8]. In expansion, cancellation arrangements are outlined to influence travelers' booking behaviors in a way that's more alluring or profitable to inns. Furthermore, some hotels use debit or credit card which tends to decrease the possibility of no appears [9].

There are many indications that the searchers for a hotel continues to search for a suitable hotel even after they have booked a hotel in order to search for another hotel that has better policies and a lower price. In the event that a hotel is found that has suitable policies and has a lower price, the researcher cancels the reservation in the previous hotel and rebook the new one. Cancellation approaches exist for the reason to avoid such "switching" behavior [10].

In [11] inspected the effect of diverse sorts of cancellation confinements on lodging guests' reservation choices. Particularly, the consider explored how cancellation due dates and costs impact the vital booking behavior of deal-seeking travelers, advanced-booking. They found out the cancellation due date influenced participants' behavior whereas the estimate of the cancellation expense had no measurably critical affect. In expansion, there was no noteworthy difference between no cancellation approach and cancellation due date.

This considered in [12] that gives experimental prove on how lodging cancellation approaches are changing in later a long time. The discoveries demonstrate that whereas lodgings are testing with stricter cancellation windows, their cancellation punishments don't appear to end up stricter. There discoveries indicate that in spite of the fact that US inns have as of late fixed their free cancellation windows, there's no clear sign that the cancellation punishments have expanded. They found that together with the tightening of cancellation windows, there appears to be a move toward "standardizing" the cancellation punishments.

Indicated a cancellation policy means choosing the terms and conditions that allow booking to be cancelled, and penalization to the clients who cancel a booking or don't appear up at the concurred date. Such penalization is known to have a non-negligible part in income administration of inn chains. In [13] stated that inns that set extensive penalties for cancellation, especially with it gets closer to the arranged entry date point to play down revenue losses. For occurrence, it has become necessary to request reservations with credit card sponsorship, so that in the event that the customer does not show up, a fee for at least one night will be charged.

Lodging cancellation approaches not only affect Income but it also influences consumer behaviors. This consider [14] confirms the fact that big cities have more Free Cancellation Approaches. The reason is twofold. One reason is due to Competition within the markets in these cities. The other reason dues to the fact that small towns have few high-end hotels Which has permissive policies.

The discovery [14] shows that high-end inns tend to have less-strict approaches because maintaining great relationships with their customers is one of the goals of most high-end hotels. A later industry report demonstrates that the cancellation arrangement is conditional based on diverse variables

such as, the larger part of the sort of clients, the area, and the economy situation. For example, amid the subsidence, the tradition lodgings have tighter cancellation arrangements.

Utilizing the proposition-based theorizing within the setting of cancellation approaches, this study provides a few suggestions that seem have wide suggestions for future investigate. In [15], they introduce a theory about hotel cancellation policies that can Future research examine the effect of these policies. One of the policies that they introduce if the tightening cancellation policy might flag solid demand because the inn anticipates to exchange the room effectively notwithstanding of the restrictive policy or a strict cancellation approach might flag frail request because the lodging does not anticipate to exchange the room, and so endeavors to recoup its losses.

B. Predict Hotel Cancellation using Machine Learning

Inn income administration employments machine learning. The lodging industry endures 20% of income due to cancellation. In Addition, the administration cannot gadget strict arrangements of non-cancellation, as the guests/customers might switch to another inn. The machine learning-based estimating employments a neural network [16].

In [17] the authors proposed a forecasting model that forecast the cancellation in hotels. The forecasting model that they have been used was based on artificial intelligence by using personal name records (PNR). This investigate has been created utilizing genuine booking records provided by a lodging accomplice found in Gran Canaria (Spain) with the point of forecasting future cancellations. There approach for forecasting the cancellation in hotels was using 13 independent variables which is a decreased number in comparison with related investigate. For the result, they used three algorithms which were Random forest with 80% of accuracy, Support vector Machine with 75% and GA for ANN with 79%.

As the cancellation effects the reservation system and the income. In [18] they used a Personal Name Records (PNR) for predict hotel cancellation. Their approach differ from other research by aiming to distinguish those people likely to create cancellations in a short-horizon of time through PNR. Their promising comes about have been accomplished with 80% accuracy for cancellations made 7 days in progress. They used a PNR data from inn found in Gran Canaria (Spain). There method gave accuracy 73% for C5.0, SVM 71% ANN 69% Boosting ensemble 80%. This approach endeavors to forecast individual cancellations likely to be made exceptionally near to the section day, from 4 to 7 days in progress, and which can be considered "critical cancellations".

It is conceivable to construct models to anticipate bookings cancellation probability. These models has never been evaluated in a real environment. Consequently, in this paper [19] the author designed a prototype of the prediction model and transferred it to two hotels. First, they developed a trained model that learns from previous reservations to see a pattern of cancellations over time. From a commercial angle, the model demonstrated its applicability, being above 84% in accuracy, 82% in precision. The framework allowed the residences to predict their net demand, thus making better options for which reservations should be acknowledged and rejected, what construction costs would be, and how many rooms to sell.

As well as, reservation cancellations within the hospitality industry not as it were creating income misfortune and influence estimating and inventory allotment choices. Moreover, in overbooking circumstances, have the potential to influence the hotel's online social notoriety. In [20] authors used data from four resort hotel Property Management Systems (PMS) to predict hotel cancellation. The creators reach the accuracy of 90%.

Furthermore, to overcome the negative affect caused by overbooking and the usage of unbending cancellation approaches to manage with booking cancellations. This paper [21] the author used a real data from four resort hotels of the Algarve, Portugal. The author used Boosted Decision Tree, Decision Forest, Decision Jungle, Locally Deep Support Vector Machine, and Neural Network models to predict hotels reservations cancellation. the author got reached accuracy values above 90%, while models of H2 and H3 reaching 98.6% and 97.4%, respectively.

The online booking system in hotels is one of the most important alluring arrangements within the hospitality industry. Cancellation of inn bookings or reservations through the online framework is as of now one of the issues within the inn administration framework. In [22] the authors used a data set from the paper "Hotel Booking Demand Datasets". That been used XGBoost, Catboost, Light Gradient Boosting Machine, and Random Forest. By using CRISP-DM framework, they got about 0.8725 of accuracy in random forest.

Supervised anomaly detection concept is using to predict hotel cancellation. In [23] the authors used a few classification models such as Decision Tree, Gradient Boost, and XGBoost. Moreover, they reached high accuracy which is 86%. Their result appears that in their circumstance, the decision tree calculation, utilizing data sets with the accurately characterized properties, may be an incredible procedure for making prescient models for booking cancellations. These discoveries too affirm Chiang (2007) articulation "as modern trade models keep on rising, the ancient estimating strategies that worked well some time recently may not work well within the future. Confronting these challenges, analysts ought to proceed to create unused and superior estimating methods".

In this paper [24], they indicated the factors that influence the cancellation behavior. The factors that they found out are lead time, country, and season. The timing of booking is vital, with early bookings showing an altogether higher cancellation likelihood. Moreover, they found out that those who booked through agencies are more likely to not cancel their cancellation while those who booked through online or offline are more likely to cancel. The lead time and country effect the reservations which made through online more than offline. The cancellation probability indicated using cluster adjusted standard errors.

In [25] they use the data from eight hotels with few data from other sources (weather, holidays, events, social reputation, and online prices/inventory). Also, they developed a model to predict reservation cancellation in hotels. Moreover, they indicated cancellation drivers. For the drivers, they found out the most important features are country, deposit type, adults, Stays at Weekend Nights, Distribution Channel and Agent. They got 86% of accuracy for XGBoost model.

We still do not have a concrete API that can evaluate each reservation process to control that process accordingly in the matter of price, payments, and priority. This study aim to provide a technique that can detect the reasons behind reservation cancelling. These reasons are to be evaluated and weighted where they can be used later to evaluate reservations with similar parameter values. This technique is to be shaped in a ready to implement model that specify the inputs, processes, and output of factor detecting processes. The outputs of this technique is a weighted factor rule that will be added to the factors knowledge base to be checked in later reservations. In this study, the factors that influence the cancellation have been found are: is repeated guest, previous cancellations, previous bookings not cancelled, required car parking spaces, and deposit type. These factors evaluated using logistics regression. Later, these factors used in API function. Table I indicates the methodology and the results of related work.

C. Logistics Regression

Logistics regression is a statistical model utilizes to Logistic equation conditional probability. In addition, it uses to model a binary dependent variable. Logistic regression is used to predict the likelihood of an event occurring with additional knowledge of the values of variables that can be explained or for that event. This modeling is widely used in commercial and commercial transactions and is one of the modeling methods most applied in the field of learning, as it is classified within the methods of supervised learning.

In order to know the effectiveness of internal control systems to increase the revenue of hotels [26]. As to know the relationship between two variable, they used logistics regression. that been found out some internal control components have positive effect on revenue in hotels such as controlling activity, information, and communication and observing action have positive and critical impact in determining the results. Moreover, checking action contains a more prominent impact on the outcomes than data communication and control action.

Moreover, to know how auxiliary and organizational components impact hotel's probability of creating service/product, process, organizational and promoting advancements. In [27], the authors used responses from 174 hotels Chi-square test was utilized to test the primary theory. An arrangement of different logistic regression investigations was performed to test the connections between the independent variables. This paper gives bits of knowledge almost the nature and degree of advancements within the inn sector. In spite of the fact that generally considered inflexible and non-innovative, around half of the responding lodgings made at slightest one sort of improvement. Most common are service/product and displaying advancements. A hotel's probability of moving forward depends to a incredible degree on fundamental flexibility (non-chain), having an unequivocal advancement strategy.

III. DATA SET

The dataset used for this project is downloaded from Kaggle website. it contains a variety of hotel reservations information collected based on some potential factors. This data set contains booking information for city and resort hotels. These information such as the date of the reservation, length of

TABLE I. SUMMARY OF RELATED WORK

Citation	Methodology	Result
[17]	SVM and random forest	predict hotel cancellation with accuracy 80%
[18]	SVM, ANN, and Boosting ensemble	forecast individual cancellations likely to be made exceptionally near to the section day, from 4 day to 7 days in progress
[24]	probit model with cluster adjusted standard errors	The factors that they found out are lead time, country, and season
[19]	they designed a prototype of a prediction model and conveyed in with two hotels	the model illustrated its viability, with comes about surpassing 84% in accuracy
[25]	they indicated cancellation drivers. For the drivers, they found out the most important features are country, deposit type, adults, Stays At Weekend Nights, Distribution Channel and Agent	They got 86% of accuracy for XGBoost model
[20]	they used data from four resort hotel Property Management Systems (PMS) to predict hotel cancellation	it is possible to construct models for anticipating booking cancellations with exactness comes about in accuracy of 90%
[21]	they used Boosted Decision Tree, Decision Forest, Decision Jungle, Locally Deep Support Vector Machine, and Neural Network models	they got reached accuracy values above 90%, while models of H2 and H3 reaching 98.6% and 97.4%, respectively
[22]	They used XG-Boost, Catboost, Light Gradient Boosting Machine, and Random Forest	they got about 0.8725 of accuracy in random forest
[23]	they used a few classification models such as Decision Tree, Gradient Boost, and XG-Boost	they reached high accuracy which is 86%

stay, number of adults, children, and/or babies, and the number of available parking spaces, among other things. All personally identifying information has been removed from the data.

The dataset contains 32 columns described in their original project¹ as the following:

- hotel: Hotel (H1 = Resort Hotel or H2 = City Hotel)
- is_cancelled: Value indicating if the booking was cancelled (1) or not (0)
- lead_time: Number of days that between the entering

¹<https://github.com/rfordatascience/tidyuesday/blob/master/data/2020/2020-02-11/readme.md>

date of the booking into the PMS and the arrival date

- arrival_date_year: Year of arrival date
- arrival_date_month: Month of arrival date
- arrival_date_week_number: Week number of year for arrival date
- arrival_date_day_of_month: Day of arrival date
- stays_in_weekend_nights: Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
- stays_in_weeknights: Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel
- adults: Number of adults
- children: Number of children
- babies: Number of babies
- meal: Type of meal booked. Categories are presented in standard hospitality meal packages: Undefined/SC – no meal package; BB – Bed & Breakfast; HB – Half board (breakfast and one other meal – usually dinner); FB – Full board (breakfast, lunch and dinner)
- country: Country of origin. Categories are represented in the ISO 3155–3:2013 format
- market_segment: Market segment designation. In categories, the term “TA” means “Travel Agents” and “TO” means “Tour Operators”
- distribution_channel: Booking distribution channel. The term “TA” means “Travel Agents” and “TO” means “Tour Operators”
- is_repeated_guest: Value indicating if the booking name was from a repeated guest (1) or not (0)
- previous_cancellations: Number of previous bookings that were cancelled by the customer prior to the current booking
- previous_bookings_not_cancelled: Number of previous bookings not cancelled by the customer prior to the current booking
- reserved_room_type: Code of room type reserved. Code is presented instead of designation for anonymity reasons
- assigned_room_type: Code for the type of room assigned to the booking. Sometimes the assigned room type differs from the reserved room type due to hotel operation reasons (e.g. overbooking) or by customer request. Code is presented instead of designation for anonymity reasons
- booking_changes: Number of changes/amendments made to the booking from the moment the booking was entered on the PMS until the moment of check-in or cancellation
- deposit_type: Indication on if the customer made a deposit to guarantee the booking. This variable can

assume three categories: No Deposit – no deposit was made; Non Refund – a deposit was made in the value of the total stay cost; Refundable – a deposit was made with a value under the total cost of stay.

- agent: ID of the travel agency that made the booking
- company: ID of the company/entity that made the booking or responsible for paying the booking. ID is presented instead of designation for anonymity reasons
- days_in_waiting_list: Number of days the booking was in the waiting list before it was confirmed to the customer
- customer_type: Type of booking, assuming one of four categories: Contract - when the booking has an allotment or other type of contract associated to it; Group – when the booking is associated to a group; Transient – when the booking is not part of a group or contract, and is not associated to other transient booking; Transient-party – when the booking is transient, but is associated to at least other transient booking
- ADR: Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights
- required_car_parking_spaces: Number of car parking spaces required by the customer
- total_of_special_requests: Number of special requests made by the customer (e.g. twin bed or high floor)
- reservation_status: Reservation last status, assuming one of three categories: Cancelled – booking was cancelled by the customer; Check-Out – customer has checked in but already departed; No-Show – customer did not check-in and did inform the hotel of the reason why
- reservation_status_date: Date at which the last status was set. This variable can be used in conjunction with the Reservation Status to understand when the booking was cancelled or when did the customer checked-out of the hotel.

The second column is the dependent variable in this study, while the other columns represents the independent variables as potential factors.

A. Data Wrangling

Start with 31 proposed variables. These included timestamps, details, and status data. For the timestamps, It can be decided what kind of time variables that should be accepted by understanding the logical scope of the problem. In other words, how would time affect the hotel reservation? and what timestamp are we talking about? If that has been considered, there might be some yearly routine that during some period of the year there would be many cancellations, then it does not need to know which year exactly that has been talking about. In addition, this cannot be observed on daily bases. It is convenient to consider the period on weekly and monthly basis. For this, other timestamps can be dropped.

Now, by talking about dates, considering the arriving date in the reservation, not the reservation date, leaving date, nor status date. The other timestamps might be eliminated too. The details of the reservation include information about the hotel, agency, travelers, reservation type, average daily rate, and meals. Any variable considering this reservation is considered as a potential factor. The status is what we are trying to predict and analyze to find its factor, namely, the calculation status. For this, we can consider all other status as not cancelled. We do not need more details about the status. Based on the criteria above, we will drop the following columns:

- lead_time: Unneeded timestamp
- arrival_date_year: Unneeded timestamp
- arrival_date_day_of_month: Unneeded timestamp
- reservation_status: Unneeded status details
- reservation_status_date: Unneeded timestamp

There are missing data in the children field, country field, agent field, and company field. From the data description in the introduction it can be found that it might be gotten the types that needed from the market segment, distribution channel, and customer type and disregard the agent and company columns. So, these two are just dropped. For the children field and country field, it might be filling the missed data with 0 and 'Not' as mark that it was not inserted.

After cleaning, the dataset contains 33,775 duplicates which is more 30% of the dataset. However, these duplicates are not real duplicates. They might be different in term of year, day, or other removed fields. For this reason, no duplicates might be eliminated.

IV. METHODOLOGY

In order to answer the question of this study, the study starting by making an exploratory data analysis. Then, creating a list of the most effective factors. Finally, evaluating what the study has done by implementing prediction model based on the factors that will be found. The EDA includes having an insight of the statistics of the dataset. This insight will help in detecting if any bias occurs. After that the study will evaluate each factor using chi-square test of independence. To create the list of the effective factors, it will calculate the correlation coefficient between all the variables to understand the relationship between the variables. This process might propose eliminating some unrelated variables. Finally, the remaining factors will be used in creating prediction models, and compare these prediction models with same models which included all the eliminated data. Fig. 1 shows the stages for the methodology.



Fig. 1. Stages.

A. Data Analysis

The dataset has total bookings cancelled about 44,224 (37%). Bookings cancelled in resort hotel are 11,122 (28%). While bookings cancelled in a City hotel are 33,102 (42%). We can see the cancellation in the dataset more than 60% as Fig. 2 shows:

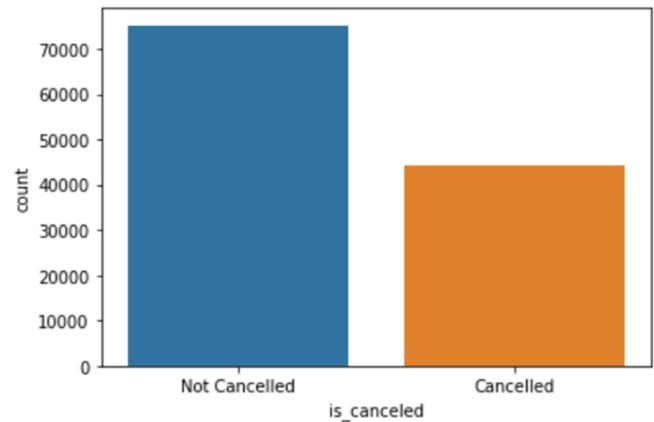


Fig. 2. Bookings Cancellation.

The question of this study is what are the main factors of hotel reservation cancelling? From this question it can be hypothetically derived a sub question for each of the other parameters. However, it can be disregarded some parameters logically or practically during data wrangling. To investigate the dataset statistics, the study starts with exploring the numerical variables. Starting with the numerical variables, the following Fig. 3 the histogram that has been found: There

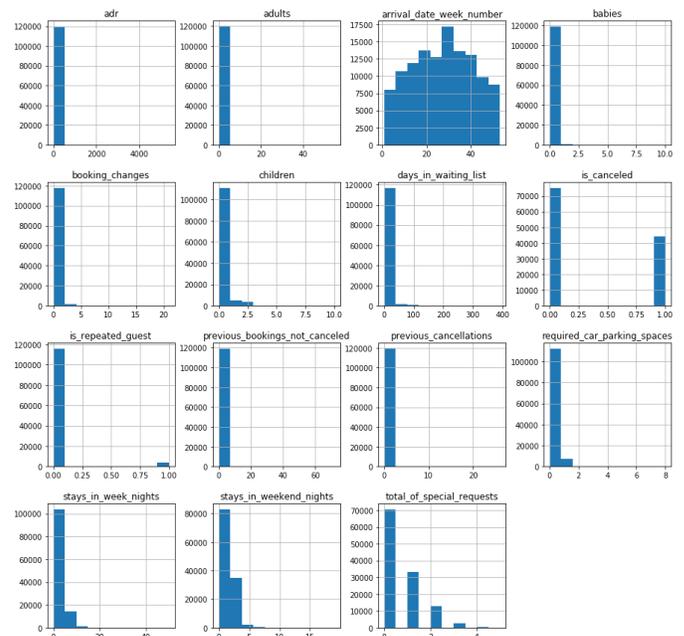


Fig. 3. Histogram of the Variable.

are 15 numerical variables including two of them are Boolean but treated as numerical which are: is_cancelled variable and

is_repeated_guest variable. Numerically, it seems that there are a lot of outliers. Logically, it is understandable that most of the reservation would have adults less than ten for example. it would not consider a reservation for group of 50 is an outlier. This also applies to the average daily rate, babies, and car parking spaces. However, for fields like the booking_changes, previous_booking_not_cancelled, previous_cancellations, stays_in_weeknights, and stays_in_weekend_nights, the concept is the important but not the quantity. This means we might look on the impact of changing booking on the reservation cancellation. This impact would be effective from the first and second change. No need for to look at those of 20 changes. It may combine all more than 5 changes as 5. In other words, the study will decrease the values more than the outlier threshold to the outlier threshold values in these fields. So, by calculating the Outliers based on quartile which is:

interquartile Range

$$IQR = Q3 - Q1$$

threshold range

$$\text{threshold_range} = IQR \times 1.5.$$

It can be seen from Table II the 1.5-based outlier threshold for most of the field is 0. This makes infeasible to use it. According to these results, it can be only used the threshold for stays_in_weekend_nights and stays_in_weeknights fields. For these two fields, we will reduce all outliers to the threshold.

TABLE II. CALCULATING OUTLIERS

Variables	Q1	Q3	IQR	Threshold range
arrival_date_week_number	16.00	38.0	22.00	33.000
stays_in_weekend_nights	0.00	2.0	2.00	3.000
stays_in_weeknights	1.00	3.0	2.00	3.000
adults	2.00	2.0	0.00	0.000
children	0.00	0.0	0.00	0.000
babies	0.00	0.0	0.00	0.000
is_repeated_guest	0.00	0.0	0.00	0.000
previous_cancellations	0.00	0.0	0.00	0.000
previous_bookings_not_cancelled	0.00	0.0	0.00	0.000
booking_changes	0.00	0.0	0.00	0.000
days_in_waiting_list	0.00	0.0	0.00	0.000
ADR	69.29	126.0	56.71	85.065
required_car_parking_spaces	0.00	0.0	0.00	0.000
total_of_special_requests	0.00	1.0	1.00	1.500

B. Factor Extraction

To extract the factors, the study uses chi-square test of independence. For nominated variables, correlation has been used to find the effective variables. After that another test has been added based on the coefficient of the variables in the linear regression.

C. Factors Evaluation

To evaluate the factors, two logistic regression models have been used. One to predict the cancellation based on all variables in the dataset. The other one predicts the cancellation based on the nominated factors. The factorizing process is considered efficient if the precision of the factors model revealed better performance than the all variables model.

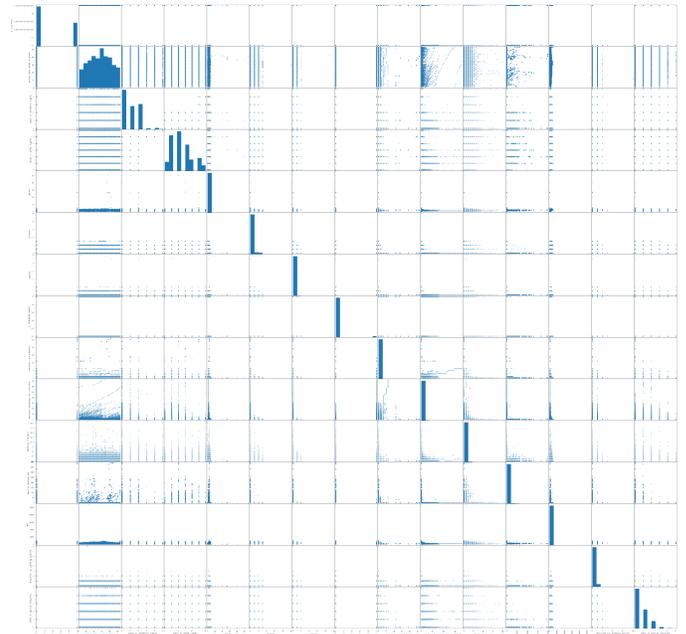


Fig. 4. Scatter Plot Matrix.

D. Scoring Function

by assuming that the cancellation process follows a linear equation, the intercept and coefficients have been used from the prediction model to create a scoring function that can calculate the sum of the factors multiplied by their coefficients in addition to the intercept. This score is to be evaluated as a probability later using the logistic function:

$$P(x) = \frac{e^x}{e^x + 1} \quad (1)$$

The result will give the probability of cancellation. If the probability is more than 0.5, the reservation process is more likely to cancel.

E. API Function

The API function is the last function that will use the scoring function to predict if a reservation would likely cancel or not. Its definition would be Boolean cancel (factorsArray). It will return true if the factors indicate probability of cancellations and false otherwise.

V. RESULTS DISCUSSION

For detecting the factors, it can be tried the scatter plot matrix which is used for seeing the connections between combinations of factors. Each scatter plot within the matrix visualizes the relationship between a combine of factors, permitting numerous relationships to be investigated in one chart. Therefore, in this project scatter plot matrix in Fig. 4 has been tried to have an insight about the relationship between the variables:

The scatter plot matrix initially suggests that there might be a relationship between the cancellation and adults' number, babies number. also, it seems that when there are previous cancellations, it is more likely it will cancel, while it might

not cancel with previous uncanceled reservations. There is also relationship with car parking and booking changes.

The study starts by initial evaluation for the factors using chi-square test of independence. The chi test is using for indicating if there is a relationship between tow variables based on the null hypothesis and alternative hypothesis. Null hypothesis assumes that there is no relationship between the two variables while the alternative hypothesis assume there is a relationship between the two variables. So, when the result from chi test less than alpha = 0.05, null hypothesis will be rejected and accept an alternative hypothesis. there are 10 categorical variables. To deal with them, they have been encoded into numerical values. The test shows the following results in Table III:

TABLE III. CHI TEST RESULT

Variable	Chi test result
0	2.78263008e-02
1	2.29104890e-02
2	1.06040302e-01
3	4.42317339e-01
4	3.20597577e-02
5	5.69348351e-02
6	2.22306546e-02
7	1.84643301e-02
8	1.68491108e-02
9	2.60087415e-02
10	1.45654133e-01
11	5.79706198e-02
12	4.51720898e-02
13	1.72514103e-02
14	1.79325050e-02
15	1.90153390e-02
16	4.20164259e-02
17	4.51832920e-02
18	5.28838342e-02
19	2.08257569e-02
20	4.67504818e-02
21	2.45671565e-02
22	2.36476367e+01
23	1.77620913e-02
24	2.77362630e-02

The variable number 1 is the ‘is cancelled’ field itself. In general, the test results indicate there is no significant dependency between arrival date month, arrival date week number, stays in week nights, country, market segment, booking changes, and ADR with is cancel variable.

Numerically, the study need to find the correlation between the numerical variables and the dependent variables. Correlation coefficient could be a factual degree of the quality of the relationship between the relative developments of two factors. In the following Table IV the correlation for numerical variable:

The correlation analysis reveals that there is significant correlation at alpha = 0.05 with the following columns: adults, is_repeated_guest, previous_cancellations, previous_bookings_not_cancelled, booking_changes, days_in_waiting_list, required_car_parking_spaces, and total_of_special_requests. At alpha = 0.01, stays_in_weeknights, babies, and ADR are added. The correlation results for the categorical variables in the following Table V:

Based on these results, we will consider the following variables as factors of reservation cancelling: adults, is repeated

TABLE IV. CORRELATION FOR NUMERICAL VARIABLES

Variables	Result
arrival_date_week_number	0.008148065395052901
stays_in_weekend_nights	-0.0017910780782611744
stays_in_weeknights	0.024764629045872715
adults	0.06001721283956815
children	0.005036254836439323
babies	-0.03249108920833264
is_repeated_guest	-0.0847934183570878
previous_cancellations	0.11013280822284255
previous_bookings_not_cancelled	-0.057357723165947075
booking_changes	-0.14438099106132224
days_in_waiting_list	0.054185824117780376
ADR	0.047556597880386124
required_car_parking_spaces	-0.1954978174945085
total_of_special_requests	-0.2346577739690198

TABLE V. CORRELATION FOR CATEGORICAL VARIABLES

Variables	Result
Hotel	0.13653126949161642
arrival_date_month	0.011822120071305441
meal	-0.01767760995132292
country	-0.10044912870002404
market_segment	0.23833549336078935
distribution_channel	0.1697270301121236
reserved_room_type	-0.04397743747112238
assigned_room_type	-0.1252105041585338
deposit_type	0.4804339866053031
customer_type	-0.13581931980513778

guest, previous cancellations, previous bookings not cancelled, booking changes, days in waiting list, required car parking spaces, total of special requests, hotel, country, market segment, distribution channel, assigned room type, deposit type, customer type.

A. Evaluation

The precision defines as the number of true positives divided by the number of true positives plus the number of false positives.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

For evaluation, two prediction model are created. The one that uses all the variables shows an average precision of 0.82 while the model used the factors shows average precision of 0.85. By exploring the coefficients in the following Table VI and eliminating the insignificant effective variables:

We remain with the following five factors: is_repeated_guest, previous_cancellations, previous_bookings_not_cancelled, required_car_parking_spaces, and deposit_type. The model using these factors show precision of 0.86. The factors are used to create a function that can evaluate these factors and predict if the reservation is more likely to cancel or not. To create this function a scoring function can be started. The intercept and coefficients are rounded to 2 decimal level and used in an equation model. The coefficient values are -0.84, 2.21, -0.76, -6.3, and 2.66. The score is then processed in a logistic function.

Based on the result of the logistic function, the cancel evaluation function returns true if the probability is more 0.5. Otherwise, it returns false. The following Table VII and Table

TABLE VI. COEFFICIENT FOR THE FACTORS

Factors	Coefficient
adults	-0.01
is_repeated_guest	-0.75
previous_cancellations	2.42
previous_bookings_not_cancelled	-0.76
booking_changes	-0.29
days_in_waiting_list	0.0
required_car_parking_spaces	-6.11
total_of_special_requests	-0.37
hotel	0.03
country	-0.0
market_segment	-0.15
distribution_channel	0.6
assigned_room_type	-0.04
deposit_type	2.61
customer_type	-0.37

VIII shows the result from the API function. Based on the result in Table VIII, it seems that there is a probability of cancelling the reservation about 83%. So, the cancel evaluation function gives a true for this reservation.

TABLE VII. EXPERIMENT 1 AFTER USING API

Scoring	7.157
API function result	0.00077
Cancel evaluation function	False

TABLE VIII. EXPERIMENT 2 AFTER USING API

Scoring	1.632
API function result	0.83
Cancel evaluation function	True

Finally, the cancel function can be interfaced in the requested form:
Boolean cancel (factorsArray)

VI. CONCLUSION AND FUTURE WORK

For a future work, it can be considered a data set from hotels in Saudi Arabia. Moreover, API function can be used with the data set that have the same parameters to evaluate it. Because of the unavailability of many data sets that can be used, only this data set was used in this study. Based on the results, the hotel management can consider a good cancellation policy. Furthermore, they can adjust their cancellation policy based on the proprieties that the API function has.

REFERENCES

- [1] A. Serra-Cantalops, D. D. Peña-Miranda, J. Ramón-Cardona, and O. Martorell-Cunill, "Progress in research on csr and the hotel industry (2006-2015)," *Cornell Hospitality Quarterly*, vol. 59, no. 1, pp. 15–38, 2018.
- [2] R. Battiti, M. Brunato, and F. Battiti, "Roomtetris: an optimal procedure for committing rooms to reservations in hotels," *Journal of Hospitality and Tourism Technology*, 2020.
- [3] M. Brunato and R. Battiti, "Combining intelligent heuristics with simulators in hotel revenue management," *Annals of mathematics and artificial intelligence*, vol. 88, no. 1, pp. 71–90, 2020.
- [4] T. Koide and H. Ishii, "The hotel yield management with two types of room prices, overbooking and cancellations," *International journal of production economics*, vol. 93, pp. 417–428, 2005.
- [5] M. Gellerstedt and T. Arvemo, "The impact of word of mouth when booking a hotel: could a good friend's opinion outweigh the online majority?," *Information Technology & Tourism*, vol. 21, no. 3, pp. 289–311, 2019.
- [6] S. Park, Y. Yin, and B.-G. Son, "Understanding of online hotel booking process: A multiple method approach," *Journal of Vacation Marketing*, vol. 25, no. 3, pp. 334–348, 2019.
- [7] N. António, "Predictive models of hotel booking cancellation: a semi-automated analysis of the literature," *Tourism & Management Studies*, vol. 15, no. 1, pp. 7–21, 2019.
- [8] T. A. Maier and C. Roberts Ph D, "Exploratory analysis of 'other revenue' impact on full and limited service hotel noi," *Perspectives in Asian Leisure and Tourism*, vol. 3, no. 1, p. 4, 2018.
- [9] C. Chen, "Cancellation policies in the hotel, airline and restaurant industries," *Journal of Revenue and Pricing Management*, vol. 15, no. 3, pp. 270–275, 2016.
- [10] F. DeKay, B. Yates, and R. S. Toh, "Non-performance penalties in the hotel industry," *International Journal of Hospitality Management*, vol. 23, no. 3, pp. 273–286, 2004.
- [11] C.-C. Chen, Z. Schwartz, and P. Vargas, "The search for the best deal: How hotel cancellation policies affect the search and booking decisions of deal-seeking customers," *International Journal of Hospitality Management*, vol. 30, no. 1, pp. 129–135, 2011.
- [12] A. Riasi, Z. Schwartz, and C.-C. Chen, "A paradigm shift in revenue management? the new landscape of hotel cancellation policies," *Journal of Revenue and Pricing Management*, vol. 18, no. 6, pp. 434–440, 2019.
- [13] B. Benítez-Aurioles, "Why are flexible booking policies priced negatively?," *Tourism Management*, vol. 67, pp. 312–325, 2018.
- [14] C.-C. Chen and K. L. Xie, "Differentiation of cancellation policies in the us hotel industry," *International Journal of Hospitality Management*, vol. 34, pp. 66–72, 2013.
- [15] A. Riasi, Z. Schwartz, and C.-C. Chen, "A proposition-based theorizing approach to hotel cancellation practices research," *International Journal of Contemporary Hospitality Management*, 2018.
- [16] E. Alotaibi, "Application of machine learning in the hotel industry: A critical review," *Journal of Association of Arab Universities for Tourism and Hospitality*, vol. 18, no. 3, pp. 78–96, 2020.
- [17] A. J. Sánchez-Medina, C. Eleazar, *et al.*, "Using machine learning and big data for efficient forecasting of hotel booking cancellations," *International Journal of Hospitality Management*, vol. 89, p. 102546, 2020.
- [18] E. C. Sánchez, A. J. Sánchez-Medina, and M. Pellejero, "Identifying critical hotel cancellations using artificial intelligence," *Tourism Management Perspectives*, vol. 35, p. 100718, 2020.
- [19] N. Antonio, A. de Almeida, and L. Nunes, "An automated machine learning based decision support system to predict hotel booking cancellations," *An automated machine learning based decision support system to predict hotel booking cancellations*, no. 1, pp. 1–20, 2019.
- [20] N. António, A. de Almeida, and L. M. Nunes, "Using data science to predict hotel booking cancellations," in *Handbook of research on holistic optimization techniques in the hospitality, tourism, and travel industry*, pp. 141–167, IGI Global, 2017.
- [21] N. Antonio, A. De Almeida, and L. Nunes, "Predicting hotel booking cancellations to decrease uncertainty and increase revenue," *Tourism & Management Studies*, vol. 13, no. 2, pp. 25–39, 2017.
- [22] Z. A. Andriawan, S. R. Purnama, A. S. Darmawan, A. Wibowo, A. Sugiharto, F. Wijayanto, *et al.*, "Prediction of hotel booking cancellation using crisp-dm," in *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)*, pp. 1–6, IEEE, 2020.
- [23] C. Timamopoulos, "Anomaly detection: Predicting hotel booking cancellations," 2020.
- [24] M. Falk and M. Vieru, "Modelling the cancellation behaviour of hotel guests," *International Journal of Contemporary Hospitality Management*, 2018.
- [25] N. Antonio, A. de Almeida, and L. Nunes, "Big data in hotel revenue management: Exploring cancellation drivers to gain insights into booking cancellation behavior," *Cornell Hospitality Quarterly*, vol. 60, no. 4, pp. 298–319, 2019.

- [26] M. Yemer, "The effect of internal controls systems on hotels revenue. a case of hotels in bahir dar and gondar cities," *Arabian Journal of Business and Management Review (Oman Chapter)*, vol. 6, no. 6, p. 19, 2017.
- [27] W. Wikhamn, J. Armbricht, and B. R. Wikhamn, "Innovation in swedish hotels," *International Journal of Contemporary Hospitality Management*, 2018.

A Technique for Constrained Optimization of Cross-ply Laminates using a New Variant of Genetic Algorithm

Huiyao Zhang¹, Atsushi Yokoyama²
Department of Fiber Science and Engineering
Kyoto Institute of Technology,
Kyoto, JAPAN

Abstract—The main challenge presented by the design of laminated composite material is the laminate layout, involving a set of fiber orientations, composite material systems, and stacking sequences. In nature, it is a combinatorial optimization problem with constraints that can be solved by the genetic algorithm. The traditional approach to solve a constrained problem is reformulating the objective function. In the present study, a new variant of the genetic algorithm is proposed for the design of composite material by using a mix of selection strategies, instead of modifying the objective function. To check the feasibility of a laminate subject to in-plane loading, the effect of the fiber orientation angles and material components on the first ply failure is studied. The algorithm has been validated by successfully optimizing the design of cross-ply laminate under different in-plane loading cases. The results obtained by this algorithm are better than works in related literature.

Keywords—Laminated composite; classical lamination theory; genetic algorithm; optimal design

I. INTRODUCTION

Composite materials offer improved strength, stiffness, fatigue, corrosion resistance, etc. over conventional materials, and are widely used as materials for applications ranging from the automotive to shipbuilding industry, electronic packaging to golf clubs, and medical equipment to homebuilding. However, the high cost of fabrication of composites is a critical drawback to its application. For example, the graphite/epoxy composite part may cost as much as 650 to 900 per kilogram. In contrast, the price of glass/epoxy is about 2.5 times less. Manufacturing techniques such as sheet molding compounds and structural reinforcement injection molding are used to lower the costs for manufacturing automobile parts. An alternative approach is using hybrid composite materials.

The mechanical performance of a laminate composite is affected by a wide range of factors such as the thickness, material, and orientation of each lamina. Because of manufacturing limitations, all these variables are usually limited to a small set of discrete values. For example, the ply thickness is fixed, and ply orientation angles are limited to a set of angles such as 0, 45, and 90 degrees in practice. So the search process for the optimal design is a discrete optimization problem that can be solved by the GA. To tailor a laminate composite, the GA has been successfully applied to solve laminate design problems [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11]. The GA simulates the process of natural evolution, including selection, crossover, and mutation according to Darwin's principle of "survival of the fittest". The known advantages of GAs are the following: (i) GAs are not easily trapped in local optima and can obtain the global optimum. (ii) GAs do not need

gradient information and can be applied to discrete optimization problems. (iii) GAs can not only find the optimal value in the domain but also maintain a set of optimal solutions. However, the GA also has some disadvantages, for example, the GA needs to evaluate the target functions many times to achieve optimization, and the cost of the search process is high. The GA consists of some basic parts, the coding of the design variable, the selection strategy, the crossover operator, the mutation operator, and how to deal with constraints. For the variable design part, there are two methods to deal with the representation of design variables, namely, binary string and real value representation [1], [4], [12], [13]. Michalewicz [14] claimed that the performance of floating-point representation was better than binary representation in the numerical optimization problem. Selection strategy plays a critical role in the GA, which determines the convergence speed and the diversity of the population. To improve search ability and reduce search costs, various selection methods have been invented, and they can be divided into four classes: proportionate reproduction, ranking, tournament, and genitor(or "steady state") selection. In the optimization of laminate composite design, the roulette wheel [1], [15], where the possibility of an individual to be chosen for the next generation is proportional to the fitness. Soremekun et al. [16] showed that the generalized elitist strategy outperformed a single individual elitism in some special cases.

The data structure, repair strategies, and penalty functions [17] are the most commonly used approaches to resolve constrained problems in the optimization of composite structures. Symmetric laminates are widely used in practical scenarios, and data structures can be used to fulfill symmetry constraints, which consists of coding half of the laminate and considering the rest with the opposite orientation. Todoroki [4] introduced a repair strategy that can scan the chromosome and repair the gene on the chromosome if it does not satisfy the contiguity constraint. The comparison of repair strategies in a permutation GA with the same orientation was presented by Liu et al. [5], and it showed that the Baldwinian repair strategy can substantially reduce the cost of constrained optimization. Haftka and Todoroki [1] used the GA to solve the laminate stacking sequence problem using a penalty function subject to buckling and strength constraints.

In typical engineering applications, composite materials are under very complicated loading conditions, not only in-plane loading but also out-of-plane loading. Most of the studies on the optimization of the laminate composite material minimized the thickness [18], [7], weight [19], [20], [21], and cost and weight [20], [22], or maximized the static strength of the composite laminates for a targeted thickness [7], [8], [23],

[24]. In the present study, the cost and weight of laminates are minimized by modifying the objective function.

To check the feasibility of a laminate composite by imposing a strength constraint, failure analysis of a laminate is performed by applying suitable failure criteria. The failure criteria of laminated composites can be classified into three classes: non-interactive theories (e.g. maximum strain), interactive theories (e.g. Tsai-wu), and partially interactive theories (e.g. Puck failure criterion). Previous researchers adopted the first-ply-failure approach using Tsai-wu failure theory [25], [26], [19], [27], [28], [29], [22], [30], Tsai-Hill [31], [32], the maximum stress [33], or the maximum strain [33] static failure criteria. Akbulut [11] used the GA to minimize the thickness of composite laminates with Tsai-Hill and maximum stress failure criteria, and the advantage of this method is it avoids spurious optima. Naik et al. [34] minimized the weight of laminated composites under restrictions with a failure mechanism-based criterion based on the maximum strain and Tsai-wu. In the present study, Tsai-wu Static failure criteria are used to investigate the feasibility of a laminate composite.

II. CLASSICAL LAMINATION THEORY

A laminate structure consists of multiple laminae bonded together through their thickness. Considering a laminate composite plate that is subject to in-plane loading of extension, shear, bending, and torsion, the classical lamination theory (CLT) is taken to calculate the stress and strain in the local and global axes of each ply, as shown in Fig. 1. Based on fiber orientation, material, and fiber thickness, there are a few special cases of laminate: the set of fiber angles in Fig. 2 only includes 0 and 90, which is called cross-ply laminate.

A. Stress and Strain in Lamina

For a single lamina, the stress-strain relation in local axis 1-2 is:

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{12} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{bmatrix}, \quad (1)$$

where Q_{ij} are the stiffnesses of the lamina that are related

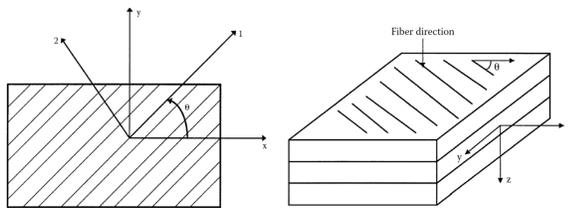


Fig. 1. Local and Global Axes of an Angle Lamina.

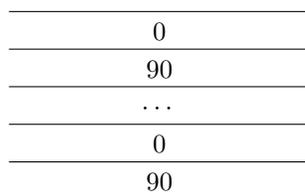


Fig. 2. Model for Cross-ply Laminate.

to engineering elastic constants given by

$$\begin{aligned} Q_{11} &= \frac{E_1}{1 - \nu_{12}\nu_{21}}, \\ Q_{22} &= \frac{E_2}{1 - \nu_{12}\nu_{21}}, \\ Q_{66} &= G_{12}, \\ Q_{12} &= \frac{\nu_{21}E_2}{1 - \nu_{12}\nu_{21}}, \end{aligned} \quad (2)$$

where $E_1, E_2, \nu_{12}, G_{12}$ are four independent engineering elastic constants, which are defined as follows: E_1 is the longitudinal Young's modulus, E_2 is the transverse Young's modulus, ν_{12} is the major Poisson's ratio, and G_{12} is the in-plane shear modulus.

Stress strain relation in the global x-y axis:

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} & \bar{Q}_{16} \\ \bar{Q}_{12} & \bar{Q}_{22} & \bar{Q}_{26} \\ \bar{Q}_{16} & \bar{Q}_{26} & \bar{Q}_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix}, \quad (3)$$

where

$$\begin{aligned} \bar{Q}_{11} &= Q_{11}c^4 + Q_{22}s^4 + 2(Q_{12} + 2Q_{66})s^2c^2, \\ \bar{Q}_{12} &= (Q_{11} + Q_{22} - 4Q_{66})s^2c^2 + Q_{12}(c^4 + s^4), \\ \bar{Q}_{22} &= Q_{11}s^4 + Q_{22}c^4 + 2(Q_{12} + 2Q_{66})s^2c^2, \\ \bar{Q}_{16} &= (Q_{11} - Q_{12} - 2Q_{66})c^3s - (Q_{22} - Q_{12} - 2Q_{66})s^3c, \\ \bar{Q}_{26} &= (Q_{11} - Q_{12} - 2Q_{66})cs^3 - (Q_{22} - Q_{12} - 2Q_{66})c^3s, \\ \bar{Q}_{66} &= (Q_{11} + Q_{22} - 2Q_{12} - 2Q_{66})s^2c^2 + Q_{66}(s^4 + c^4). \end{aligned} \quad (4)$$

The c and s denote $\cos\theta$ and $\sin\theta$, respectively.

The local and global stresses in an angle lamina are related to each other through the angle of the lamina θ

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix} = [T] \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} \quad (5)$$

where

$$[T] = \begin{bmatrix} c^2 & s^2 & 2sc \\ s^2 & c^2 & -2sc \\ -sc & sc & c^2 - s^2 \end{bmatrix}. \quad (6)$$

B. Stress and Strain in a Laminate

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{16} \\ A_{12} & A_{22} & A_{26} \\ A_{16} & A_{26} & A_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^0 \\ \varepsilon_y^0 \\ \gamma_{xy}^0 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{16} \\ B_{11} & B_{12} & B_{16} \\ B_{16} & B_{26} & B_{66} \end{bmatrix} \begin{bmatrix} k_x \\ k_y \\ k_{xy} \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{16} \\ B_{12} & B_{22} & B_{26} \\ B_{16} & B_{26} & B_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^0 \\ \varepsilon_y^0 \\ \gamma_{xy}^0 \end{bmatrix} + \begin{bmatrix} D_{11} & D_{12} & D_{16} \\ D_{11} & D_{12} & D_{16} \\ D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} k_x \\ k_y \\ k_{xy} \end{bmatrix}$$

N_x, N_y - normal force per unit length

N_{xy} - shear force per unit length
 M_x, M_y - bending moment per unit length
 M_{xy} - twisting moments per unit length
 ε^0, k - mid-plane strains and curvature of a laminate in x-y coordinates

The mid-plane strain and curvature is given by

$$\begin{aligned} A_{ij} &= \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k - h_{k-1}) i = 1, 2, 6, j = 1, 2, 6, \\ B_{ij} &= \frac{1}{2} \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k^2 - h_{k-1}^2) i = 1, 2, 6, j = 1, 2, 6, \\ D_{ij} &= \frac{1}{3} \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k^3 - h_{k-1}^3) i = 1, 2, 6, j = 1, 2, 6, \end{aligned} \quad (8)$$

where the [A], [B], and [D] matrices are called the extensional, coupling, and bending stiffness matrices.

III. FAILURE THEORY

A. Failure Process

A laminate will fail under increasing mechanical loading; however, the procedure of laminate failure may not be catastrophic. In some cases, some layers fail first, and the rest are able to continue to take additional loading until all the plies fail. A ply is fully discounted when a ply fails; then, the ply is replaced by a near-zero stiffness and strength. The procedure for finding the first ply failure in the present study follows the fully discounted method:

- 1) Compute the reduced stiffness matrix [Q] referred to as the local axis for each ply using its four engineering elastic constants E_1, E_2, E_{12} , and G_{12} .
- 2) Calculate the transformed reduced stiffness $[\bar{Q}]$ referring to the global coordinate system (x, y) using the reduced stiffness matrix [Q] obtained in step 1 and the ply angle for each layer.
- 3) Given the thickness and location of each layer, the three laminate stiffness matrices [A], [B], and [D] are determined.
- 4) Apply the forces and moments, $[N]_{xy}, [M]_{xy}$ solve Equation 7, and calculate the middle plane strain $[\sigma^0]_{xy}$ and curvature $[k]_{xy}$.
- 5) Determine the local strain and stress of each layer under the applied load.
- 6) Use the ply-by-ply stress-strain and related failure theories to determine the strength ratio.

B. Tsai-wu Failure Theory

Many different theories about the failure of an angle lamina have been developed for a unidirectional lamina, such as the maximum stress failure theory, maximum strain failure theory, Tsai-Hill failure theory, and Tsai-Wu failure theory. The failure theories of a lamina are based on the stresses in the local axes in the material. There are four normal strength parameters and one shear stress for a unidirectional lamina. The five strength parameters are:

$$\begin{aligned} (\sigma_1^T)_{ult} &= \text{ultimate longitudinal tensile strength} \\ (\sigma_1^C)_{ult} &= \text{ultimate longitudinal compressive strength} \end{aligned}$$

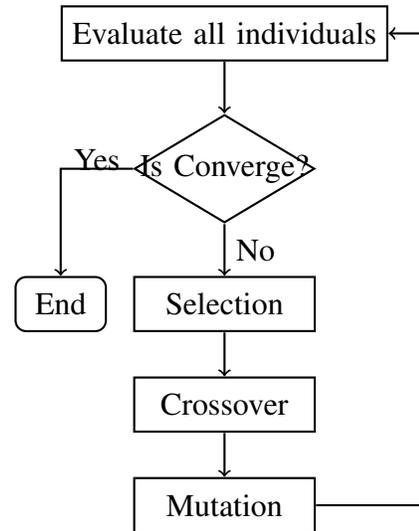


Fig. 3. Traditional GA Model.

$(\sigma_2^T)_{ult}$ = ultimate transverse tensile strength

$(\sigma_2^C)_{ult}$ = ultimate transverse compressive strength

$(\tau_{12})_{ult}$ = and ultimate in-plane shear strength

In the present study, Tsai-wu failure theory is taken to decide whether a lamina fails, because this theory is more general than the Tsai-Hill failure theory, which considers two different situations, the compression and tensile strengths of a lamina. A lamina is considered to fail if

$$\begin{aligned} H_1\sigma_1 + H_2\sigma_2 + H_6\tau_{12} + H_{11}\sigma_1^2 + H_{22}\sigma_2^2 \\ + H_{66}\tau_{12}^2 + 2H_{12}\sigma_1\sigma_2 < 1 \end{aligned} \quad (9)$$

is violated, where

$$SR = \frac{\text{Maximum Load}}{\text{Load Applied}} \quad (10)$$

The maximum load refers to that can be applied using Tsai-wu failure theory.

IV. GENETIC ALGORITHM MODEL

The genetic algorithm starts with multiple individuals with limited chromosome lengths, in which maybe none of these individuals fulfill the constraints. The GA is assumed to derive appropriate offspring based on the initial population as the GA continues. The traditional way to handle the constrained search of the GA is either to introduce repair strategies or to use a penalty function. Fig. 3 shows the classic flow chart of a GA framework, which includes selection, crossover, and mutation operators. However, GA is originally proposed to solve unconstrained problems; therefore, we suggest a new approach to address the constrained GA search problem in an unconstrained way.

Because of the existence of constraints, the population not only can be sorted by the fitness (obtained by the objective function) but also sorted by the constraint value obtained by the constraint function (assuming a constraint function exists), so the parents of the next generation can be chosen by the

following three approaches. First, the population is sorted by fitness in ascending order, and individuals with smaller fitness are selected. These selected individuals form a group named as a proper group. Second, remove individual which satisfies constraints, and sort population by the difference between the individual's constraint value and the threshold of the constraint in descending order, and individuals with bigger differences are chosen to be the parents of the next generation. The group which forms are called potential group, and an individual from this group is referred to as a potential individual. Third, the population is sorted by fitness from low to high after removing individuals which fail to fit the constraints, select individuals with bigger fitness, and these individuals form the proper group. So the final parents' pool consists of three groups, active group, potential group, and proper group. The number of active individuals, potential individuals, and proper individuals are called, respectively, active number, potential numbers, and proper number.

Each group in the parents' population has its role in the searching process. The problem within traditional GA is premature and has weak local search ability, therefore, traditional GAs are more likely to get stuck in a local optimum. To prevent the GA from experiencing early convergence and to improve the local search performance, the active group is proposed to overcome this problem. As its name suggests, this group would always live in the population. Because both active individual's fitness and constraint value are small, each individual can be treated as an independent gene clip. So their existence enriches the gene clip variety of the mating pool. The offspring of two active individuals are more likely to be active individuals, which can maintain the active group.

For an individual in the potential group, it doesn't satisfy the constraint, however, it's supposed to evolve into a proper individual after multiple generations by modifying its chromosome structure or length. The crossover operation could happen between a potential individual with an active individual, or a potential individual, or a proper individual. The child of an active individual and a potential individual is more likely to be a potential individual, and this active individual could inject a new gene clip into this potential individual, therefore providing a new evolution direction.

A proper individual means a feasible solution, which fulfills all the constraints. However, there are still some drawbacks within it, for example, its fitness is low. The crossover operation could happen between a proper individual and any other individuals.

The mutation operator for an active group is different from the potential group and proper group because their roles in the searching process are different: the target of the potential group and proper group is to obtain a feasible solution; however, the role of the active group is to maintain the variety of gene clips in the mating pool.

Fig. 4 shows the flow chart of the proposed GA. First, the population is divided into three groups, active group, potential group, and proper group by the above-mentioned method. Second, select an appropriate number of individuals from each group as parents, and the various selection scheme can be taken for each group.

The searching process can be divided into two phases according to whether proper individuals are generated or not. During the initial stage, no individual in the population is appropriate, which means the number of individuals in the proper group is zero. Both active group and potential group

are full. After a couple of generations, some proper individuals could be produced. Then, the GA comes to its second phase, the number of proper individuals begins to increase, finally, the number in the proper group reaches its upper bound.

V. EXPERIMENT

First, we formulate a constrained problem by searching the optimal stacking sequence of cross-ply laminate under in-plane loading under the constraint whose strength ratio is not less than two. Each lamina dimensions $1000 \times 1000 \times 0.165mm^3$ is adopted in this experiment, each graphite/epoxy, and glass/epoxy layer is assumed to cost 2.5 and 1 monetary unit, respectively. The other material properties are shown in Table I.

TABLE I. COMPARISON OF THE GRAPHITE/EPOXY AND GLASS/EPOXY PROPERTIES

Property	Symbol	Unit	Graphite/Epoxy	Glass/Epoxy
Longitudinal elastic modulus	E_1	GPa	181	38.6
Transverse elastic modulus	E_2	GPa	10.3	8.27
Major Poisson's ratio	ν_{12}		0.28	0.26
Shear modulus	G_{12}	GPa	7.17	4.14
Ultimate longitudinal tensile strength	$(\sigma_1^T)_{ult}$	MPa	1500	1062
Ultimate longitudinal compressive strength	$(\sigma_1^C)_{ult}$	MPa	1500	610
Ultimate transverse tensile strength	$(\sigma_2^T)_{ult}$	MPa	40	31
Ultimate transverse compressive strength	$(\sigma_2^C)_{ult}$	MPa	246	118
Ultimate in-plane shear strength	$(\tau_{12})_{ult}$	MPa	68	72
Density	ρ	g/cm^3	1.590	1.903
Cost			2.5	1

A. Problem Formulation

In the present experiment, the optimal composite sequences, and the number of layers for a targeted strength

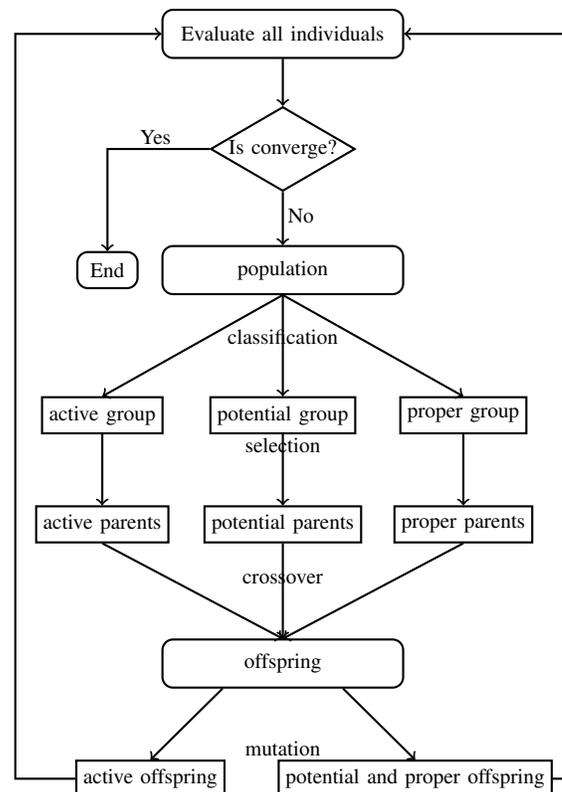


Fig. 4. General Flowchart of Proposed GA Model in which the Parents consist of three Different Groups.

ratio under in-plane loading conditions are investigated. The aim is to minimize the mass of a laminate composite for a targeted strength ratio based on the Tsai-wu failure theory. The design variables are the ply angles and the number of layers. Ply orientation restricted to a discrete set of angles (0, and 90 degrees). The problem can be formulated as the following equation:

Find: $\{\theta_k, n\} \theta_k \in \{0, 90\}$,
 Minimize: weight,
 Subject to: strength ratio.

B. GA Operation

In the present study, floating encoding is adopted to represent a solution for the layup design of cross-ply laminate, Fig. 5(a) shows two parents P_1 and P_2 represent two cross-ply laminates, the corresponding laminates layups are $[0_3/90_7]$ and $[0_6/90_4]$, respectively; Fig. 5(b) shows two offspring of parents P_1 and P_2 are consist of half of each parent’s chromosome; Fig. 5(c) and 5(d) display the offspring after length and angle mutation, respectively.

For the length mutation of chromosome, calculate the chromosome’s strength ratio based on its sequence, if its strength ratio is less than the threshold, then increase its length; otherwise, reduce it. We introduce the term length mutation coefficient to control the length mutation. As shown in Fig. 5(b), the strength ratio of O_1 chromosome is 0.0854, and the strength ratio threshold is five. Suppose the length mutation coefficient takes two, then the corresponding increase length is the multiplication result of length mutation coefficient and the difference between current strength ratio and the threshold: the result is $5 \times (2 - 0.0854)$, round this number to its closest integer, which is 9. So the length of offspring O_1 changes from 10 to 19 after length mutation. For the angle mutation, randomly swap the gene from 0 to 90 in the chromosome, or the otherwise.

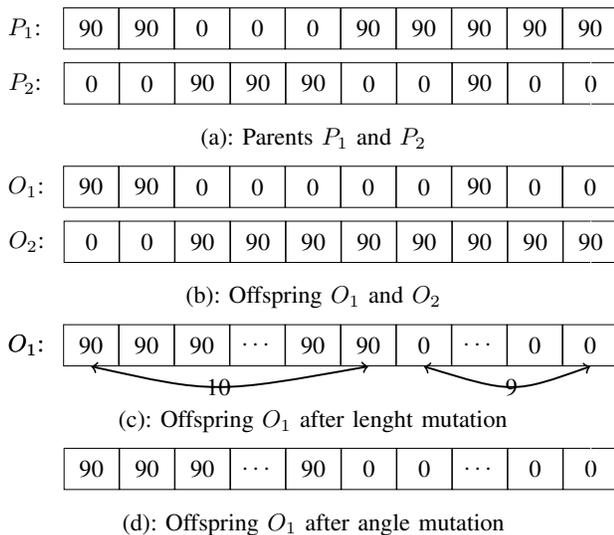


Fig. 5. Examples of Crossover, Length Mutation, Angle Mutation Operator for Proposed GA.

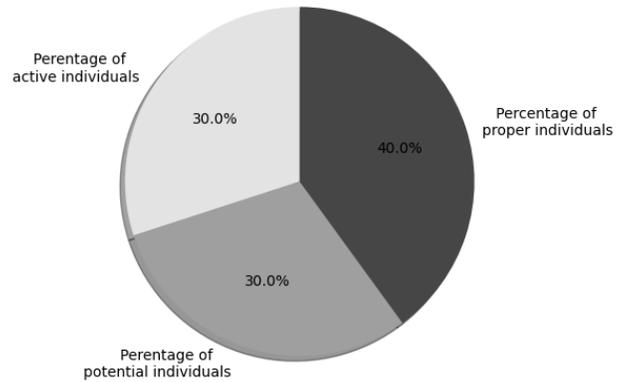


Fig. 6. Percentage of Active Individuals from Active Group, Potential Individuals from Potential Group, and Proper Individuals from Proper Group in Parent Population.

C. GA Parameters

Table II shows related GA parameters: the population is 40, and 50 percent is as the mating pool, so the parent population is 20; as shown in Fig. 6, the percentage of active individuals from the active group, potential individuals from the potential group, and proper individuals from the proper group are 0.3, 0.3, and 0.4, which means the corresponding number of these three types of individuals are 6, 6, and 8, respectively.

VI. NUMERICAL RESULTS AND DISCUSSION

To figure out how the number of individuals in each group varies during the GA process, we conduct a one-time experiment and show the number of individuals in each generation with respect of GA generation. Second, to verify its performance and stability, the GA is run one hundred times: the best, worst case, and average results are presented, respectively. Finally, we compare the results with the work in the other literature.

TABLE II. PARAMETERS OF PROPOSED GA MODEL

Parameter	Value
Population	40
Initial length range	[3-15]
Encoding	Integer
Percentage of parent	0.5
Percentage of active group	0.3
Percentage of potential group	0.3
Percentage of proper group	0.4
Selection strategy for active group	Ranking
Selection strategy for potential group	Ranking
Selection strategy for proper group	Ranking
Crossover strategy	One-point
Mutation strategy	Mass mutation
Length mutation coefficient	5
Angle mutation rate	0.1
Crossover rate	0.3

Fig. 7 shows the number of individuals in each group during the one-time GA process. For both active group and potential group, the number of individuals is fixed, and equal to its upper bound from the beginning to the end of the searching process. However, for the proper group, at the initial stage of GA, no individual fulfills the constraint, so the number of proper individuals is zero. As seen from Fig. 7, after forty generations, proper individuals appear, and its population increases very quickly up to its upper bound.

There are two approaches that the GA could obtain a better solution: the first is increasing the length of the chromosome; the second one is adjusting the internal structure of a chromosome. The GA process can be divided into two phases by whether there are proper individuals or not. Fig. 8 shows the GA process in which the dashed vertical line is the watershed between the initial phase and the last phase. In the initial phase, no individual's strength ratio is over the specified threshold, and the main reason that the fitness gets smaller gradually is the increase of chromosome's length; At point 1 on the fitness curve, the fitness suddenly goes up, however, the corresponding strength ratio of point 1, denoted by the point 1' on the generation-strength ratio curve, also increases. it is because of the adjustment of a chromosome's layup. Then GA comes to its second phase. During this phase, the GA already found proper individuals which could satisfy the constraint, so the target in this stage is to improve fitness. This means GA needs to adjust its inner structure, at the point 2 and 3 on the generation-fitness curve, the fitness curves go up, and the corresponding strength ratio of these two points slightly go down, but both of them still satisfy the constraint.

Table III shows the searching results after conducting this experiment one hundred times in two length mutation coefficient cases for glass-epoxy and graphite-epoxy material, respectively. The best, worst case, and average experiment results are showed in this table. For the glass-epoxy material, if the length mutation coefficient takes one, the best and worst sequences are $[0_{40}/90_{26}]_s$, $[90_{24}/0_{38}/90]_s$; the average mass, cost, and the number of layers are 7.83, 123, 123. If we increase its length mutation coefficient, suppose it is five, the number of layers for best and worst cases are 125 and 136; the average mass, cost, and number of layers are 8.55, 131, 131. When graphite-epoxy is taken as the experiment

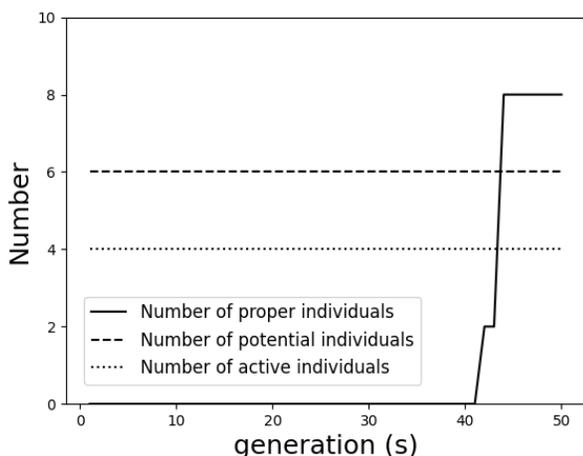


Fig. 7. Number of Individuals in Each Group as a Function of Generation.

b

TABLE III. THE OPTIMUM LAYUP FOR THE LOADING $N_x = 1e^6$ N WHEN CHANGING THE LENGTH MUTATION COEFFICIENT, THE PERFORMANCE OF THE GA CAN BE IMPROVED WHEN THE LENGTH MUTATION COEFFICIENT IS SMALLER.

Length mutation coefficient	Material	case	Stacking sequence	Strength ratio	Mass	Cost	Layer
1	glass-epoxy	worst	$[0_{40}/90_{26}]_s$	2.010	8.58	132	132
		best	$[90_{24}/0_{38}/90]_s$	2.078	8.12	125	125
		average		2.012	7.83	123	123
	graphite-epoxy	worst	$[0_{09}/90_{14}/0]_s$	2.17	1.41	68	27
		best	$[0_{09}/90_{11}/0]_s$	2.15	1.10	53	21
		average		2.018	1.47	70	28
5	glass-epoxy	worst	$[0_{36}/90_{32}]_s$	2.009	8.84	136	136
		best	$[0_{36}/90_{26}/90]_s$	2.003	8.12	125	125
		average		2.008	8.55	131	131
	graphite-epoxy	worst	$[0_{09}/90_{12}]_s$	2.006	2.20	105	42
		best	$[0_{08}/90_{3}/0]_s$	2.001	1.20	57	23
		average		2.022	1.54	73	29

material, similar experiment results are found. Comparing these two results, we see that a bigger value of length mutation coefficient can improve this GA's performance. This is because the mutation coefficient can control both the convergence speed and search performance, a small mutation coefficient would slow the convergence speed, however, it would lead to small-grained exploitation in the local space.

Table IV shows the optimal cross-ply sequences by the proposed GA and Choudhury and Mondal's [30] study. For the loading case $N_x = 1$ MPa m, the optimal layups are a $[0_{68}/90_{72}]$ cross-ply laminate if glass-epoxy is taken; however, in the present study, a $[90_{24}/0_{38}/90]_s$ glass-epoxy cross-ply laminate is found which significantly reduces both the cost and weight, and it satisfies the constraint. Similarly, if graphite-epoxy is taken, compared with the $[0_{17}/90_{18}]$ cross-ply laminate, an alternative solution is found, its layup is $[0_{09}/90_{11}/0]_s$. For both cases, we can see that the experiment results show that using the present proposed GA can obtain better results.

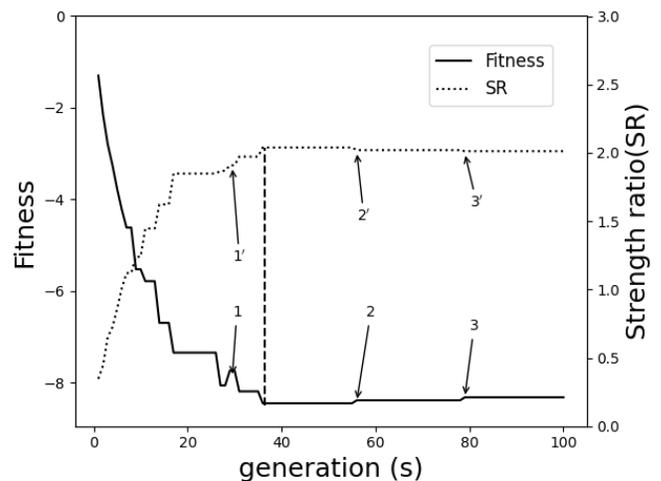


Fig. 8. The Fitness is the Negation of the Individual's Mass. The Solid Curve is the Fitness of the Best Individual in the Population in Respect to the Generations; and the Dotted Line Denotes its Corresponding Strength Ratio. If no Individuals in the Population Satisfy the Constraint, the best Individual is the one with the Biggest Strength Ratio; if not, the best Individual is the one with the Smallest Mass.

TABLE IV. COMPARISON OF EXPERIMENT RESULTS OF CHOUDHURY AND MONDAL'S [30] AND CURRENT STUDY UNDER IN-PLANE LOADING $N_x = 1e6$ N. THE RESULTS OF PRESENT STUDY IS FROM PREVIOUS EXPERIMENT.

Cross Ply [$0_M/90_N$]	Choudhury and Mondal's study		Present study	
	Glass-Epoxy	Graphite-Epoxy	Glass-Epoxy	Graphite-Epoxy
Material				
M	68	17	76	19
N	72	18	49	2
no. of lamina(n)	140	35	125	21
SR	2.01	2.10	2.078	2.15
weight	9.10	1.84	8.12	1.10
cost	140	87.5	125	53

VII. CONCLUSIONS

In this paper, we reviewed the use of the proposed genetic algorithm framework, classical lamination theory, and Tsai-Wu failure theory for the layup design for cross-ply laminate under different loading cases. The principal applications of this genetic algorithm are in the design of composite laminate material with imposed constraints.

The main contribution of the present work is the genetic algorithm framework for constrained problems since the traditional genetic algorithm is primarily concerned with solving the unconstrained problem, which is not suitable for a constrained case. We deal with these constrained problems by the use of mixing strategies of selection methods instead of adding any punishment terms into the objective function. The performance of this algorithm is heavily influenced by the coefficient of the length mutation. Both for the glass-epoxy and graphite-epoxy material cases, if the coefficient takes a relatively low value, this algorithm can obtain better results than when the coefficient value is high. However, The algorithm converged more quickly with a high coefficient value of length mutation.

This variant of the genetic algorithm provides a new approach to address the constrained search, and this method can be easy to apply in other domains. A drawback of the proposed genetic algorithm is more parameters involving in this GA, which makes fine-tuning more difficult.

ACKNOWLEDGMENT

The paper was supported by China Scholarship Council with the code number 201806630112

REFERENCES

- [1] R. L. Riche and R. T. Haftka, "Optimization of laminate stacking sequence for buckling load maximization by genetic algorithm," *AIAA journal*, vol. 31, no. 5, pp. 951–956, 1993.
- [2] S. Nagendra, D. Jestin, Z. Gürdal, R. T. Haftka, and L. T. Watson, "Improved genetic algorithm for the design of stiffened composite panels," *Computers & Structures*, vol. 58, no. 3, pp. 543–555, 1996.
- [3] D. Sadagopan and R. Pitchumani, "Application of genetic algorithms to optimal tailoring of composite materials," *Composites Science and Technology*, vol. 58, no. 3–4, pp. 571–589, 1998.
- [4] A. Todoroki and R. T. Haftka, "Stacking sequence optimization by a genetic algorithm with a new recessive gene like repair strategy," *Composites Part B: Engineering*, vol. 29, no. 3, pp. 277–285, 1998.
- [5] B. Liu, R. T. Haftka, M. A. Akgün, and A. Todoroki, "Permutation genetic algorithm for stacking sequence design of composite laminates," *Computer methods in applied mechanics and engineering*, vol. 186, no. 2–4, pp. 357–372, 2000.
- [6] K. Sivakumar, N. Iyengar, and K. Deb, "Optimum design of laminated composite plates with cutouts using a genetic algorithm," *Composite Structures*, vol. 42, no. 3, pp. 265–279, 1998.

- [7] M. Walker and R. E. Smith, "A technique for the multiobjective optimisation of laminated composite structures using genetic algorithms and finite element analysis," *Composite structures*, vol. 62, no. 1, pp. 123–128, 2003.
- [8] C.-C. Lin and Y.-J. Lee, "Stacking sequence optimization of laminated composite structures using genetic algorithm with local improvement," *Composite structures*, vol. 63, no. 3–4, pp. 339–345, 2004.
- [9] J.-H. Kang and C.-G. Kim, "Minimum-weight design of compressively loaded composite plates and stiffened panels for postbuckling strength by genetic algorithm," *Composite structures*, vol. 69, no. 2, pp. 239–246, 2005.
- [10] M. Murugan, S. Suresh, R. Ganguli, and V. Mani, "Target vector optimization of composite box beam using real-coded genetic algorithm: a decomposition approach," *Structural and Multidisciplinary Optimization*, vol. 33, no. 2, pp. 131–146, 2007.
- [11] M. Akbulut and F. O. Sonmez, "Optimum design of composite laminates for minimum thickness," *Computers & Structures*, vol. 86, no. 21–22, pp. 1974–1982, 2008.
- [12] C. R. Houck, J. Joines, and M. G. Kay, "A genetic algorithm for function optimization: a matlab implementation," *Ncsu-ie tr*, vol. 95, no. 09, pp. 1–10, 1995.
- [13] M. D. Vose, *The simple genetic algorithm: foundations and theory*. MIT press, 1999.
- [14] M. Zbigniew, "Genetic algorithms+ data structures= evolution programs," *Computational Statistics*, pp. 372–373, 1996.
- [15] O. Seresta, Z. Gürdal, D. B. Adams, and L. T. Watson, "Optimal design of composite wing structures with blended laminates," *Composites Part B: Engineering*, vol. 38, no. 4, pp. 469–480, 2007.
- [16] G. Soremekun, Z. Gürdal, R. Haftka, and L. Watson, "Composite laminate design optimization by genetic algorithm with generalized elitist selection," *Computers & structures*, vol. 79, no. 2, pp. 131–143, 2001.
- [17] R. Le Riche and R. Haftka, "Improved genetic algorithm for minimum thickness composite laminate design," *Composites Engineering*, vol. 5, no. 2, pp. 143–161, 1995.
- [18] A. Y. Abu-Odeh and H. L. Jones, "Optimum design of composite plates using response surface method," *Composite structures*, vol. 43, no. 3, pp. 233–242, 1998.
- [19] C. Fang and G. S. Springer, "Design of composite laminates by a monte carlo method," *Journal of composite materials*, vol. 27, no. 7, pp. 721–753, 1993.
- [20] D. J. Deka, G. Sandeep, D. Chakraborty, and A. Dutta, "Multiobjective optimization of laminated composites using finite element method and genetic algorithm," *Journal of reinforced plastics and composites*, vol. 24, no. 3, pp. 273–285, 2005.
- [21] C. H. Park, W. I. Lee, W. S. Han, and A. Vautrin, "Improved genetic algorithm for multidisciplinary optimization of composite laminates," *Computers & structures*, vol. 86, no. 19–20, pp. 1894–1903, 2008.
- [22] S. Omkar, R. Khandelwal, S. Yathindra, G. N. Naik, and S. Gopalakrishnan, "Artificial immune system for multi-objective design optimization of composite structures," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 8, pp. 1416–1429, 2008.
- [23] J.-S. Kim, "Development of a user-friendly expert system for composite laminate design," *Composite Structures*, vol. 79, no. 1, pp. 76–83, 2007.
- [24] M. Gholami, A. Fathi, and A. M. Baghestani, "Multi-objective optimal structural design of composite superstructure using a novel monmpso algorithm," *International Journal of Mechanical Sciences*, p. 106149, 2020.
- [25] T. N. Massard, "Computer sizing of composite laminates for strength," *Journal of reinforced plastics and composites*, vol. 3, no. 4, pp. 300–345, 1984.
- [26] J. Reddy and A. Pandey, "A first-ply failure analysis of composite laminates," *Computers & Structures*, vol. 25, no. 3, pp. 371–393, 1987.
- [27] A. Soeiro, C. C. António, and A. T. Marques, "Multilevel optimization of laminated composite structures," *Structural optimization*, vol. 7, no. 1–2, pp. 55–60, 1994.
- [28] J. L. Pelletier and S. S. Vel, "Multi-objective optimization of fiber reinforced composite laminates for strength, stiffness and minimal mass," *Computers & structures*, vol. 84, no. 29–30, pp. 2065–2080, 2006.
- [29] P. Jadhav and P. R. Mantena, "Parametric optimization of grid-stiffened composite panels for maximizing their performance under transverse loading," *Composite structures*, vol. 77, no. 3, pp. 353–363, 2007.

- [30] A. Choudhury, S. Mondal, and S. Sarkar, "Failure analysis of laminated composite plate under hygro-thermo mechanical load and optimisation," *International Journal of Applied Mechanics and Engineering*, vol. 24, no. 3, pp. 509–526, 2019.
- [31] P. Martin, "Optimum design of anisotropic sandwich panels with thin faces," *Engineering optimization*, vol. 11, no. 1-2, pp. 3–12, 1987.
- [32] C. M. Soares, V. F. Correia, H. Mateus, and J. Herskovits, "A discrete model for the optimal design of thin composite plate-shell type structures using a two-level approach," *Composite structures*, vol. 30, no. 2, pp. 147–157, 1995.
- [33] R. Watkins and A. Morris, "A multicriteria objective function optimization scheme for laminated composites for use in multilevel structural optimization schemes," *Computer Methods in Applied Mechanics and Engineering*, vol. 60, no. 2, pp. 233–251, 1987.
- [34] G. N. Naik, S. Gopalakrishnan, and R. Ganguli, "Design optimization of composites using genetic algorithms and failure mechanism based failure criterion," *Composite Structures*, vol. 83, no. 4, pp. 354–367, 2008.

Kinematic Analysis for Trajectory Planning of Open-Source 4-DoF Robot Arm

Han Zhong Ting¹, Mohd Hairi Mohd Zaman², Mohd Faisal Ibrahim³, Asraf Mohamed Moubark⁴

Department of Electrical, Electronic and Systems Engineering
Faculty of Engineering and Built Environment,
Universiti Kebangsaan Malaysia
43600 Bangi, Selangor, Malaysia

Abstract—Many small and large industries use robot arms to establish a range of tasks such as picking and placing, and painting in today's world. However, to complete these tasks, one of the most critical problems is to obtain the desired position of the robot arm's end-effector. There are two methods for analyzing the robot arm: forward kinematic analysis and inverse kinematic analysis. This study aims to model the forward and inverse kinematic of an open-source 4 degrees of freedom (DoF) articulated robotic arm. A kinematic model is designed and further evaluated all the joint parameters to calculate the end-effector's desired position. Forward kinematic is simple to design, but as for the inverse kinematic, a closed-form solution is needed. The developed kinematic model's performance is assessed on a simulated robot arm, and the results were analyzed if the errors were produced within the accepted range. At the end of this study, forward kinematic and inverse kinematic solutions of a 4-DoF articulated robot arm are successfully modeled, which provides the theoretical basis for the subsequent analysis and research of the robot arm.

Keywords—Robot arm; kinematic analysis; forward kinematic; inverse kinematic; open-source

I. INTRODUCTION

A robot arm is made up of a movable chain of links that are linked together by joints. A hand or end-effector that can move freely in space is typically connected to one end, which is fixed to the ground. Robot arms are capable of performing repetitive tasks at speeds and precision well above those of human operators. Robot arms are used in various fields requiring high precision, for instance medical, industry, and some hazardous places. However, controlling the robot arm has been challenging with a higher degrees of freedom (DoF). Position analysis and trajectory planning of robot arm is an essential step in design and control.

Robotics research has gotten much traction over the years, thanks to advances in robot technology and the growing use of robotics in various fields. The robotics researchers have been focusing on more current configurations, intelligent actions, autonomous robotics, and high-level intelligence. All of these areas are related to robot kinematic, both forward and inverse. However, there are no excellent universal algorithms for producing the forward and inverse kinematic of a robot arm, especially finding the correlation between both kinematic models. The problem involves an error in achieving an accurate correlation between the forward and the inverse kinematic of a robot arm [1], [2], [3], [4], [5], [6], [7], [8], [9]. To plan an accurate robot arm's movement, we have to understand the relationship between the forward and inverse kinematic of the actuators to control the robot's resulting position in an accurate manner. Hence, there remains a scope to investigate further and work towards finding better solutions.

This study aims to develop a suitable methodology for solving the forward and inverse kinematic problem for a 4-DoF articulated robotic manipulator with relative ease. A kinematic model with an accurate correlation between the forward and inverse kinematics of a 4-DoF articulated robotic manipulator is developed using MATLAB software. The performance of the developed kinematic model is assessed and evaluated in Robot Operating System (ROS) on Ubuntu. The robot arm assessed in this research is the OpenMANIPULATOR-X robot arm manufactured by Robotis, which is an open-source robotic platform.

A. Kinematic Modeling

Kinematic is a branch of mechanics that studies the motion of bodies and structures without taking force into account. Geometry is used to research the movement of multi-DoF kinematic chains that make up the robot arm's structure. The relationship between the robot arm's linkages and its position, orientation, and acceleration is studied in robot kinematics. Kinematic analysis is an effective method when planning the robot arm's trajectory can be divided into two types: forward kinematic and inverse kinematic.

Forward kinematic refers to using the kinematic equations of a robot arm to compute the position of the end-effector's frame and joint variables relationship. Meanwhile, inverse kinematic is the reverse process that calculates the joint parameters that achieve an end-effector's position. The reverse operation, on the other hand, is much more challenging in general. The reverse operation is, in general, much more challenging. Forward kinematic has a simple and straightforward solution if compared with the inverse kinematic solution, which has many equations with a highly complex form to be solved [10].

In robotics, the inverse kinematic approach uses the kinematic equations to find the joint parameters that give each of the robot arm's end-effectors the desired configuration (position and rotation). Motion planning determines the robot arm's movement so that its end-effectors move from an initial configuration to the desired configuration. Inverse kinematic transforms the motion plan into joint actuator trajectories for the robot arm. Forward kinematic measures the chain's configuration using joint parameters, while inverse kinematic reverses this calculation to find the joint parameters that achieve the desired configuration [11].

B. Mechanism of Robot Arm

A robot arm can be either serial configuration with an open-loop structure or parallel configuration with a closed-loop structure. The joint type for an industrial robot arm can

be revolute or prismatic, and the link type can be rigid or flexible. Furthermore, a hybrid structure comprising both open and closed-loop mechanical chains is possible. The serial robot arm is characterized by the fact that the first joint is always starting from the fixed base and the end of the chain is free to move in space. Several different configurations of the robot arm can be formed due to the revolute and prismatic joints, and the axes of two adjacent joints can be either parallel or orthogonal. Orthogonal joints intersect by 90° with respect to their common normal, and it can be parallel when one axis rotates 90° [12]. Some of the robot arm mechanisms arise from open, closed, and hybrid open and closed kinematic chains.

In addition, the DoF can be defined as the particular motion of links related to any mechanism or machine. When performing a specific task, the DoF often plays an important role. The total number of DoF is always equal to the number of independent displacements of links. The number of DoF permitted by a joint and their characteristics can be determined by the design constraints of the body or link.

C. Articulated Robot Arm

The articulated robot arm is a robot arm with revolute joints and is also called a revolute, or anthropomorphic robot arm. The anthropomorphic resembles the human arm's design, including shoulder, elbow, and wrist joints [13]. The articulated robot arm can range from a simple two-joints structure to systems with 10 or more interacting joints [14]. The configuration of a 4-DoF articulated manipulator is shown in Fig. 1.

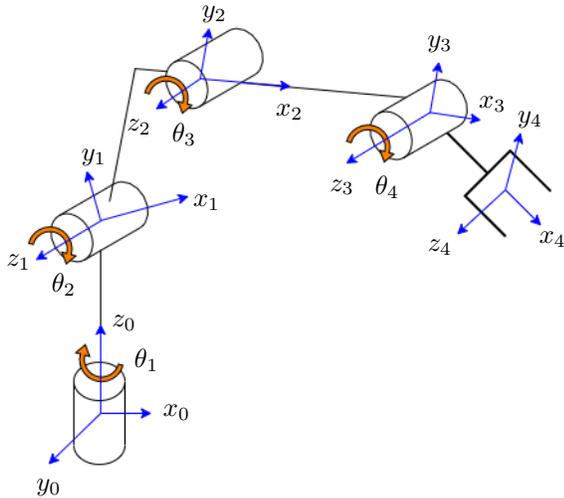


Fig. 1. Configuration of a 4-DoF Articulated Robot Arm.

D. Denavit-Hartenberg (D-H) Convention

In the Denavit-Hartenberg (D-H) convention, the link notation can describe the spatial relationship between two relative joints. The frame that connects with another joint has a relationship to the position and geometry with another joint. Fig. 2 shows that the link parameters are α_i and a_i , and the joint offset (d_i) is fixed to provide the manipulator configuration. This configuration achieves a specific posture using n of θ_i [15]. The posture of every configuration or end-effector can be changed if the value of θ_i changes.

Once the link frame has been set, the position and orientation of the i -frame in relation to $i - 1$ are completely

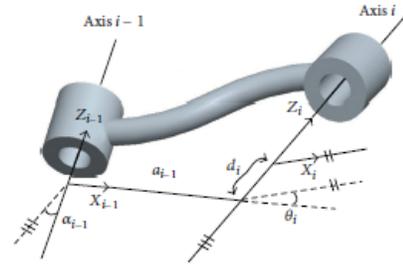


Fig. 2. Definition Parameters of a Robot Arm [15].

defined using four parameters, known as the D-H parameters. The parameters taken are as follows:

- 1) Joint offset, d_i : This parameter is the intersection point's length at the common normal line on the Z_i joint axis. The measured value is at a distance between x_i and x_{i+1} , which is measured along Z_i .
- 2) Joint angle, θ_i : This parameter is the angle between the orthogonal projection on the common normal line, x_i and x_{i+1} , and the normal plane to the Z_i joint axis. The rotation direction is closely related to the parameter value; when the rotation is counterclockwise, it is positive. This parameter is taken at degrees between x_i and x_{i+1} , measured approximately Z_i .
- 3) Link length, a_i : This parameter is taken at the distance of the common normal to Z_i and Z_{i+1} , measured along x_{i+1} .
- 4) Twist angle, α_i : This parameter is the angle between the orthogonal projections on the axis of the joint Z_i and Z_{i-1} to the normal plane on the common normal line. This value can be obtained from the degree between Z_i and Z_{i+1} , measured by x_i .

A transformation matrix can be obtained based on the D-H parameters, which define the transformation of the i -frame relative to the $i - 1$ frame. This matrix can be represented as T_i^{i-1} , and can be calculated as,

$$T_i^{i-1} = \begin{bmatrix} c(\theta_i) & -s(\theta_i)c(\alpha_i) & s(\theta_i)s(\alpha_i) & a_i c(\theta_i) \\ s(\theta_i) & c(\theta_i)c(\alpha_i) & -c(\theta_i)s(\alpha_i) & a_i s(\theta_i) \\ 0 & s(\alpha_i) & c(\alpha_i) & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where c and s represent \sin and \cos , respectively. For n DoF consisting of $n + 1$ frames, the sum of the numbers for the transformation matrix is n [16]. The concatenated matrix provides the required transformation from the n frame that corresponds to the end-effector to the 0-frame mounted on the base as follows:

$$T_n^0 = T_1^0 \cdot T_2^1 \cdot T_3^2 \dots T_n^{n-1} \quad (2)$$

E. Closed-Form Solution

A closed-form solution is an expression for an exact solution given with a finite amount of data [17]. If an equation solves a given problem in terms of functions and mathematical operations from a given commonly accepted set, it is said to be a closed-form solution. An infinite sum, for example, will not be considered a closed-form solution. The closed-form solution

can be applied in solving the inverse kinematic of the robot arm. The solution obtained has the advantages of being exact, includes all solution sets, and low computational cost.

II. METHODS

Some considerations should be emphasized to ensure the robot arm can maneuver correctly, including its software and hardware. Accordingly, software should be stable and easy to use to prevent difficulties. Furthermore, hardware plays a role in forming a solid robot arm and able to move as planned. The MATLAB software (MathWorks) has been used to develop the algorithm in this study.

The primary hardware utilized in this study is the OpenMANIPULATOR-X robot arm manufactured by Robotis. The OpenMANIPULATOR-X robot arm is made up of pieces from the DYNAMIXEL-X series and 3D printing. The daisy chain method is adopted by DYNAMIXEL, which has a modular design. It enables users to add or remove joints for their own use easily. Fig. 3 shows the structure of the OpenMANIPULATOR-X robot arm [18].

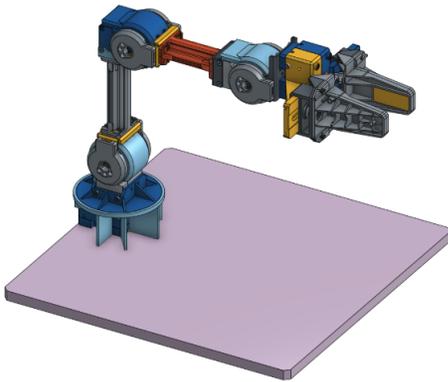


Fig. 3. Structure of the OpenMANIPULATOR-X Robot Arm [18].

A. Forward and Inverse Kinematic Analysis

Fig. 4 shows the flowchart of the forward and inverse kinematic analysis. Forward kinematic modeling was performed using the Denavit-Hartenberg (D-H) convention method. For inverse kinematic modeling, the closed-form solution method was used to obtain all sets of possible solutions. Next, correlation analysis between the forward and inverse kinematic was performed to determine whether there are any relationship between both kinematic models. Finally, performance evaluation was carried out using trajectory planning and waypoint tracking algorithm in MATLAB and ROS on Ubuntu.

B. Modeling of Forward Kinematic

The D-H convention method used to model the forward kinematic of the robot arm can be divided into three processes. The first process is the configuration analysis of the robot arm to obtain its D-H parameters. Fig. 5 shows the OpenMANIPULATOR-X robot arm's configuration analysis, while Table I shows the D-H parameters obtained after the configuration analysis is performed.

There is an offset between joint 2 and joint 3 on the robot arm, resulting in an offset angle. The offset angle is identified and considered in D-H parameters. Fig. 6 shows the calculation of the offset angle, θ_0 , between joints 2 and 3.

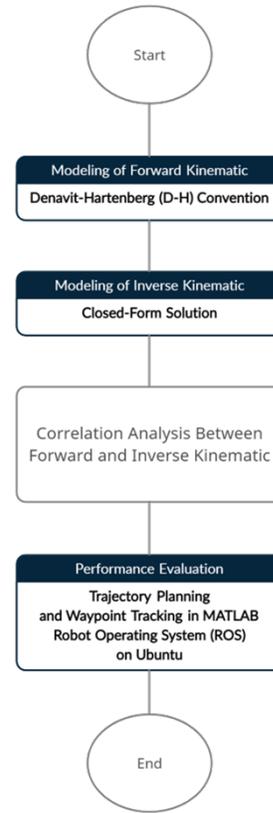


Fig. 4. Flowchart of the Process of Kinematic Modeling.

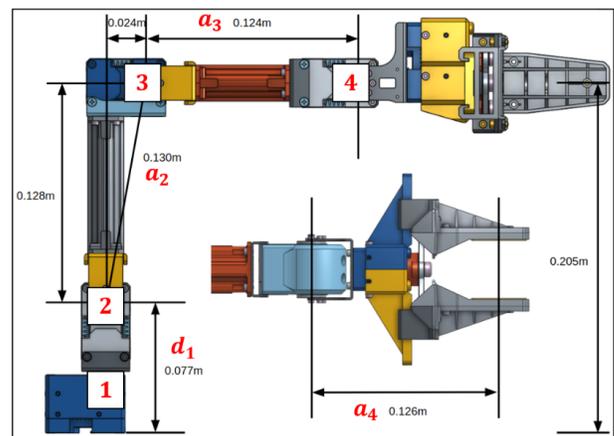


Fig. 5. Configuration Analysis of the OpenMANIPULATOR-X Robot Arm.

TABLE I. D-H PARAMETERS

Joint	θ_i ($^\circ$)	α_i ($^\circ$)	a_i (m)	d_i (m)
1	θ_1	90	0	0.077
2	$\theta_1 - \theta_0$	0	0.130	0
3	$\theta_3 + \theta_0$	0	0.135	0
4	θ_4	0	0.126	0

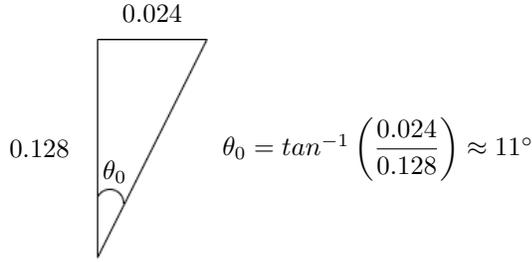


Fig. 6. Calculation of Offset Angle, θ_0 between Joint 2 and 3.

The second process in the D-H convention involves the calculation of the transformation matrix. After obtaining the value of the D-H parameters, the parameters will be included in the transformation matrix. Since 4 frames or 4 DoF were used, the limit for the linked matrix is T_4^0 . The transformation matrix equations can be formulated as follows:

$$T_1^0 = \begin{bmatrix} c(\theta_1) & 0 & s(\theta_1) & 0 \\ s(\theta_1) & 0 & -c(\theta_1) & 0 \\ 0 & 1 & 0 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$T_2^1 = \begin{bmatrix} c(\theta_2) & -s(\theta_2) & 0 & a_2c(\theta_2) \\ s(\theta_2) & c(\theta_2) & 0 & a_2s(\theta_2) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$T_3^2 = \begin{bmatrix} c(\theta_3) & -s(\theta_3) & 0 & a_3c(\theta_3) \\ s(\theta_3) & c(\theta_3) & 0 & a_3s(\theta_3) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$T_4^3 = \begin{bmatrix} c(\theta_4) & -s(\theta_4) & 0 & a_4c(\theta_4) \\ s(\theta_4) & c(\theta_4) & 0 & a_4s(\theta_4) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

After obtaining values for T_1^0 , T_2^1 , T_3^2 and T_4^3 , the next step is to combine them in a sequence to simplify them.

$$T_4^2 = T_3^2 \cdot T_4^3 \quad (7)$$

$$T_4^1 = T_2^1 \cdot T_4^2 \quad (8)$$

$$T_4^0 = T_1^0 \cdot T_4^1 \quad (9)$$

The final process in the D-H convention is to obtain the position vector for the end-effector. After completing the simplification process, the transformation matrix equation is as follows:

$$T_4^0 = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

where $p(p_x, p_y, p_z)$ is the position vector of the end-effector while $n(n_x, n_y, n_z)$, $o(o_x, o_y, o_z)$ and $a(a_x, a_y, a_z)$ are orthogonal unit vectors that determine the orientation of the frame for the end-effector. p_x , p_y , and p_z are the values for x , y , and z coordinates of the end-effector.

C. Modeling of Inverse Kinematic

The closed-form solution method involving geometric and algebraic solutions is used to model the robot arm's inverse kinematic. The solution is divided into two processes. The first process is the calculation to get the angle for joint 1. The movement of the robot arm on the $x-y$ surface depends only on the angle of joint 1.

Regarding the $x-y$ surface projection of the robot arm movement, the angle for joint 1, θ_1 can be calculated as follows:

$$\theta_1 = \tan^{-1} \left(\frac{p_y}{p_x} \right) \quad (11)$$

Since θ_1 is the rotational angle for the robot arm's base, the angle range is between -180° and 180° . The quadrant for θ_1 is identified to determine the sign for each trigonometric ratio in a given quadrant. Table II shows the signs for the trigonometric ratio in each quadrant.

TABLE II. SIGNS FOR THE TRIGONOMETRIC RATIO IN EACH QUADRANT

p_x	p_y	Quadrant	θ_1
+	+	1	θ_1
-	+	2	$\theta_1 + 180^\circ$
-	-	3	$\theta_1 - 180^\circ$
+	-	4	θ_1

The second process involves calculating the angles for joints 2, 3, and 4 (θ_2 , θ_3 , and θ_4). For obtaining the solution for θ_2 , θ_3 , and θ_4 , the 3-dimensional (3D) space which is consisting of the x , y , and z coordinate axes, is simplified to a 2-dimensional (2D) surface. The x and y -axis are merged as a new axis and known as the r -axis, as shown in Fig. 7.

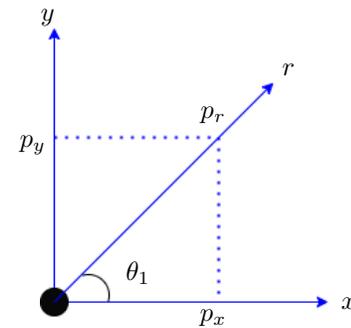


Fig. 7. Combination of x and y -Axis as r -Axis.

The x -coordinate and the y -coordinate for the end-effector are combined as r -coordinate using the Pythagorean equation as follows:

$$p_r^2 = p_x^2 + p_y^2 \quad (12)$$

$$p_r = \sqrt{p_x^2 + p_y^2} \quad (13)$$

Fig. 8 shows the projection of the r-z surface for the movement of the robot arm. The r_3 coordinates and the z_3 coordinates can be calculated as follows:

$$r_3 = p_r \quad (14)$$

$$z_3 = p_z - d_1 \quad (15)$$

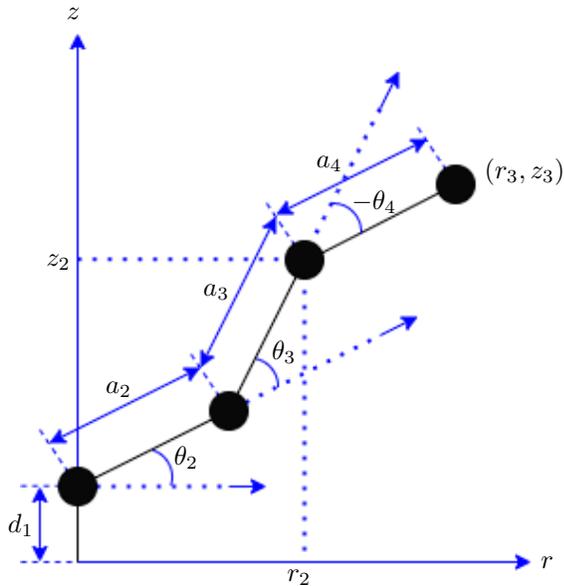


Fig. 8. Projection of the r - z Surface for the Movement of the Robot Arm.

The range of the total sum of angles for θ_2 , θ_3 , and θ_4 are between -90° and 90° as shown in Fig. 9.

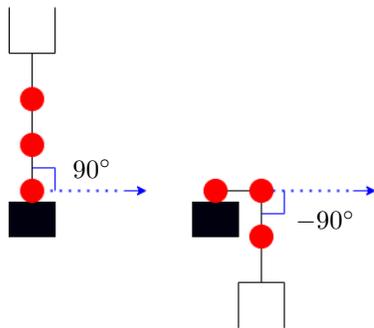


Fig. 9. Maximum and Minimum of the Sum of Angle for Joint 2, 3 and 4.

θ_2 , θ_3 , and θ_4 can be calculated using the following equations:

$$\phi = \theta_2 + \theta_3 + \theta_4 \quad (16)$$

$$r_2 = r_3 - a_4 \cos \phi \quad (17)$$

$$z_2 = z_3 - a_4 \sin \phi \quad (18)$$

$$\cos \theta_3 = \frac{r_2^2 + z_2^2 - (a_2^2 + a_3^2)}{2 a_2 a_3} \quad (19)$$

$$\theta_3 = \pm \cos^{-1} \left(\frac{r_2^2 + z_2^2 - (a_2^2 + a_3^2)}{2 a_2 a_3} \right) \quad (20)$$

$$r_2 = a_2 \cos \theta_2 + a_3 \cos (\theta_2 + \theta_3) \quad (21)$$

$$z_2 = a_2 \sin \theta_2 + a_3 \sin (\theta_2 + \theta_3) \quad (22)$$

$$r_2 = \cos \theta_2 (a_2 + a_3 \cos \theta_3) - \sin \theta_2 (a_3 \sin \theta_3) \quad (23)$$

$$z_2 = \cos \theta_2 (a_3 \sin \theta_3) + \sin \theta_2 (a_2 + a_3 \cos \theta_3) \quad (24)$$

$$\cos \theta_2 = \frac{(a_2 + a_3 \cos \theta_3) r_2 + (a_3 \sin \theta_3) z_2}{r_2^2 + z_2^2} \quad (25)$$

$$\sin \theta_2 = \frac{(a_2 + a_3 \cos \theta_3) z_2 + (a_3 \sin \theta_3) r_2}{r_2^2 + z_2^2} \quad (26)$$

$$\theta_2 = \tan^{-1} \left(\frac{\sin \theta_2}{\cos \theta_2} \right) \quad (27)$$

$$\theta_4 = \phi - (\theta_2 + \theta_3) \quad (28)$$

There are two possible configurations for the robot arm, namely “elbow up” and “elbow down”. Both configurations are considered in the calculation of θ_2 , θ_3 , and θ_4 .

The angle range for θ_2 , θ_3 , and θ_4 is between -90° and 90° , respectively. Based on the configuration of the robot arm in ROS, the angle range for θ_2 becomes 0° and 180° and the angle range for θ_3 becomes -180° and 0° . The angle limit for each joint in ROS ranges between -90° and 90° . Thus, the angle range for θ_2 to be considered in kinematic modeling is between 0 and 90° , while for θ_3 is between -90° and 0° .

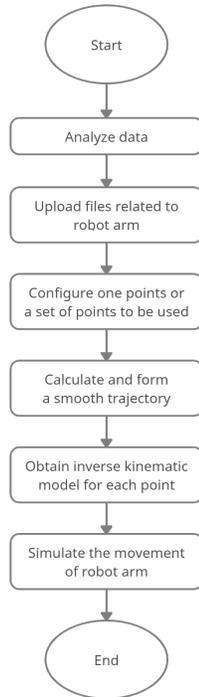


Fig. 10. Simulation Procedures for Trajectory Planning and Waypoint Tracking.

D. Trajectory Planning and Waypoint Tracking

The simulation of trajectory planning and waypoint tracking is done using MATLAB software with the help of Robotics System Toolbox, Simulink and Simscape Multibody. Fig. 10 shows the simulation procedure for trajectory planning and waypoint tracking. The first step to be completed is to analyze the data using MATLAB software. This procedure aims to identify and overcome the dependencies of the simulation files. Next, the algorithm includes uploading files related to the robot arm into MATLAB software. After that, the algorithm configures one point or a set of points to be used. The points are obtained from the modeling of the robot arm and incorporated into the algorithm. This step is very important to determine the similarity with the actual robot arm.

The next step is to calculate and form a smooth trajectory. This process runs the trajectory process and uses the points configured in the previous step. Trajectory planning is one of the critical processes in robots, where it involves the smooth movement of the end-effector from the initial position to the target position [18]. When an arbitrary endpoint is given, the algorithm computes the optimal feasible path for the robot arm's movement based on the kinematic constraints.

The final process involves performing inverse kinematic of the points found on the robot arm workspace. There are six values in the weight vector. The first three values are for rotational purposes, and the last three values are for transition purposes. Then, the algorithm configures a number greater than the number of points previously configured to ensure the robot arm's smooth movement. After that, it uses a kinematic solver for each end-effector position and uses the previous configuration to make an initial guess. After completing all the processes, the movement of the robot arm is simulated.

III. RESULTS AND DISCUSSION

A. Modeling of Forward Kinematic

Fig. 11 shows the flowchart of the forward kinematic modeling algorithm. D-H parameters are included in the developed algorithm to obtain the end-effector's coordinate of the robot arm, and the transformation matrix is calculated to obtain the equations for the end-effector's coordinates. Next, random values of the four joint angles are generated by using the "randi" function. This function is able to generate uniformly distributed random integers from the specified interval. Finally, the end-effector's coordinates are calculated by inserting the random values for the four joint angles into the transformation matrix equations. Table III shows the results of the forward kinematic modeling developed.

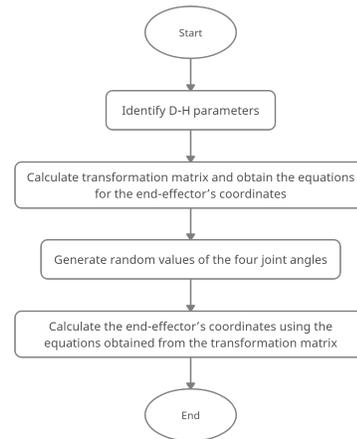


Fig. 11. Flowchart of Forward Kinematic Modeling Algorithm.

TABLE III. RESULTS OF FORWARD KINEMATIC MODELING

Case	Joint Angles (°)				End-Effector's Coordinates		
	θ_1	θ_2	θ_3	θ_4	p_x	p_y	p_z
1	103	15	-5	21	-0.1489	0.3508	0.115
2	56	3	-13	79	0.1716	0.2544	0.1531
3	65	68	-23	-20	0.1185	0.2542	0.3347
4	-22	34	-21	53	0.2804	-0.1133	0.2733
5	48	8	-65	32	0.2125	0.236	-0.0963
6	11	70	-8	-67	0.2512	0.0488	0.2966
7	166	42	-35	-40	-0.3407	0.0849	0.0918
8	82	38	-58	36	0.0632	0.3583	0.1246
9	-158	56	-79	14	-0.3158	-0.1276	0.0965
10	-54	75	-37	9	0.1465	-0.2017	0.3691

B. Modeling of Inverse Kinematic

Fig. 12 shows the flowchart of the inverse kinematic modeling algorithm. The coordinates of the end-effector are entered into the algorithm to obtain the joint angles that achieve a particular end-effector position. Next, the angle value for joint 1 is calculated, followed by calculating the angle values for

joints 2, 3, and 4. Finally, all solution sets for the joint angles are obtained.

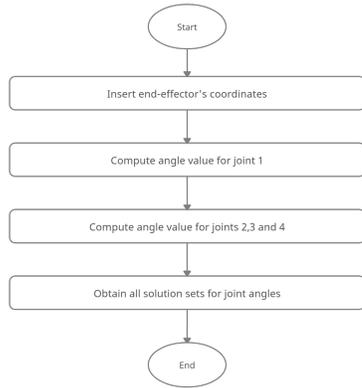


Fig. 12. Flowchart of Inverse Kinematic Modeling Algorithm.

Table IV shows the set of solutions for joint angles in integers obtained using MATLAB software. Referring to case 1 from the results of the forward kinematic modeling in Table V, a total of 66 sets of solutions can be obtained for the end-effector's coordinates ($p_x = -0.0834$, $p_y = 0.3611$ and $p_z = 0.1744$) including the desired solution set, which is the 65th set of solution ($\theta_1 = 103^\circ$, $\theta_2 = 15^\circ$, $\theta_3 = -5^\circ$ and $\theta_4 = 21^\circ$).

TABLE IV. SET OF SOLUTIONS FOR JOINT ANGLES IN THE FORM OF INTEGER

Solution	Joint Angles ($^\circ$)			
	θ_1	θ_2	θ_3	θ_4
1:	103	28	-2	-28
2:	103	38	-20	-19
3:	103	26	2	-28
4:	103	39	-24	-15
5:	103	24	5	-28
6:	103	40	-27	-13
7:	103	23	7	-28
8:	103	41	-29	-10
9:	103	21	9	-27
10:	103	41	-31	-8
11:	103	20	10	-26
12:	103	42	-32	-5
13:	103	19	12	-25
14:	103	42	-34	-3
15:	103	18	13	-25
16:	103	42	-35	-1
17:	103	17	14	-24
18:	103	42	-36	1
19:	103	16	14	-22
20:	103	42	-36	2
21:	103	15	15	-21
22:	103	42	-37	4
23:	103	14	16	-20
24:	103	42	-38	6
25:	103	14	16	-19
26:	103	41	-38	8
27:	103	13	16	-18
28:	103	41	-38	9
29:	103	12	17	-16
30:	103	41	-39	11
31:	103	12	17	-15
32:	103	40	-39	13
33:	103	11	17	-13

Solution	Joint Angles ($^\circ$)			
	θ_1	θ_2	θ_3	θ_4
34:	103	40	-39	14
35:	103	11	17	-12
36:	103	39	-39	16
37:	103	11	17	-10
38:	103	39	-39	17
39:	103	10	16	-9
40:	103	38	-38	18
41:	103	10	16	-7
42:	103	37	-38	20
43:	103	10	15	-5
44:	103	37	-37	21
45:	103	10	15	-3
46:	103	36	-37	22
47:	103	9	14	-1
48:	103	35	-36	23
49:	103	9	13	0
50:	103	34	-35	24
51:	103	9	12	2
52:	103	33	-34	25
53:	103	10	11	5
54:	103	32	-33	26
55:	103	10	9	7
56:	103	31	-31	27
57:	103	10	8	9
58:	103	29	-30	27
59:	103	11	6	11
60:	103	28	-28	28
61:	103	12	3	14
62:	103	26	-25	28
63:	103	13	0	17
64:	103	24	-22	28
65:	103	15	-5	21
66:	103	21	-17	27

To obtain more accurate angles with more decimal points, the number of solution sets increases. The number of solution sets increases by about 10 times with the increase of one

decimal point for the joint angles, as shown in Table V.

TABLE V. RESULTS OF INVERSE KINEMATIC MODELING

Case	End-Effector's Coordinates			Joint Angles ($^\circ$)				Number of Solution Sets		
	p_x	p_y	p_z	θ_1	θ_2	θ_3	θ_4	0 DP	1 DP	2 DP
1	-0.0833	0.361	0.1744	103	15	-5	21	66	668	6672
2	0.1716	0.2544	0.1531	56	3	-13	79	179	1769	17686
3	0.1185	0.2542	0.3347	65	68	-23	-20	76	762	7616
4	0.2804	-0.1133	0.2733	-22	34	-21	53	134	1346	13454
5	0.2125	0.236	-0.0963	48	8	-65	32	132	1318	13174
6	0.2512	0.0488	0.2966	11	70	-8	-67	147	1462	14617
7	-0.3407	0.0849	0.0918	166	42	-35	-40	156	1546	15468
8	0.0632	0.3583	0.1246	82	38	-58	36	120	1192	11902
9	-0.3158	-0.1276	0.0965	-158	56	-79	14	176	1752	17518
10	0.1465	-0.2017	0.3691	-54	75	-37	9	64	630	6310

C. Correlation Analysis of Kinematic Modeling

Referring to the forward and inverse kinematic modeling results, as shown in Tables III and V, the solution is precisely the same when comparing the joint angles from the forward and inverse kinematic. The same solution set of joint angles could be obtained using the inverse kinematic model, as applied in the forward kinematic model. Thus, the correlation between the developed forward and inverse kinematic model is said to be accurate.

D. Performance of End-Effector Coordinates

Fig. 13 and 14 display the example of simulation results of trajectory planning and waypoint tracking for the set of solutions for the joint angles. The simulation results show that the solution set of joint angles which produce the most feasible trajectory are $\theta_1 = 103^\circ$, $\theta_2 = 42^\circ$, $\theta_3 = -35^\circ$ and $\theta_4 = 0^\circ$. Furthermore, Table VI depicts the comparison of the end-effector's coordinates obtained from the modeling process in MATLAB and output from ROS. The most significant errors occurred in case 1 and case 5. Errors occur between the end-effector's coordinates obtained from the results of modeling and output from ROS due to the rounding calculations that unavoidable in modeling and the kinematic constraints on the simulated robot arm in ROS as follows:

- 1) Joint Bounds: This constraint is satisfied if the robot configuration vector maintains all joint positions within the bounds specified.
- 2) Cartesian bounds: This constraint is satisfied if the end-effector origin's position relative to the target frame remains within the bounds specified.
- 3) Orientation target: This constraint requires the end-effector orientation to match a target orientation within an angular tolerance in any direction. The target orientation is specified relative to the body frame of the reference body.
- 4) Pose target: This constraint requires the end-effector's pose to match a target pose within a distance and angular tolerance in any direction. The target pose is specified relative to the body frame of the reference body.
- 5) Position target: This constraint requires the end-effector position to match a target position within a distance tolerance in

any direction. The target position is specified relative to the body frame of the reference body.

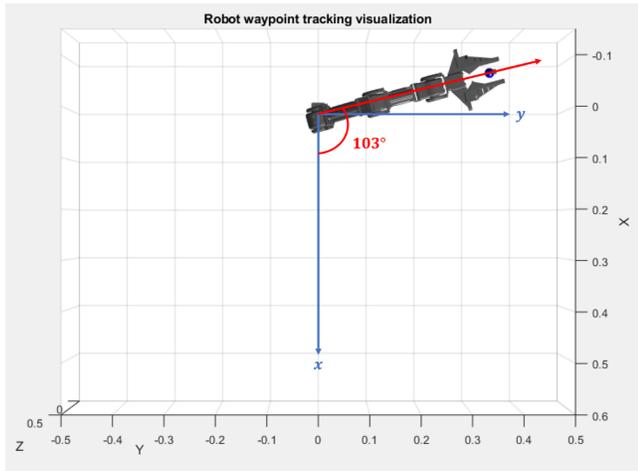


Fig. 13. Simulation Results of Trajectory Planning and Waypoint Tracking for Joint Angle 1.

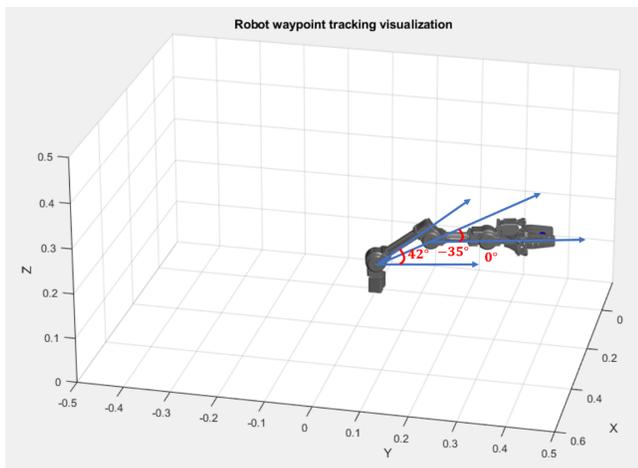


Fig. 14. Simulation Results of Trajectory Planning and Waypoint Tracking for Joint Angles 2, 3 and 4.

TABLE VI. PERFORMANCE OF END-EFFECTOR COORDINATES FROM MODELING AND ROS.

Case	End-Effector's Coordinates						Error (%)		
	Modeling			Output from ROS					
	p_x	p_y	p_z	p_x	p_y	p_z	p_x	p_y	p_z
1	-0.0833	0.361	0.1744	-0.073	0.366	0.1784	14.25	1.34	2.24
2	0.1716	0.2544	0.1531	0.178	0.246	0.155	3.6	3.41	1.23
3	0.1185	0.2542	0.3347	0.127	0.247	0.327	6.69	2.91	2.35
4	0.2804	-0.1133	0.2733	0.282	-0.109	0.271	0.57	3.94	0.85
5	0.2125	0.236	-0.0963	0.221	0.232	-0.087	3.85	1.72	10.69
6	0.2512	0.0488	0.2966	0.257	0.048	0.287	2.26	1.67	3.34
7	-0.3407	0.0849	0.0918	-0.318	0.082	0.091	7.14	3.54	0.88
8	0.0632	0.3583	0.1246	0.061	0.35	0.129	3.61	2.37	3.41
9	-0.3158	-0.1276	0.0965	-0.294	-0.124	0.101	7.41	2.9	4.46
10	0.1465	-0.2017	0.3691	0.153	-0.194	0.362	4.25	3.97	1.96

IV. CONCLUSION

This study successfully modeled both the forward and inverse kinematic model for the 4-DoF articulated robot arm. Correlation between the forward and inverse kinematic model is analyzed, and the performance of the kinematic model is evaluated on the simulated robot arm. The developed kinematic model serves as a theoretical framework for further research, for instance robot arm trajectory planning, and structural optimization of robot design.

ACKNOWLEDGMENT

This work was financially supported by Universiti Kebangsaan Malaysia (grant no. GGPM-2019-051 and GUP-2018-103).

REFERENCES

- [1] G. Bao, S. Liu, and H. Zhao, "Kinematics simulation of 4 DOF manipulator," vol. 123, pp. 400–408, 2017.
- [2] Z. Chao, F. Wang, C. Zhang, and H. Li, "Inverse kinematics solution and verification of 4-DOF hydraulic manipulator," Journal of Physics: Conference Series, vol. 916, no. 1, 2017.
- [3] T. Dewi, S. Nurmaini, P. Risma, Y. Oktarina, and M. Roriz, "Inverse kinematic analysis of 4 DOF pick and place arm robot manipulator using fuzzy logic controller," Int. Journal of Electrical and Computer Engineering, vol. 10, no. 2, 2020.
- [4] S. Gómez, G. Sánchez, J. Zarama, M. C. Ramos, J. E. Alcántar, A. Nuñez, J. Torres, S. Santana, and F. Nájera, "Design of a 4-Dof robot manipulator with optimized algorithm for inverse kinematics," Int. Journal of Mechanical and Mechatronics Engineering, vol. 9, no. 6, pp. 929–934, 2016.
- [5] A. A. Mohammed, and M. Sunar, "Kinematics modeling of a 4-DOF robotic arm," in Proceedings of the 2015 Int. Conf. on Control, Automation and Robotics, pp. 87–91, 2015.
- [6] A. Novitarini, Y. Aniroh, D. Y. Anshori, and S. Budiprayitno, "A closed-form solution of inverse kinematic for 4 DOF tetrix manipulator robot," in Proceedings of the 2017 Int. Conf. on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation, pp. 25–29, 2017.
- [7] L. Rónai, and T. Szabó, "Kinematical investigation and regulation of a 4DOF model robot," Acta Mechanica Slovaca, vol. 20, no. 3, pp. 50–56, 2016.
- [8] A. Singh, and A. Singla, "Kinematic Modeling of Robotic Manipulators," in Proceedings of the 2017 National Academy of Sciences India Section A - Physical Sciences, pp. 303–319, 2017.
- [9] T. Singh, P. Suresh, and S. Chandan, "Forward and inverse kinematic analysis of robotic manipulators," Int. Research Journal of Engineering and Technology, vol. 4, no. 2, pp. 1459–1469, 2017.
- [10] A. El-Sherbiny, M. Elhosseini, and A. Haikal, "A comparative study of soft computing methods to solve inverse kinematics problem," Ain Shams Engineering Journal, vol. 9, no. 4, pp. 2535–2548, 2018.
- [11] P. Jha, "Inverse kinematic analysis of robot manipulators," India: National Institute of Technology Rourkela, 2015.
- [12] J. M. McCarthy, and G. S. Soh, "Geometric design of linkages," New York: Springer, 2010.
- [13] L. Luthsamy, H.F. AL-Qrimli, S.Shazzana Wan Taha, and N.A. Raj, "Design and control of an anthropomorphic robotic arm," Journal of Industrial Engineering Research, vol. 2, no. 1, pp. 1–8, 2016.
- [14] I. Al-Naimi, "Introduction to robot manipulators," Robotics and Automation, 2018.
- [15] P. Corke, "Robotics, vision and control: Fundamental algorithms in MATLAB," 2nd ed. Switzerland: Springer, 2017.
- [16] S. Singh, and E. Singla, "Service arms with unconventional robotic parameters for intricate workstations: Optimal number and dimensional synthesis," Journal of Robotics, pp. 1–11, 2016.
- [17] M. v. Hoeij, "Closed form solutions," Florida: Florida State University, 2017.
- [18] Robotis, "OpenMANIPULATOR-X," Available online: <https://emanual.robotis.com> [accessed on 18 May 2021].

Fake News Detection in Arabic Tweets during the COVID-19 Pandemic

Ahmed Redha Mahlous¹

College of Computer and Information Sciences
Prince Sultan University
Riyadh, Saudi Arabia

Ali Al-Laith²

Center for Language Engineering - KICS
University of Engineering and Technology
Lahore, Pakistan

Abstract—In March 2020, the World Health Organization declared the COVID-19 outbreak to be a pandemic. Soon afterwards, people began sharing millions of posts on social media without considering their reliability and truthfulness. While there has been extensive research on COVID-19 in the English language, there is a lack of research on the subject in Arabic. In this paper, we address the problem of detecting fake news surrounding COVID-19 in Arabic tweets. We collected more than seven million Arabic tweets related to the corona virus pandemic from January 2020 to August 2020 using the trending hashtags during the time of pandemic. We relied on two fact-checkers: the France-Press Agency and the Saudi Anti-Rumors Authority to extract a list of keywords related to the misinformation and fake news topics. A small corpus was extracted from the collected tweets and manually annotated into fake or genuine classes. We used a set of features extracted from tweet contents to train a set of machine learning classifiers. The manually annotated corpus was used as a baseline to build a system for automatically detecting fake news from Arabic text. Classification of the manually annotated dataset achieved an F1-score of 87.8% using Logistic Regression (LR) as a classifier with the n-gram-level Term Frequency-Inverse Document Frequency (TF-IDF) as a feature, and a 93.3% F1-score on the automatically annotated dataset using the same classifier with count vector feature. The introduced system and datasets could help governments, decision-makers, and the public judge the credibility of information published on social media during the COVID-19 pandemic.

Keywords—Fake news; Twitter; social media; Arabic corpus

I. INTRODUCTION

The rise of social networks such as Facebook, Twitter, and many others has enabled the rapid spread of information. Any user on social media can publish whatever they want without considering the truthfulness and reliability of the published information, which introduces challenges in information reliability assurance. Twitter is one of the most popular social media platforms. It is designed to allow users to send information as short texts, known as tweets, with no more than 280 characters, and each user on Twitter can follow as many accounts as he or she wants. Nowadays, and with the outbreak of the COVID-19 pandemic, millions of tweets are generated daily, which has caused some adverse effects that impact individuals and society. For example, the spread of misinformation about COVID-19 symptoms may harm people [1]. For instance, it could be anxiety-inducing for a person who experiences COVID-19 like symptoms even if they have not been infected with the virus. The terms fake news and misinformation are closely related and are often

used interchangeably. Authors in [2] defined rumors as: “a hypothesis offered in the absence of verifiable information regarding uncertain circumstances that are important to those individuals who are subsequently anxious about their lack of control resulting from this uncertainty.” Another definition presented in [3] is: “unverified and instrumentally relevant information statements in circulation that arise in contexts of ambiguity, danger or potential threat, and that function to help people make sense and manage risk.”

Detecting fake news in English tweets is an active research area and many studies and datasets have been published during the COVID-19 pandemic [4]. In Arabic, fake news detection is fairly new and has a long way to go to reach the level achieved in other languages, especially English. Therefore, the fight against fake news requires a system that automatically assists in verifying the truthfulness of shared information about the COVID-19 pandemic on social media. Fake news detection is a very challenging task, especially with the lack of available datasets related to the pandemic. An automated fake news detection system is necessary by utilizing human annotation, machine/deep learning, and Natural Language Processing techniques [5]. These techniques help to determine whether a given text is fake news or not by comparing the text with some pre-known corpora that contain both fake and truthful information [6].

In this paper, we address the problem of fake news detection on Twitter during the COVID-19 pandemic period. Our focus is to build a manually annotated dataset for fake news detection from Twitter’s social media platform. We rely on fact-checking sources to manually annotate a sample dataset. The consideration of these fact-checking sources could help in reducing the spread of misinformation [7], [8], [9], [10]. As manual annotation is expensive and time-consuming [11], we also developed a system to expand the manually annotated dataset by automatically annotating a large and unlabeled dataset. We use a supervised learning classification to train and test both the manually and automatically annotated datasets to ensure the quality of our annotation. We use six different machine learning algorithms, four different features with each algorithm, and three pre-processing techniques. The rest of the paper is organized as follows. In Section 2, we cover related work. Section 3 presents our methodology to annotate and automatically detect fake news related to COVID-19. In Section 4, we present the results and discussion. Finally, the conclusion and future work are presented in Section 5.

II. RELATED WORK

Recently, many works have been done to tackle the issue of detecting fake news, rumors, misinformation, or disinformation in social media networks. Most of these studies can be categorized into supervised and unsupervised learning approaches. Moreover, there are fewer works that tackled the problem using semi-supervised techniques.

For the supervised approach, a system based on machine learning techniques for detecting fake news or rumors in the Arabic language from social media during the COVID-19 pandemic is presented in [12]. The authors collected one million Arabic tweets using Twitter's Streaming API. The collected tweets were analyzed by identifying the topics discussed during the pandemic, detecting rumors, and predicting the source of the tweets. A sample of 2,000 tweets was labeled manually into false information, correct information, and unrelated. Different machine learning classifiers were applied, including Support Vector Machine, Logistic Regression, and Naïve Bayes. They obtained 84% accuracy in identifying rumors. The limitations of this research include the unavailability of the dataset, and the fact that it relies on a single source of rumors: the Saudi Arabian Ministry of Health.

Identifying breaking news rumors on Twitter has been proposed in [13]. The authors built a word2vec model and an LSTM-RNN model to detect rumors from news published on social media. The proposed model is capable of detecting rumors based on a tweet's text, and experiments showed that the proposed model outperforms state-of-the-art classifiers. As rumors can be deemed later to be true or false, their model is unable to memorize the facts across time; it only looks at the tweet at the current time. Detecting rumors from Arabic tweets using features extracted from the user and content has been proposed in [14]. The authors obtained rumors and non-rumors topics from anti-rumors and Ar-Riyadh websites. More than 270K tweets were collected, containing 89 and 88 rumour and non-rumour events, respectively. A supervised Gaussian Naïve Bayes classification algorithm reported an F1-score of 78.6%. This research's limitation is that the proposed dataset is not verified using any of the benchmark datasets.

In [15], a supervised learning approach for Twitter credibility detection is proposed. A set of features including content-based and source-based features, were used to train five machine learning classifiers. The Random Forests classifier outperformed the other classifiers when used with a combined set of features. A total of 3,830 English tweets were manually annotated with credible or non-credible classes. The textual features were not studied to examine their impact on credibility detection. Another supervised machine learning approach was proposed in [16] to detect rumors from business reviews. A publicly available dataset was used to conduct rumour detection experiments. Different supervised learning classifiers were used to classify business reviews. The experimental results showed that the Naïve Bayes classifier achieved the highest accuracy and outperformed three classifiers, namely, the Support Vector Classifier, K-Nearest Neighbors, and Logistic Regression. This work's limitation is the small size of the dataset used to train machine learning classifiers.

Detection of fake news using n-gram analysis and machine learning techniques was proposed in [17]. Two different feature

extraction techniques and six machine learning algorithms were investigated and compared based on a dataset from political articles that were collected from Reuters.com and kaggle.com for real and fake news. Another Arabic corpus for the task of detecting fake news on YouTube is presented in [18]. The authors introduced a corpus that covered topics most concerned by rumors. More than 4,000 comments were collected to build the corpus. Three different machine learning classifiers (Support Vector Machine, Decision Tree, and Multinomial Naïve Bayes) were used to differentiate between rumour and non-rumour comments with the n-gram TF-IDF feature. The SVM classifier achieved the highest results. Authors in [19] proposed identifying fake news on social media. They used several pre-processing steps on the textual data, and then used 23 supervised classifiers with the TF weighting feature. The combined text pre-processing and supervised classifiers were tested on three different real-world English datasets, including BuzzFeed Political News, Random Political News, and ISOT Fake News.

An automatic approach to detecting fake news from Arabic and English tweets using machine learning classifiers has been proposed in [4]. The authors developed a large and continuous dataset for Arabic and English fake news during the COVID-19 pandemic. Information shared on official websites and Twitter accounts were considered a source of real information. Along with the data collected from official websites and Twitter accounts, they also relied on various fact-checking websites to build the dataset. A set of 13 machine learning classifiers and seven other feature extraction techniques were used to build fake news models. These models were used to automatically annotate the dataset into real and fake information. The dataset was collected for 36 days, from the 4th of February to the 10th of March 2020.

A large corpus for fighting the COVID-19 infodemic on social media has been proposed in [11]. The authors developed a schema that covers several categories including advice, cure, call for action, or asking a question. They considered such categories to be useful for journalists, policymakers, or even the community as a whole. The collected dataset contains tweets in Arabic and English. Three classifiers were used to perform classification experiments using three input representations: word-based, FastText, and BERT. The authors only made 210 of the classified tweets public.

Two Arabic corpora have also been constructed, without manual annotation. In [20], more than 700,000 Arabic tweets were collected from Twitter during the COVID-19 period. The corpus covers prevalent topics discussed in that period and is publicly available to enable research under different domains such as NLP, information retrieval, and computational social media. They used the Twitter API to collect the tweets on a daily basis, covering the period from January 27, 2020, to March 31, 2020.

The second corpus is presented in [21]. The tweets were collected during the period of the COVID-19 pandemic to study the pandemic from a social perspective. The corpus was developed to identify information influencers during the month of March 2020, and contains nearly four million tweets. Different algorithms were used to analyze the influence of information spreading and compare the ranking of users.

For fake news detection in other languages, there are many corpora that are publicly available to tackle the spread of false information. A multilingual cross-domain fact-checking news dataset for COVID-19 has been introduced in [22]. The collected dataset covered 40 languages and relies on fact-checked articles from 92 different fact-checking websites to manually annotate the dataset. The dataset is available on GitHub. Another publicly available dataset called “TweetsCOV19” was introduced in [23]. This dataset contains more than eight million English tweets about the COVID-19 pandemic. The dataset can be used for training and testing a wide range of NLP and machine learning methods and is available online. A novel Twitter dataset is presented in [24], which was developed to characterize COVID-19 misinformation communities. Authors categorized the tweets into 17 classes, including fake cure, fake treatment, and fake facts or prevention. They performed different tasks on the developed dataset, including identifying communities, network analysis, bot detection, sociolinguistic analysis, and vaccination stance. This study’s limitations are that only one person performed annotation, the analyses are correlational and not causal, and the collected data covered a short period of only three weeks. MM-COVID is a multilingual and multidimensional fake news data repository [25]. The dataset contains 3981 fake and 7192 genuine news contents from English, Spanish, Portuguese, Hindi, French, and Italian. The authors explored the collected dataset from different perspectives including social engagements and user profiles on social media.

Sentiment analysis has also been used in fake news detection has also been facilitated. In [26], the authors used sentiment analysis to eliminate neutral tweets. They claimed that tweets related to fake news are more negative and have strong sentiment polarity in comparison with genuine news. The main issue in using this approach to detect fake news from Arabic text is the lack of Arabic sentiment resources, including sentiment lexicons and corpora [27]. Testing whether emotions play a role in the formation of beliefs in online political misinformation is presented in [28]. The authors explore emotional responses as an under-explored mechanism of belief in political misinformation. Understanding emotions helps in different domains including capturing the public’s sentiments about social events such as the spreading of misinformation on social media [29].

Text classification using machine/deep learning provides a good results over many NLP applications including, sentiment analysis [30], [31], emotion detection [32], hate speech detection [33], sarcasm detection [34], and other applications.

To summarize, most of the existing datasets target the English language, with only a few targeting Arabic. Furthermore, most of the Arabic datasets related to COVID-19 are published without annotation. Datasets that are annotated were annotated automatically and collected during a short period of time. Additionally, not all of these datasets are publicly available. In this research, we address these issues by employing three annotators to manually perform the annotation task.

III. METHODOLOGY

Fig. 1 presents the architecture of the proposed fake news detection system. In the first step of the framework, we collect

data from Twitter using the Twitter Streaming API. In the second step, we perform the extraction of tweets which discuss rumors or fake news topics during the pandemic, annotate a small sample of tweets manually, and develop a system to annotate a large dataset of unlabeled tweets automatically.

In the last step, we store the dataset in a database and use it to accomplish our experiments and analysis. This research intends to build an Arabic fake news corpus that can be used for analyzing the spread of fake news on social media during the COVID-19 pandemic. To address this need, we perform the following four steps: 1) data collection, 2) rumor/misinformation keyword extraction, 3) data pre-processing, and 4) fake news annotation.

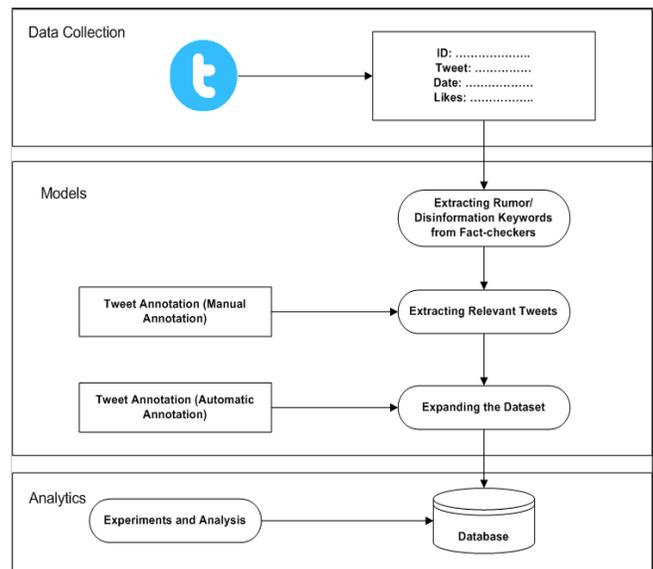


Fig. 1. Fake New Detection Architecture.

A. Data Collection

In this section, we describe the process of data collection from Twitter. In the first instance, we prepared a list of hashtags that appeared during the COVID-19 outbreak, as shown in Table I. Armed with the Tweepy Python library and using Twitter’s API, we proceeded to collect Arabic tweets related to COVID-19 from January 1, 2020, until May 31, 2020. We then searched for tweets containing one or more of the defined hashtags in the tweet’s text. This step allowed us to collect more than seven million unique tweets. After applying some filters such as removing the short and repeated tweets, the remaining tweets are 5.5 million tweets. However, as some of the collected tweets were irrelevant, we decided to keep only those tweets relevant to the COVID-19 pandemic and containing fake news keywords.

B. Fake News Keywords Extraction

To collect a list of keywords relevant to the rumours circulating during the pandemic, we used two sources:

- Agence France-Presse (AFP)¹ with its newly formed health investigation team which have the responsibility

¹<https://factuel.afp.com/ar/CORNA%20COMPILATION%202-20>

TABLE I. LIST OF HASHTAGS USED TO COLLECT THE DATASET

#	Hashtag	English Translation
1	#كورونا	Corona
2	#فيروس_كورونا	Corona virus
3	#كورونا_المستجد	New Corona
4	#كوفيد_١٩	COVID 19
5	#الفيروس_التاجي	Corona virus
6	#الحجر_المنزلي	Home Quarantine
7	#الحجر_الصحي	Quarantine
8	#التباعد_الاجتماعي	Social distancing
9	#كلنا_مسؤول	We are all responsible
10	#ابقوا_في_منازلكم	Stay in your homes

of dealing with large amounts of fake news in various languages and indicating their error or inaccuracy.

- The Anti-Rumours Authority (No Rumours)², an independent project established in 2012 to address and contain rumours and sedition to prevent them from causing any harm to society.

After reading and analyzing rumours and misinformation circulated on social media using the above-mentioned sources, a list of 40 keywords was extracted and used to prepare our dataset, as shown in Table II. These keywords cover a variety of topics associated with fake news, rumours, racism, unproven cure methods, false information. For example, there was a rumour that herbal tea is used to treat COVID-19. Another topic circulated was that Cristiano Ronaldo offered to transform his hotels into hospitals and give free treatment to COVID-19 patients. One alleged that the corona virus targets only those who have yellow-skin and Asian people to reduce population density. Other topics include the conversion of non-Muslims to Islam.

We extracted a corpus of more than 37,000 unique tweets related to rumours and misinformation topics during the COVID-19 pandemic. The tweets were written by 24,117 users with an average of 1.5 tweets per user. Statistical information details about the corpus are presented in Table III.

C. Data Pre-processing

We performed several text pre-processing steps based on the procedure described in [35] in order to sanitize the collected tweets before annotation and classification. Our dataset, which is a mixture of modern standard Arabic and dialectal Arabic, requires further filtering such as removing duplicated letters, strange words, and non-Arabic words. The following is a complete list of the steps performed:

- Removing mentions, hyperlinks, and hashtags.
- Removing non-Arabic and strange words.
- Text normalization.

²https://twitter.com/No_Rumors

TABLE II. LIST OF VERIFIED RUMORS AND MISINFORMATION TOPICS

#	Keyword	English Translation
1	شبيكات الجيل الخامس	5G networks
2	العجز الجنسي	Impotence
3	اكتشاف علاج	Discover a cure
4	درجة حرارة ٢٦	A temperature of 26
5	حبس النفس	Holding your breath
6	الغرغرة	Gargle
7	الموز	Banana
8	بخار الماء	Water vapor
9	تشخيص حالتك اونلاين	Online diagnosis of your condition
10	الشاي العشبي	Herbal tea
11	يلقون باموالهم	Throw in their money
12	يرمون اموالهم	Throw in their money
13	محففات الايدي	Hand dryers
14	الساونا	Sauna
15	الرئيس يبكي	The president is crying
16	١٥ دقيقة	Water 15 minutes
17	يصفون	Spit
18	الزندانى كورونا	Zindani Corona
19	الاوغور	Uyghurs
20	الأعشاب	Herbs
21	يدخلون الإسلام	They enter Islam
22	الرئيس الصيني	Chinese President
23	البشرة السوداء	Black skin
24	يعتنون	Embrace
25	توزيع مصاحف	Distributing the Qur'an
26	صدام حسين	Saddam Hussein
27	مستشفيات رونالدو	Ronaldo Hospitals
28	الاذان غرناطة	Call to prayer in Granada
29	الاذان لندن	Call to prayer in London
30	هروب صيني الى اليمن	Chinese escape to Yemen
31	اسلمو	Entered Islam
32	جميع الاديان	All religions
33	تاسوكو	Tasuco
34	أوباما	Obama
35	بيل غيتس	Bill Gates
36	رفع اذان المانيا	Call to prayer in Germany
37	يصلون	They pray
38	متروك للسماء	Up to the sky
39	نقص الأوكسجين	Lack of oxygen
40	المغرب يصدر الكمامات	Morocco exports masks

- Removing punctuations and Arabic diacritics.
- Removing repeated characters which add noise and influence the mining process.

Two libraries were used to perform further pre-processing on the corpus text, including stemming and rooting. The first library is NLTK, which was used to perform stemming on the corpus text using ISRIStemmer4³. The second library is

³<https://www.kite.com/python/docs/nltk.ISRIStemmer>

Tashaphyne⁴, which was used to get the root of each word in the corpus.

IV. CORPUS ANNOTATION

A. Manual Annotation

A sample of 2,500 tweets was manually annotated into fake or genuine classes. We developed a small application to facilitate the annotation process, as shown in Fig. 2. We involved three annotators in the annotation of the sample dataset. Two of the annotators performed annotation while the third was tasked with evaluating their output and resolving conflicts. We requested the annotators to read and understand the list of guidelines and informed them to skip tweets in which there are mixture of fake news and genuine topics and only annotate tweets that have a clear and distinct fake news topic. The following are the guidelines:

- Tweets are generally considered fake if one fake news topic is discussed in the tweet.
- Tweets are considered to not be fake if one fake news topic is discussed in the tweet, and the topic is negated.
- Tweets that contain a mixture of both fake and genuine news are skipped.



Fig. 2. Fake News Annotation Interface.

The annotation process resulted in a corpus containing 1,537 tweets (835 fake and 702 genuine), after excluding duplicated tweets, tweets that contain mixed fake and genuine news, and tweets where the fake news was meant as sarcasm. Statistical information about the manually annotated corpus is shown in Table III. We used Cohen's kappa coefficient to measure the inter-annotator agreement, obtaining a value of 0.91. Table IV shows an example of some annotated tweets.

B. Automatic Annotation

Initially, we trained different machine learning classifiers on the manually annotated corpus and used the best performing classifier to automatically predict the fake news classes of remaining unlabeled tweets. The outcome of the prediction process is 34,529 tweets (19,582 fake and 19,582 genuine) as shown in Table III.

During the annotation process, the annotators found some tweets containing fake news keywords but carrying sarcasm. In this case, the annotators were requested to annotate them as genuine. Table V shows a sample of such tweets.

⁴<https://pypi.org/project/Tashaphyne/>

TABLE III. CORPUS STATISTICS

Manually Annotated Corpus		
	Fake Tweets	Not Fake Tweets
Total Tweets	835	702
Total Words	20,395	19,852
Unique Words	6,246	7,115
Total Characters	117,630	113,121

Automatically Annotated Corpus		
	Fake Tweets	Not Fake Tweets
Total Tweets	19,582	14,947
Total Words	479,349	463,768
Unique Words	79,383	88,037
Total Characters	2,855,454	2,680,067

TABLE IV. FAKE AND GENUINE TWEETS EXAMPLES

#	Tweet	Class
1	نفسى اكتب اشاعة علي فيروس كورونا اللي تخلي الناس مرزوعة في بيوتها اقول فيها كورونا يسبب الضعف الجنسي حتى لو شفيت منها. I want to write a rumour about COVID-19 that would make people stay in their homes, and say that it causes impotence even if you recover from it.	Genuine
2	ربما مفرك مثل كلام صدام حسين عن كورونا قبل اكثر من عشرين عام. Perhaps fabricated like Saddam Hussein's video about COVID-19 more than 20 years ago.	Genuine
3	تحذير صيني فيروس كورونا يسبب الضعف الجنسي والعقم. Chinese warning: COVID-19 causes impotence and sterility.	Fake
4	كورونا الفيروس مصنع لاصحاب البشرة الصفراء والاسويين لتقليص الكثافة السكانية على الارض والدليل انه لم يصاب اي احد من ذوي البشرة السوداء. COVID-19 is made for yellow-skinned people and Asians to reduce population density, and the evidence for that is that no black-skinned person has been infected.	Fake

V. EXPERIMENTS

In this section, we present the results of the fake news classification after describing the employed feature extraction techniques, experimental setup, classifier model training, and evaluation measures.

A. Feature Extraction

The next step after performing text pre-processing is to prepare the features to build classification models. To accomplish that, we used the following features:

- Count Vector: The text in our corpus was converted into a vector of term counts.
- Word-Level TF-IDF: Each term in our corpus is represented in a TF-IDF matrix.

TABLE V. ANNOTATION CONFUSION

#	Tweet	English Translation
1	بالفيديو ابنة صدام حسين تحسم جدل فيديو والدها عن كورونا.	In the video, Saddam Hussein's daughter resolves the controversy of her father's video about COVID-19
2	اقرا عن بيل غيتس ومسالة لقاح كورونا ومبلغ ال ٧٥ مليون	Read about Bill Gates, the issue of the COVID-19 vaccine, and the 75 million[dollar] amount

TABLE VI. STATISTICS IN EXPERIMENTAL DATASET

Manually Annotated Corpus			
	Fake Tweets	Not Fake Tweets	Total Tweets
Training	668	562	1,230
Testing	167	140	307
Total Tweets	835	720	1,537

Automatically Annotated Corpus			
	Fake Tweets	Not Fake Tweets	Total Tweets
Training	15,666	11,958	27,624
Testing	3,926	2,989	6,905
Total Tweets	19,582	14,947	34,529

- N-gram-Level TF-IDF: We used unigram, bigram, and trigram models in our experiments. We then represented these terms in a matrix containing TF-IDF scores.
- Character-Level TF-IDF: We represent TF-IDF character scores for each tweet in our corpus.

These features are used to train multiple classifiers in order to build machine learning models with the ability to decide the most probable category for new, unseen tweets.

B. Experimental Setup

This section describes the experimental configurations used to perform the text classification task. We designed a set of experiments aiming to validate and ensure the quality of manually and automatically generated annotations. We also explored fake news detection as a binary classification problem (fake and genuine). The total tweets in our fake news dataset are 1,537 and 34,529 tweets in both manually and automatically annotated corpora, respectively. We divided both datasets into 80% for training and 20% for testing. Table VI shows detail about the manual and automatic annotated datasets.

Six machine learning classifiers were used to perform fake news classification for both datasets: Naïve Bayes, Logistic Regression (LR), Support Vector Machine (SVM), Multilayer Perceptron (MLP), Random Forest Bagging Model (RF), and eXtreme Gradient Boosting Model (XGB). The following are the hyper-parameters used with each classifier:

- NB: alpha=0.5
- LR: with default values
- SVM: c=1.0, kernel=linear, gamma=3
- MLP: activation function=ReLU, maximum iterations=30, learning rate=0.1
- RF: with default values
- XGB: with default values

C. Models Training

Once the numerical form of the textual tweets was complete, the data frame containing the count vector, word-level TF-IDF, n-gram level TF-IDF, and character-level TF-IDF representations for each tweet in our corpus were used to train six different classifiers. We used scikit-learn, a Python library for classifier implementation and prediction of the classes of the unlabeled dataset. K-fold cross-validation was used to select the classifier that provides the highest results and shows

the best ability to generalize. The collection was split into five-folds, four of which were used for training on each iteration, and the fifth for evaluation.

D. Evaluation Measures

The evaluation was carried out using three measures: Precision, Recall, and F1-score as follows:

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (1)$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (2)$$

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

Where:

- True Positive: the number of fake tweets that are correctly predicted as fake tweets.
- True Negative: the number of genuine tweets that are correctly predicted as genuine tweets.
- False Positive: the actual class is genuine, but the predicted class is fake.
- False Negative: the actual class is fake, but the predicted class is genuine.

E. Experimental Results

We present the experimental results on the Arabic fake news dataset. Six machine learning classifiers (NB, LR, SVM, MLP,RF, and XGB) were used to perform our experiments on the manually and automatically annotated datasets. We used count vector and TF-IDF vectorization (word-level, n-gram-level, and character-level) to train the classifiers. Precision, recall, and F1-score are the measures that have been used to evaluate each classifier using 5-fold cross-validation. Bold values indicate which setting yielded the best classification performance of fake tweets.

The results showed that using the LR classifier with the n-gram TF-IDF feature and without applying further pre-processing on the text (such as stemming or rooting) yielded a significantly better classification performance. The classifier gave an 87.8% F1-score classification result with the manually annotated corpus, as shown in Table VII. The same classifier, with the word count feature and without applying stemming or rooting, obtained the best classification performance when applied to the automatically annotated corpus, as shown in Table VIII. It achieved an F1-score of 93.3%.

As shown in Fig. 3, the highest precision value was obtained using the n-gram TF-IDF feature with the LR classifier (87.8%) and the count vector feature with the LR classifier (93.4%) on manually and automatically annotated corpora, respectively. The results obtained using raw text is better than with the corpus text after applying stemming and rooting. We can conclude that performing further pre-processing did not

TABLE VII. PRECISION (P), RECALL (R), AND F1-SCORE (F1) CLASSIFICATION RESULTS (MANUAL ANNOTATED DATASET)

Classifier	Feature Measure	Word Count			TF-IDF (word-level)			TF-IDF (n-gram-level)			TF-IDF (character-level)		
		Raw Text	Stemming	Rooting	Raw Text	Stemming	Rooting	Raw Text	Stemming	Rooting	Raw Text	Stemming	Rooting
NB	P	77.4	78.1	75.65	82.61	80.9	78.4	80.4	85.8	82.07	83.8	83.3	77.72
	R	77.1	77.5	74.68	73.38	75.6	70.13	74	77.5	73.38	72.7	74.9	69.16
	F1	77.2	77.6	74.98	75.13	76.7	71.9	75.5	78.9	75.06	75.2	76.5	71.04
LR	P	84	79.9	79.8	82.1	74	81.4	87.8	76	80.9	81.3	68.7	78
	R	84	79.8	79.9	81.8	74	81.2	87.7	76	79.9	80.5	68.2	77.3
	F1	84	79.9	79.9	81.5	74	81.2	87.8	76	80.1	80.2	68.3	77.5
SVM	P	80.8	76.3	79.2	78.5	79.9	82.6	85.7	81.6	86.1	80.6	78.1	76.7
	R	79.9	76	79.2	78.6	79.9	82.5	85.7	81.2	85.7	80.5	77.9	76
	F1	80.1	76.1	79.2	78.4	79.9	82.4	85.7	81	85.7	80.3	77.7	76
MLP	P	83.6	78.64	78.37	78.7	76.91	76.21	80.4	78.22	79.01	86.4	77.14	77
	R	83.1	78.57	77.92	78.6	75.65	75.97	78.6	78.25	75.32	86.4	77.27	76.62
	F1	83.2	78.6	78.1	78.6	76.06	76.09	78.9	78.23	76.06	86.4	77.15	76.79
RF	P	81.16	80.44	79.21	74.96	74.44	78.31	77.45	75.35	77.43	79.39	73.79	75.92
	R	77.6	79.55	78.57	75	74.03	78.25	77.27	74.68	75.65	79.22	73.38	75
	F1	78.27	79.69	78.73	74.96	74.12	78.27	77.33	74.81	76.04	79.28	73.47	75.23
XGB	P	74.99	78.2	72.08	73.26	79.53	73.34	79.77	77.39	75.82	76.59	76.66	74.05
	R	74.03	76.95	70.78	71.75	76.95	73.05	77.6	75	75	75.97	75	73.05
	F1	74.26	77.23	71	72.11	77.44	73.11	78	75.46	75.13	76.13	75.35	73.21

TABLE VIII. PRECISION (P), RECALL (R), AND F1-SCORE (F1) CLASSIFICATION RESULTS (AUTOMATIC ANNOTATED DATASET)

Classifier	Feature Measure	Word Count			TF-IDF (word-level)			TF-IDF (n-gram-level)			TF-IDF (character-level)		
		Raw Text	Stemming	Rooting	Raw Text	Stemming	Rooting	Raw Text	Stemming	Rooting	Raw Text	Stemming	Rooting
NB	P	76.7	75.6	72.6	76.2	76	74	74.4	75.9	75.1	72.8	73.6	70.1
	R	76.4	75.5	72.5	76.2	74.8	69.4	74.4	76	75	72.8	73.6	70
	F1	76.6	75.4	72.4	76.2	75.1	70.5	74.4	75.9	74.9	72.8	73.6	70
LR	P	93.4	84.6	76.3	91.8	85.8	77.1	90.8	85.8	79.7	84.3	83.1	76.9
	R	93.3	84.6	76	91.7	85.8	77	90.7	85.7	79.6	84.2	83.1	76.9
	F1	93.3	84.6	76.1	91.7	85.8	77.1	90.7	85.7	79.6	84.3	83.1	76.9
SVM	P	92.1	82.9	78.3	91.2	84.2	79.8	90.6	85.4	82	89.7	83.8	76
	R	92	82.8	78	91.2	84.2	79.6	90.5	85.3	81.8	89.1	83.1	75.1
	F1	92	82.8	78	91.2	84.2	79.7	90.5	85.3	81.9	89.4	83.3	75.4
MLP	P	88.8	79.6	70.1	87.1	77	70.4	71.7	73.2	72.2	80.5	78	72.3
	R	88.5	79.5	70	87.1	77	70.4	71.7	73.2	72.2	80.5	78	72.3
	F1	88.6	79.5	70.1	87.1	77	70.4	71.7	73.2	72.2	80.5	78	72.3
RF	P	84.9	82.5	77.7	84.6	82.1	77.7	84.9	81.5	78.3	78.3	79	76.1
	R	84.7	82.3	77.5	84.7	82.1	77.6	84.9	81.6	78.3	78.3	78.9	76.1
	F1	84.7	82.4	77.6	84.6	82.1	77.6	84.9	81.5	78.3	78.3	78.9	76.1
XGB	P	82.8	82.1	76.2	82.8	81.7	75.8	82.3	82	76.3	80.1	80.1	76.5
	R	80.2	80.4	74.7	80.9	80.5	75.2	79.9	80.4	75.2	79.3	79.3	75.8
	F1	80.7	80.7	75.1	81.2	80.7	75.3	80.4	80.7	75.5	79.5	79.5	76

enhance the classification results with the text from social media.

As shown in Fig. 4, the highest recall was obtained using the count vector TF-IDF feature with the LR classifier (87.7%), and the count vector feature with the LR classifier (93.3%) on manually and automatically annotated corpora, respectively. The highest F1-score, as shown in Fig. 5 was obtained using the n-gram-level TF-IDF feature with the MLP classifier (87.8%) and the count vector feature with the LR classifier (93.3%) on manually and automatically annotated corpora, respectively.

VI. DISCUSSION

The primary objective of this research was to build a benchmark dataset for fake news in Arabic related to the COVID-19 pandemic. We introduce a new fake news corpus in the Arabic language, collected from Twitter. It is clear from the experimental results that the manually annotated corpus can be used as a baseline for further research in the domain of fake news and misinformation. As there remains no

benchmark dataset for fake news detection in Arabic related to the COVID-19 pandemic, this corpus will help the research community once the dataset is publicly available. The proposed corpus was manually annotated by three annotators to ensure the quality and usefulness of the developed corpus. We used a set of machine learning classifiers to train different machine learning models on the manually annotated corpus. The best model was selected to predict fake news classes of unlabeled tweets (more than 35,000 tweets). The statistical analysis showed lower precision, recall, and F1-score values in the classification of the manually annotated corpus, while the automatically annotated corpus showed improved results. From the results presented in the previous section, we notice that increasing the size of the dataset leads to an improvement in the classification results using precision, recall, and F1-score measures.

The use of machine learning methods to classify the fake news corpus using content-based features gives better results than the user-based features. The corpus can be further expanded using two methods: 1) increasing the number of

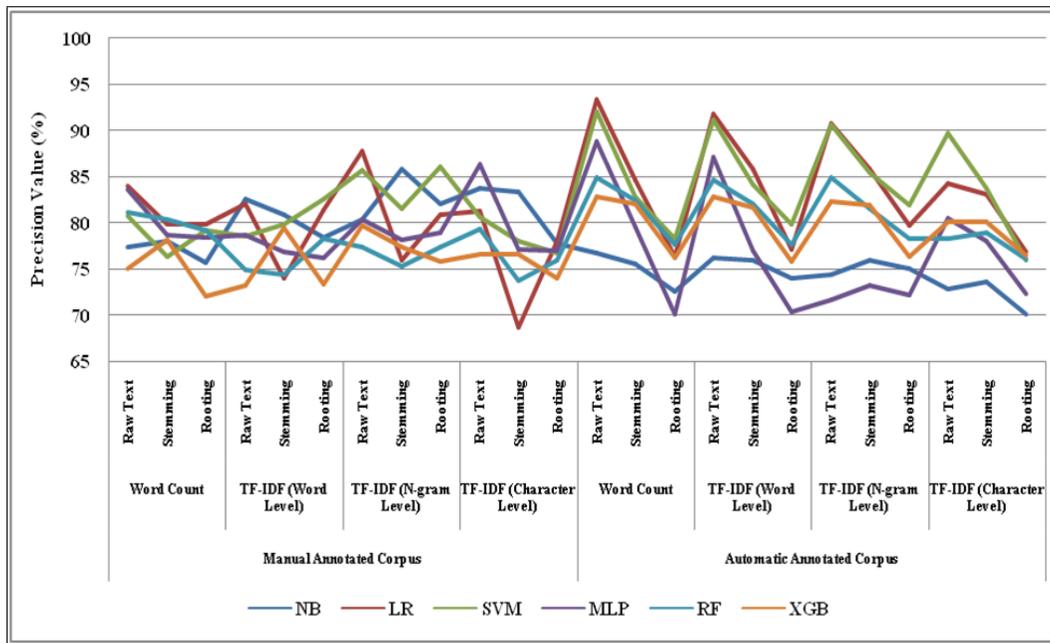


Fig. 3. Precision Results.

verified rumour misinformation topics, or 2) performing classification on more unlabeled tweets related to the COVID-19 pandemic. After then, the deep learning approach can be used to enhance fake news classification.

VII. CONCLUSION AND FUTURE WORK

In this paper, we introduced a new Arabic corpus of fake news that will be made publicly available for research purposes on this link: (<https://github.com/yemen2016/FakeNewsDetection>), after preparing the tweets ID's and their associated classes. We explained the collection process of fake news and gave details about how we select rumors and misinformation topics during the COVID-19 pandemic. The classification task was performed using six classifiers (Naïve Bayes, Logistic Regression, Support Vector Machine, Multilayer Perceptron, Random Forest Bagging, and eXtreme Gradient Boosting) to test the possibility of recognizing fake and genuine tweets. We used four feature types: count vector, word-level TF-IDF, n-gram-level TF-IDF, and character-level TF-IDF. We noticed that the achieved performance varies depending on the features and classifiers used. Along with considering the raw text as an input to the machine learning classifiers, we also used two pre-processing methods: stemming and rooting. Both techniques failed to improve the classification results as the corpus text was collected from Twitter, which includes various dialects and language mistakes. Therefore, the stemming and rooting procedures did not produce correct results. The study concluded that we can achieve higher performance with more annotated data.

In the future, we plan to expand our corpus with additional verified rumour and misinformation topics. We also look forward to investigating the performance of new classification methods such as deep learning. In this research, we only

used content-based features to classify and analyze fake news, though user-based features may also be utilized.

ACKNOWLEDGMENT

“This work was supported by the research Project PSU-COVID19 Emergency Research Program; Prince Sultan University; Saudi Arabia [PSU-COVID19 Emergency Research Program-CCIS-2020-57]”.

REFERENCES

- [1] J. AlHumaid, S. Ali, and I. Farooq, “The psychological effects of the covid-19 pandemic and coping with them in saudi arabia?” *Psychological Trauma: Theory, Research, Practice, and Policy*, vol. 12, no. 5, p. 505, 2020.
- [2] H. B. Dunn and C. A. Allen, “Rumors, urban legends and internet hoaxes,” in *Proceedings of the Annual Meeting of the Association of Collegiate Marketing Educators*, 2005, p. 85.
- [3] N. DiFonzo and P. Bordia, “Rumor, gossip and urban legends,” *Diogenes*, vol. 54, no. 1, pp. 19–35, 2007.
- [4] M. K. Elhadad, K. F. Li, and F. Gebali, “Covid-19-fakes: a twitter (arabic/english) dataset for detecting misleading information on covid-19,” in *International Conference on Intelligent Networking and Collaborative Systems*. Springer, 2020, pp. 256–268.
- [5] R. Oshikawa, J. Qian, and W. Y. Wang, “A survey on natural language processing for fake news detection,” *arXiv preprint arXiv:1811.00770*, 2018.
- [6] M. K. Elhadad, K. F. Li, and F. Gebali, “Fake news detection on social media: a systematic survey,” in *2019 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*. IEEE, 2019, pp. 1–8.
- [7] P. L. Liu and L. V. Huang, “Digital disinformation about covid-19 and the third-person effect: examining the channel differences and negative emotional outcomes,” *Cyberpsychology, Behavior, and Social Networking*, vol. 23, no. 11, pp. 789–793, 2020.
- [8] N. M. Krause, I. Freiling, B. Beets, and D. Brossard, “Fact-checking as risk communication: the multi-layered risk of misinformation in times of covid-19,” *Journal of Risk Research*, vol. 23, no. 7-8, pp. 1052–1059, 2020.

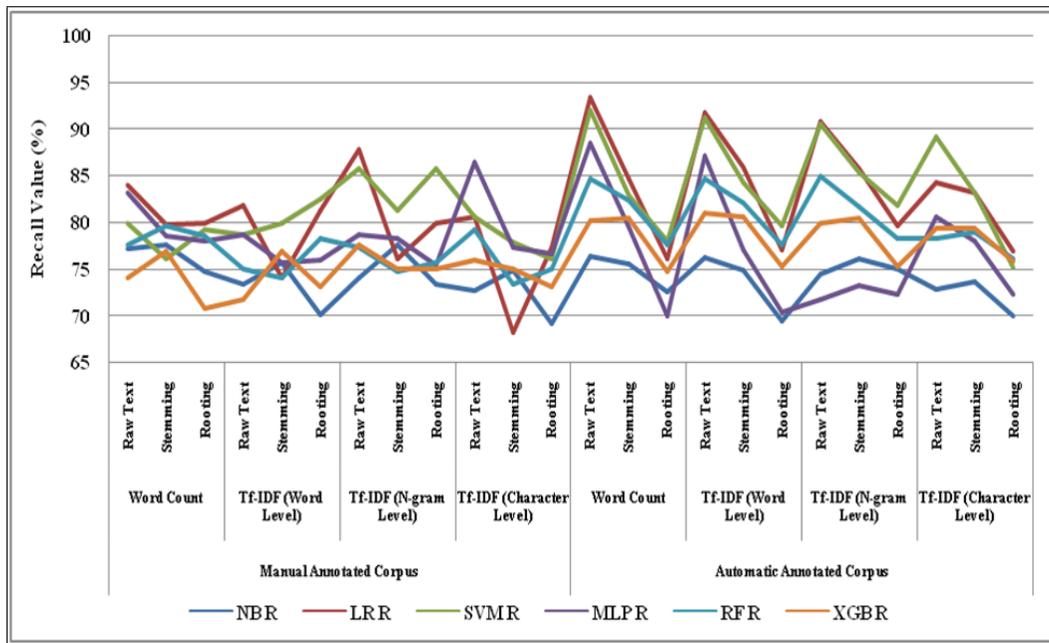


Fig. 4. Recall Results.

- [9] J. Donovan, "Concrete recommendations for cutting through misinformation during the covid-19 pandemic," 2020.
- [10] M. Luengo and D. García-Marín, "The performance of truth: politicians, fact-checking journalism, and the struggle to tackle covid-19 misinformation," *American Journal of Cultural Sociology*, vol. 8, no. 3, pp. 405–427, 2020.
- [11] F. Alam, F. Dalvi, S. Shaar, N. Durrani, H. Mubarak, A. Nikolov, G. D. S. Martino, A. Abdelali, H. Sajjad, K. Darwish *et al.*, "Fighting the covid-19 infodemic in social media: a holistic perspective and a call to arms," *arXiv preprint arXiv:2007.07996*, 2020.
- [12] L. Alsudias and P. Rayson, "Covid-19 and arabic twitter: How can arab world governments and public health organizations learn from social media?" in *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*, 2020.
- [13] S. A. Alkhodair, S. H. Ding, B. C. Fung, and J. Liu, "Detecting breaking news rumors of emerging topics in social media," *Information Processing & Management*, vol. 57, no. 2, p. 102018, 2020.
- [14] S. M. Alzanin and A. M. Azmi, "Rumor detection in arabic tweets using semi-supervised and unsupervised expectation-maximization," *Knowledge-Based Systems*, vol. 185, p. 104945, 2019.
- [15] N. Y. Hassan, W. H. Gomaa, G. A. Khoriba, and M. H. Haggag, "Supervised learning approach for twitter credibility detection," in *2018 13th International Conference on Computer Engineering and Systems (ICCES)*. IEEE, 2018, pp. 196–201.
- [16] A. Habib, S. Akbar, M. Z. Asghar, A. M. Khattak, R. Ali, and U. Batool, "Rumor detection in business reviews using supervised machine learning," in *2018 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESCom)*. IEEE, 2018, pp. 233–237.
- [17] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques," in *International conference on intelligent, secure, and dependable systems in distributed and cloud environments*. Springer, 2017, pp. 127–138.
- [18] M. Alkhair, K. Meftouh, K. Smaïli, and N. Othman, "An arabic corpus of fake news: Collection, analysis and classification," in *International Conference on Arabic Language Processing*. Springer, 2019, pp. 292–302.
- [19] F. A. Ozbay and B. Alatas, "Fake news detection within online social media using supervised artificial intelligence algorithms," *Physica A: Statistical Mechanics and its Applications*, vol. 540, p. 123174, 2020.
- [20] F. Haouari, M. Hasanain, R. Suwaileh, and T. Elsayed, "Arcov-19: The first arabic covid-19 twitter dataset with propagation networks," in *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, 2021, pp. 82–91.
- [21] S. Alqurashi, A. Alashaikh, and E. Alanazi, "Identifying information superspreaders of covid-19 from arabic tweets," *Preprints*, 2020.
- [22] G. K. Shahi and D. Nandini, "Fakecovid—a multilingual cross-domain fact check news dataset for covid-19," *arXiv preprint arXiv:2006.11343*, 2020.
- [23] D. Dimitrov, E. Baran, P. Fafalios, R. Yu, X. Zhu, M. Zloch, and S. Dietze, "Tweetscov19—a knowledge base of semantically annotated tweets about the covid-19 pandemic," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 2020, pp. 2991–2998.
- [24] S. A. Memon and K. M. Carley, "Characterizing covid-19 misinformation communities using a novel twitter dataset," *arXiv preprint arXiv:2008.00791*, 2020.
- [25] Y. Li, B. Jiang, K. Shu, and H. Liu, "Mm-covid: A multilingual and multidimensional data repository for combating covid-19 fake news," *arXiv preprint arXiv:2011.04088*, 2020.
- [26] L. Cui and D. Lee, "Coaid: Covid-19 healthcare misinformation dataset," *arXiv preprint arXiv:2006.00885*, 2020.
- [27] O. Oueslati, E. Cambria, M. B. HajHmida, and H. Ounelli, "A review of sentiment analysis research in arabic language," *Future Generation Computer Systems*, vol. 112, pp. 408–430, 2020.
- [28] L. Rosenzweig, B. Bago, A. J. Berinsky, and D. Rand, "Misinformation and emotions in nigeria: The case of covid-19 fake news," 2020.
- [29] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, "Affective computing and sentiment analysis," in *A practical guide to sentiment analysis*. Springer, 2017, pp. 1–10.
- [30] A. Al-Laith and M. Shahbaz, "Tracking sentiment towards news entities from arabic news on social media," *Future Generation Computer Systems*, vol. 118, pp. 467–484, 2021.
- [31] A. Al-Laith, M. Shahbaz, H. F. Alaskar, and A. Rehmat, "Arasencorpus: A semi-supervised approach for sentiment annotation of a large arabic text corpus," *Applied Sciences*, vol. 11, no. 5, p. 2434, 2021.
- [32] A. Al-Laith and M. Alenezi, "Monitoring people's emotions and symptoms from arabic tweets during the covid-19 pandemic," *Information*, vol. 12, no. 2, p. 86, 2021.
- [33] M. Z. Ali, Ehsan-Ul-Haq, S. Rauf, K. Javed, and S. Hussain, "Improving

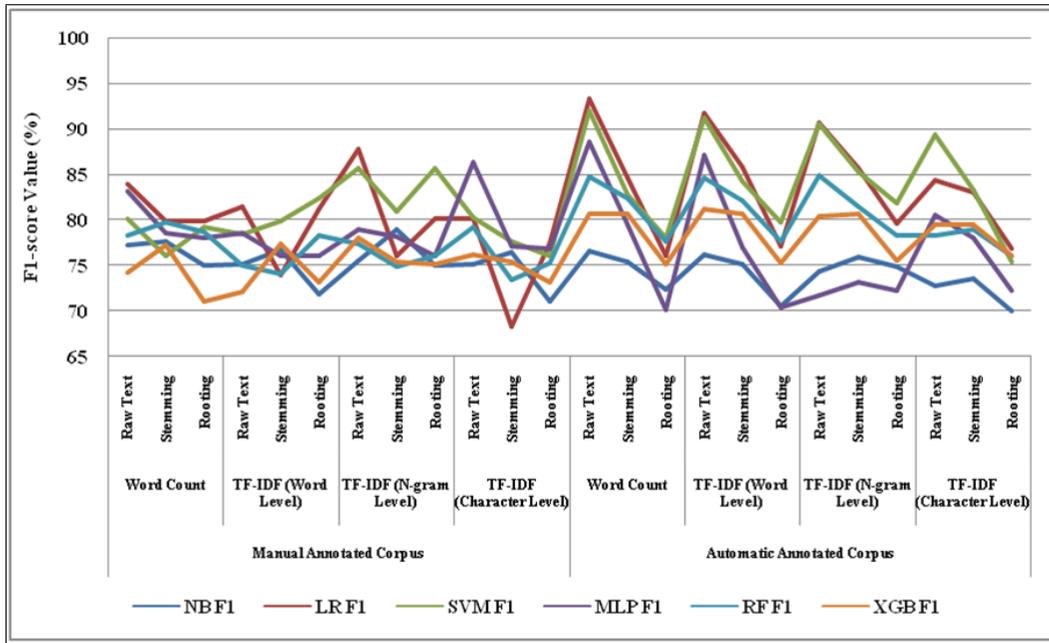


Fig. 5. F1-score Results.

hate speech detection of urdu tweets using sentiment analysis,” *IEEE Access*, vol. 9, pp. 84 296–84 305, 2021.

[34] A. Allaith, M. Shabbaz, and M. Alkoli, “Neural network approach for irony detection from arabic text on social media.” in *FIRE (Working*

Notes), 2019, pp. 445–450.

[35] S. R. El-Beltagy, M. E. Kalamawy, and A. B. Soliman, “Niletmg at semeval-2017 task 4: Arabic sentiment analysis,” *arXiv preprint arXiv:1710.08458*, 2017.

Detecting COVID-19 Utilizing Probabilistic Graphical Models

Emad Alsuwat¹, Sabah Alzahrani²
Department of Computer Science
College of Computers and Information Technology
Taif University

Hatim Alsuwat³
Department of Computer Science
College of Computer and Information Systems
Umm Al Qura University

Abstract—Probabilistic graphical models are employed in a variety of areas such as artificial intelligence and machine learning to depict causal relations among sets of random variables. In this research, we employ probabilistic graphical models in the form of Bayesian network to detect coronavirus disease 2019 (denoted as COVID-19) disease. We propose two efficient Bayesian network models that are potent in encoding causal relations among random variable, i.e., COVID-19 symptoms. The first Bayesian network model, denoted as BN_1 , is built depending on the acquired knowledge from medical experts. We collect data from clinics and hospitals in Saudi Arabia for our research. We name this authentic dataset DS_{covid} . The second Bayesian network model, denoted as BN_2 , is learned from the real dataset DS_{covid} depending on Chow-Liu tree approach. We also implement our proposed Bayesian network models and present our experimental results. Our results show that the proposed approaches are capable of modeling the issue of making decisions in the context of COVID-19. Moreover, our experimental results show that the two Bayesian network models we propose in this work are effective for not only extracting casual relations but also reducing uncertainty and increasing the effectiveness of causal reasoning and prediction.

Keywords—Coronavirus disease 2019; COVID-19; artificial intelligence; machine learning; probabilistic graphical models; causal models; Bayesian networks; detection methods

I. INTRODUCTION

Over the last two years, the world has experienced events that changed the whole scenario and would be considered as an extremely dark phase when written as history. This is because of the outbreak of a coronavirus disease 2019 pandemic denoted as COVID-19 that results from the family of viruses and causes issues related to gastrointestinal and respiratory diseases. There are different symptoms associated with the COVID-19 including shortness of breath, cough, fever, and fatigue, etc. It is a challenge to detect COVID-19 as these symptoms are also present in other diseases such as the common cold. The testing or detection of COVID-19 depends on the past or current presence of the SARS-CoV-2. If there is no virus, then the patient cannot be declared as a COVID-19 patient. Moreover, there is a presence of fear in people that can lead to psychological responses.

Medical diagnostic errors, on the other hand, can be considered as the leading error of injury and death. The error in diagnosis of the diseases results in the wrong treatment of the patient, which delays the patient's recovery time, give time for actual diseases to spread, and form another reaction from the wrong medications. There is a need for a strong patient

safety culture to improve the quality and reducing the risks associated with the diagnosis and then the treatment of the patient. For coping with all the mentioned above challenging COVID-19 detection issues, Artificial Intelligence (AI) needs to get involved in order to make the process of detecting the virus become accurate and efficient.

Artificial intelligence and machine learning have been widely used in many real-world applications in recent years. It has been reported that machine learning techniques could be used to deliver automated tools that are able to efficiently predict and detect risks. Artificial intelligence is the way that combines large data, intelligent algorithms, and iterative processing. If artificial intelligence is utilized in the right way, it can be beneficial and powerful. In this regard, this research focuses on the detection of COVID-19 utilizing machine learning, namely using Bayesian networks which is the expanded representation of the probability distribution. Bayesian networks are now established as the indispensable tools of artificial intelligence and are vital in the fields of engineering and science.

Probability theory has provided the base for different engineering and scientific tasks. Artificial intelligence and more specifically machine learning are considered as one of the major fields that have used probability for the development of new algorithms and theorems. In terms of the probabilistic graphical models or PGM's, the Bayesian networks are used at a wide-scale [1]. Bayesian networks are the extensive range class of probabilistic models. Initially, Bayesian networks were the type of statistical model that can show the probability distribution at a complex scale. They are well-developed modeling systems that are useful for monitoring, diagnosing, and making predictions even if uncertainties are present.

In this research, we utilize the soundness of probabilistic graphical models to both build a Bayesian network model using expert's knowledge and learn the structure of the Bayesian network model from authentic data. That is, we propose two probabilistic graphical models that are effective in the context of detecting COVID-19. The proposed probabilistic models are potent in encoding causal relations among random variables, i.e., COVID-19 symptoms, and thus can extract the required knowledge in order to reduce uncertainty and increase effectiveness and accuracy of causal reasoning and prediction.

The main contributions of our research are as follows: We build a naive Bayesian network model using experts' knowledge that is potent in dealing with the uncertainty in

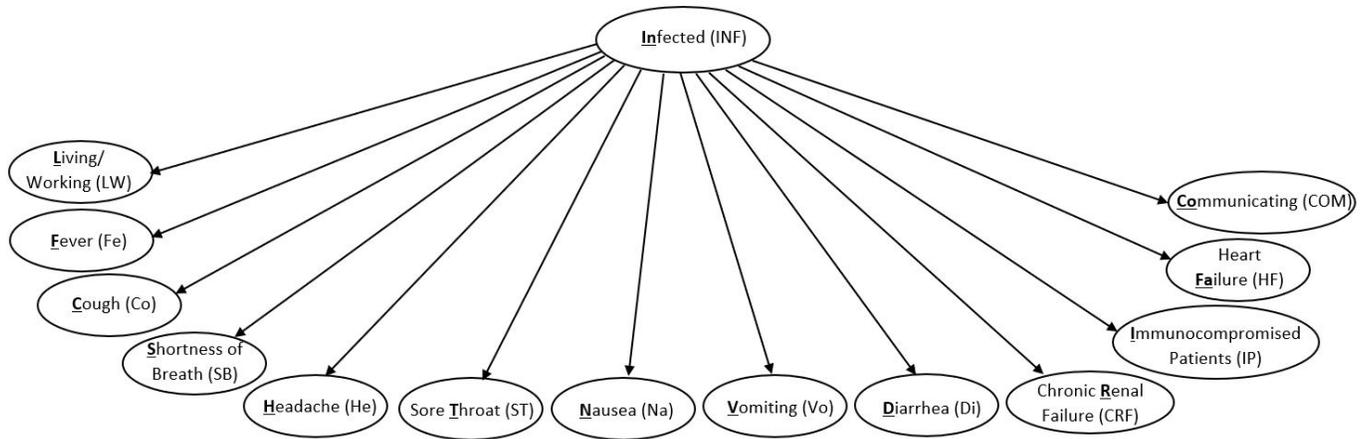


Fig. 1. A Causal Bayesian Network Model that was Built based on the Acquired knowledge from Medical Experts, Denoted as BN_1 . The Direction of the Causal Relations in from the Class Variable to Attributes, i.e., top to Down. Each Node in the Diagnostic Bayesian Network Model has Two States, Either 'yes' or 'no'.

diagnosing the arriving patients in a clinic with COVID-19. We name this causal model BN_1 . We collect data from clinics and hospitals in Saudi Arabia for our research. We name this authentic dataset DS_{covid} . We construct the “best possible” Bayesian network model from the dataset DS_{covid} using Chow-Liu tree approach (we refer the readers to [2], [3], [4] for complete information about Chow-Liu algorithm). We name this causal model BN_2 . We have implemented our approaches and evaluate the correctness of BN_1 using the authentic dataset DS_{covid} . Our results show that the naive Bayesian network model, BN_1 , and the learned model using Chow-Liu tree approach, BN_2 , are identical despite using different ways to recover these causal models. In addition, our experimental results indicate that the two causal models, BN_1 and BN_2 , are promising for reasoning under uncertainty, that is, the proposed models not only are capable to detect COVID-19 positive cases but also can be efficiently used make real-time reasoning and predictions.

The rest of this paper is structured as follows. In Section II, we present our coronavirus diagnostic naive Bayesian network model built using Expert Knowledge. In Section III, we present our coronavirus diagnostic Bayesian network model recovered from the collected authentic data. In Section IV, we present implementations of the proposed causal models and show our empirical results. In Section V, we give a brief overview of background information. In Section VI, we present the conclusion and future work of our research.

II. CORONAVIRUS DIAGNOSTIC NAIVE BAYESIAN NETWORK MODEL FROM EXPERT KNOWLEDGE

In this section, we acquire medical knowledge from experts and then build a Bayesian network model that captures the acquired knowledge. The fundamental purpose is to use the developed Bayesian network model to diagnose the condition of arriving patients suffering from one or some of COVID-19 symptoms.

The acquired medical knowledge is as follows: Coronavirus disease (COVID-19) is an acute respiratory disease that can be detected via investigating either exposure risks or clinical signs

and symptoms. Exposure risks for coronavirus disease increase when communicating with a reported case of coronavirus in the last 14 days before onset of symptoms or living/working in a place identified to be suffering of occurrences of coronavirus in the last 14 days. Clinical signs and symptoms may include one or more of the following:

- (1) Patients suffering from fever above 37.5 degrees Celsius,
- (2) Patients suffering from cough,
- (3) Patients suffering from shortness of breath,
- (4) Patients suffering from sore throat or headache,
- (5) Patients suffering from vomiting, nausea, or diarrhea, and
- (6) Patients suffering from Heart failure, chronic renal failure, and immunocompromised patients.

This knowledge needs to be applied to the situation where a patient is presented in a clinic, and we need to assure whether the patient is infected with coronavirus disease or not. Give the set of symptoms that the patient is suffering from, the doctor would like to know the probability that the patient is infected with coronavirus disease. For instance, if the patients suffers from fever, cough, and shortness of breath, what is the chance that this patient is infected with coronavirus disease and must be quarantined immediately? When getting all information from the patient (i.e., by directly asking the patient about what symptoms the patient suffers from), the developed Bayesian network model will give the final judgement, which will significantly help the doctor and significantly reduce the diagnose time.

Therefore, the goal is to transfer this acquired knowledge into a medical network (i.e., Bayesian network model) that can be used to diagnose arriving patients in a clinic. In order to do so, we first have to determine the variables of the Bayesian network model. In the context of the above given medical knowledge from experts, the variables are either symptoms or conditions, i.e., diseases. There are, indeed, fourteen symptoms and a condition, which is the case whether the patient is infected with coronavirus disease or not. The variables in this Bayesian network model are discrete, i.e., their values are either True or False.

TABLE I. CONDITIONAL PROBABILITY TABLES FOR THE CAUSAL BAYESIAN NETWORK MODEL BN_1

Node	Probability Value(s)		Type
INF	$P(in) = 0.01$	-	Prior Probability
LW	$P(l in) = 0.05$	$P(l in') = 0.01$	Conditional Probability
Fe	$P(f in) = 0.1$	$P(f in') = 0.01$	Conditional Probability
Co	$P(c in) = 0.1$	$P(c in') = 0.01$	Conditional Probability
SB	$P(s in) = 0.1$	$P(s in') = 0.05$	Conditional Probability
He	$P(h in) = 0.02$	$P(h in') = 0.01$	Conditional Probability
ST	$P(t in) = 0.02$	$P(t in') = 0.01$	Conditional Probability
Na	$P(n in) = 0.02$	$P(n in') = 0.01$	Conditional Probability
Vo	$P(v in) = 0.02$	$P(v in') = 0.01$	Conditional Probability
Di	$P(d in) = 0.02$	$P(d in') = 0.01$	Conditional Probability
CRF	$P(r in) = 0.02$	$P(r in') = 0.01$	Conditional Probability
IP	$P(i in) = 0.02$	$P(i in') = 0.01$	Conditional Probability
HF	$P(fa in) = 0.02$	$P(fa in') = 0.01$	Conditional Probability
COM	$P(co in) = 0.05$	$P(co in') = 0.01$	Conditional Probability

The structure of the acquired medical knowledge from experts is represented by the qualitative naive Bayesian network model shown in Fig. 1. There are two layers in the proposed Bayesian network model. The edges in this model go from the first layer, the condition or the variable class layer, to the second layer, the attributes or symptoms layer. One should note that give the condition or the class variable in this Bayesian network model, every two attributes are d-separated, i.e., independent.

We obtain the Conditional Probability Tables (CPTs) from medical experts depending on their subjective beliefs that reflect the practical experience they have gained. The assessment of the prior probability of the class variable and the assessment of the probability of each attribute conditional on its parents, i.e., conditional on the class variable, are give in Table I. We use the notation f to indicate a positive response on the node Fe , 'Fever' and f' to indicate a negative response on the node Fe , 'Fever'. The notation $P(f)$ stands for $P(Fe = f)$. Similar notations are applied to each node in the developed Bayesian network model as follows:

- in stands for a positive response on the node INF ;
- l stands for a positive response on the node LW ;
- c stands for a positive response on the node Co ;
- s stands for a positive response on the node SB ;
- h stands for a positive response on the node He ;
- t stands for a positive response on the node ST ;
- n stands for a positive response on the node Na ;
- v stands for a positive response on the node Vo ;
- d stands for a positive response on the node Di ;
- r stands for a positive response on the node CRF ;
- i stands for a positive response on the node IP ;
- fa stands for a positive response on the node HF ; and
- co stands for a positive response on the node COM .

One should note that Table I includes only positive probabilities while negative probabilities can be easily derived from this table. For instance, the probability of heart failure given that the patient has coronavirus disease is 0.02, formally can be

written as $P(fa|in) = 0.02$. One can derive the probability of not having heart failure given that the patient has coronavirus disease, which is $1 - 0.02 = 0.98$, formally can be written as $P(fa'|in) = 0.98$.

The proposed causal model shows that the joint probability can be written as follows:

$$P(INF, LW, Fe, Co, SB, He, ST, Na, Vo, Di, CRF, IP, HF, COM) = P(INF) P(LW|INF) P(Fe|INF) P(Co|INF) P(SB|INF) P(He|INF) P(ST|INF) P(Na|INF) P(Vo|INF) P(Di|INF) P(CRF|INF) P(IP|INF) P(HF|INF) P(COM|INF)$$

The joint probability is the product of class variable prior probability times all conditional probabilities of all other variables. On the other hand, if we want to calculate the probability that a patient has coronavirus disease given that this patient has fever and has communicated with a person who was reported to have coronavirus disease, we would do the following:

$$P(INF|Fe, COM) = P(INF, Fe, COM)/P(Fe, COM)$$

For more information about this inferential equation, we refer the readers to Bayes theorem [4]. The implementation of the proposed causal model in this section and more experiments will be presented in Section IV.

III. CORONAVIRUS DIAGNOSTIC BAYESIAN NETWORK MODEL FROM DATA

In this section, we learn a Bayesian network model from data. We have collected data from clinics and hospitals in Saudi Arabia for our research. We name our dataset DS_{covid} , which contains the data for 60,000 patients. The dataset DS_{covid} is available at https://emadalsuwat.github.io/resources/Covid19_dataset.dat.

The dataset DS_{covid} consists of cases, a.k.a. tuples. Each case represent a unique object in the collection process. Objects or cases are patients in this research context. Each case in our dataset is characterized by fourteen variables. That is, for each tuple in our dataset, there are fourteen cells, a.k.a

data items, that are required to be filled such that each cell corresponds to a defined random variable. The dataset DS_{covid} is a discrete dataset, which means that the answer to each variables is either “yes” or “no”. For example, case #1 contains information about patient #1 in our dataset and so on. If patient #1 has fever, we shall see “yes” in the cell that corresponds to the random variable “Fever” and so on. If the value of a cell in a tuple is missing (i.e., there is no reported information), the cell is left blank.

We construct the “best possible” tree-shaped Bayesian network model from the learning dataset DS_{covid} . We then determine whether or not the best possible tree-shaped belief network learned using the collected dataset DS_{covid} is the naive Bayesian network model. If so, we say that the naive Bayesian network model is the best possible tree-shaped belief network that fits the collected dataset DS_{covid} .

We construct the “best possible” Bayesian network model from the dataset DS_{covid} using Chow-Liu tree approach, which is an approximation of the probability distribution with tree-shaped belief network. In Chow-Liu tree approach, the links are directed away from the root of the tree-shaped Bayesian network model.

Let DS_{covid} be a discrete dataset over variables $\{V_1, V_2, \dots, V_n\}$. Let $BN_2 = (V, E)$ be the Bayesian network model constructed from the dataset DS_{covid} using Chow-Liu tree approach. Let $V = \{V_1, V_2, \dots, V_n\}$ and $E = \{(v_1, v_2) : v_1, v_2 \in V\}$ be the set of nodes and edges, respectively, in the belief network BN_2 . The causal network BN_2 can be learned as follows [3]:

Step 1: For each pair of vertices $(V_i, V_j) \in BN_2$, we calculate the mutual information as follows.

$$MI(V_i, V_j) = \sum_{V_i, V_j} P(V_i, V_j) \log_2 \left(\frac{P(V_i, V_j)}{P(V_i)P(V_j)} \right) \quad (1)$$

Step 2: Construct the mutual information weighted graph, which we denote as G_{MI} , over the set of vertices $\{V_1, V_2, \dots, V_n\}$, where the weight of the edge (V_1, V_2) in G_{MI} is $MI(V_i, V_j)$.

Step 3: Given the G_{MI} graph, build a maximum weight spanning tree, which we denote as G_{SPT} .

Step 4: Direct the resulting maximum spanning tree G_{SPT} . This can be accomplished by selecting any node in the resulting graph as the root of the tree and then set the direction of the remaining edges to point outward from it. We denote the resulting tree as BN_2 .

In this section, we have shown that it is feasible to estimate the Bayesian network model from medical recorders of patients, i.e., from the collected data. That is, the result of the formal steps presented above is the Bayesian network directed acyclic graph. An important point to be mentioned is that one should use Expectation–Maximization algorithm (a.k.a EM algorithm) to estimate the conditional probability distributions of the constructed causal model, BN_2 , from data. For more information about the EM algorithm, we refer the readers to [5]. The implementation of learning the structure of the causal model BN_2 from the collected data using Chow-Liu tree approach as proposed in this section and more experiments will be presented in Section IV.

IV. IMPLEMENTATIONS OF CAUSAL MODELS AND EXPERIMENTAL RESULTS

In this section, we implement our proposed two Bayesian network models, BN_1 and BN_2 , using *HuginTM Researcher* 8.4. We then present some experimental results to show the practical capability of the proposed causal models.

A. Implementation of the Causal Model BN_1

In this subsection, expert knowledge presented in section II can be modeled by the implemented Bayesian network BN_2 shown in Figure 2. The Bayesian network BN_2 consists of fourteen nodes: Infected (INF), Living/Working (LW), Fever (Fe), and so on. Each node has two states: “Yes” or “No”. For instance, if the presented patient has Fever, we choose “Yes” as the answer for this symptom and vice versa.

Fig. 2 shows the running mode of the Bayesian network model BN_1 . The figure is split into two parts: Node List Pane on the left hand side and Network Pane on the right hand side. It is important to point out that we have expanded all nodes getting ready to enter evidence or facts. For instance, if a patients has Nausea, we then just click on the state “Yes” under the node Nausea in the Node List Pane and so on.

One should note that any two variables in the Bayesian network model BN_2 are independent given the state of the hypothesis variable. That is, if the state of the hypothesis variable Infect is given, then any two variables of the symptoms nodes are independent. Another important note is that causality is obvious and clear in the implemented model BN_2 . That is, the class variable, Infect in this case, is the cause of the other variables, i.e., symptoms. Entering information about the state symptoms (either “Yes” or “No”) will increase/decrease the joint probability that the patient has COVID-19 disease.

B. Implementation of the Causal Model BN_2

In this subsection, we use the four steps presented in Section III to learn the structure of the Bayesian network model from data. That is, we use the Chow-Liu algorithm to recover the structure of BN_2 from the collected dataset DS_{covid} . We present the first few lines of the authentic dataset DS_{covid} in Fig. 3. The data file of the dataset DS_{covid} we use in this implementation consists of 60,000 cases, i.e., tuples.

We use *HuginTM Researcher* 8.4 to learn the Bayesian network model BN_2 from the dataset DS_{covid} . We use *HuginTM* implementation of Chow-Liu algorithm with the default significance level at 0.05 [6]. It is important to keep the level of significance at the default setting and not to increase this value in order to reduce the running time of the Bayesian network structure learning algorithm to the best possible running time.

The recovered Bayesian network model from the dataset DS_{covid} is as shown in Fig. 4. When comparing the result of Chow-Liu algorithm, BN_2 , with the result of the original model we built using experts’ knowledge, BN_1 , we observe the following: 1) All nodes have been recovered correctly, 2) All links have been recovered correctly, and 3) There is no missing or misoriented links, i.e., the original and the learned Bayesian network models indicate the same causal

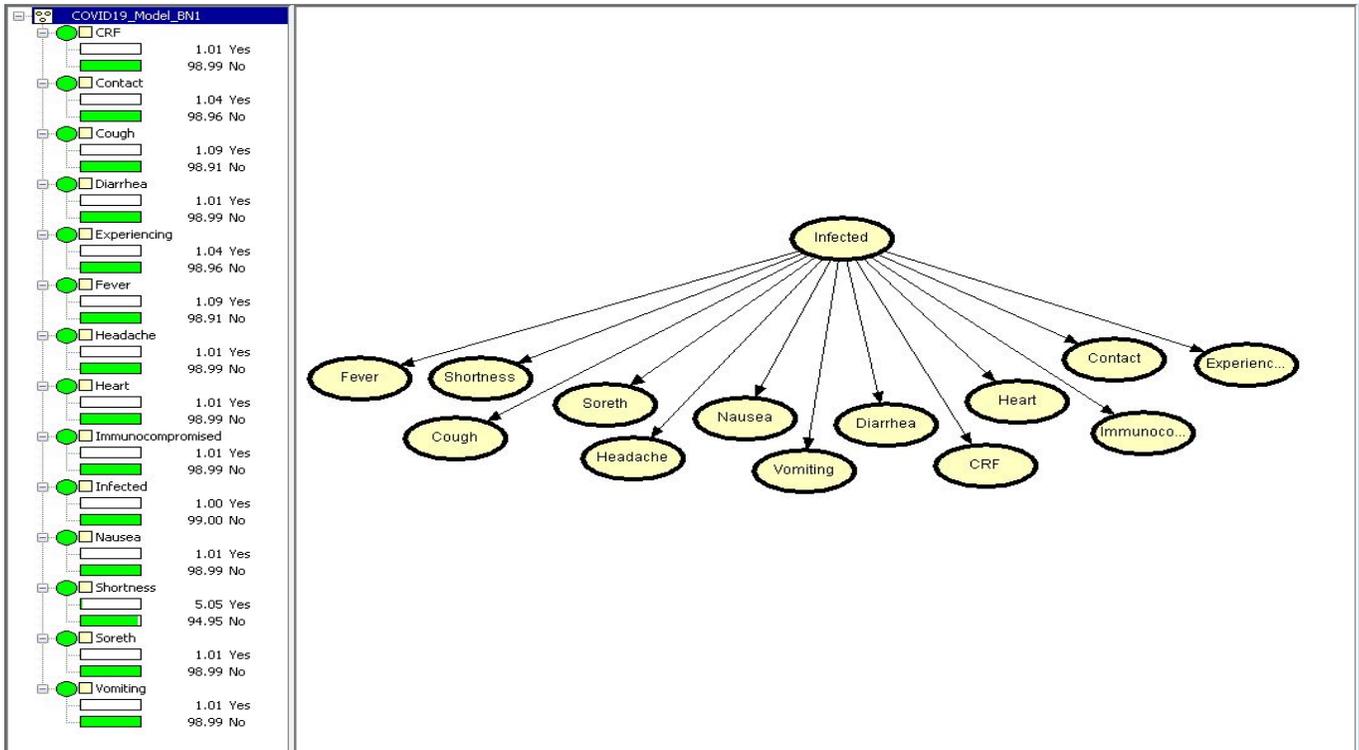


Fig. 2. The Implemented Bayesian Network Model BN_1 .

#	Fever	Headache	Soreth	Nausea	Vomiting	CRF	Heart	Infected	Immunoco...	Diarrhea	Shortness	Cough	Experiencing	Contact
0	No	No	No	No	No	No	No	No	No	No	No	No	No	No
1	No	No	No	No	No	No	No	No	No	No	No	No	No	No
2	No	No	No	No	No	No	No	No	No	No	No	No	No	No
3	No	No	No	No	No	No	No	No	No	No	No	No	No	No
4	No	No	No	No	No	No	No	No	No	No	No	No	No	No
5	No	No	No	No	No	No	No	No	No	No	No	No	No	No
6	No	No	No	No	No	No	No	No	No	No	No	No	No	No
7	No	No	No	No	No	No	No	No	No	No	No	No	No	No
8	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No
9	No	No	No	No	No	No	No	No	No	No	No	No	No	No
10	No	No	No	No	No	No	No	No	No	No	No	No	No	No
11	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No
12	No	No	No	No	No	No	No	No	No	No	No	No	No	No
13	No	No	No	No	No	No	No	No	No	No	No	No	No	No
14	No	No	No	No	No	No	No	No	No	No	No	No	No	No
15	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No

Fig. 3. The First Few Cases (Tuples) of the Data File of the Authentic Dataset DS_{covid} .

relationships. We thus conclude that Chow-Liu structure learning algorithm was able to perfectly recover the structure of the Bayesian network model BN_2 from the collected dataset DS_{covid} .

C. Reasoning using the Proposed Models

In this subsection, we employ the implemented Bayesian network models, BN_1 and BN_2 , for reasoning under uncertainty. That is, given some symptoms that a presented patient in a clinic suffers from, the medical practitioner can then utilize one of these models to find out the probability that this patients is infected with COVID-19.

Using BN_1 for Reasoning under Uncertainty: Assume that a patient presented in a clinic where this patient has been in contact with a person who has COVID-19 and suffers from Fever and Sore Throat. The medical practitioner employs the

Bayesian network model BN_1 to calculate the inferential (conditional) probability $P(INF = "Yes" | Fe = "Yes", ST = "Yes", COM = "Yes")$, which can be calculated by dividing the joint probability $P(INF = "Yes", Fe = "Yes", ST = "Yes", COM = "Yes")$ by the probability of evidence $P(Fe = "Yes", ST = "Yes", COM = "Yes")$.

The medical practitioner can simply enter the evidence on the nodes "Fever", "Soreth" and "Contact". That is, the medical practitioner chooses the state "Yes" on these three nodes as shown in Fig. 5.

We eventually propagate these pieces of evidence in order to calculate the conditional probability. The conditional probability is calculated by dividing the joint probability by the probability of the evidence. BN_1 can be used to obtain the joint probability as follows: $P(INF = "Yes", Fe = "Yes", ST = "Yes", COM = "Yes") = 8.5916e - 7$.

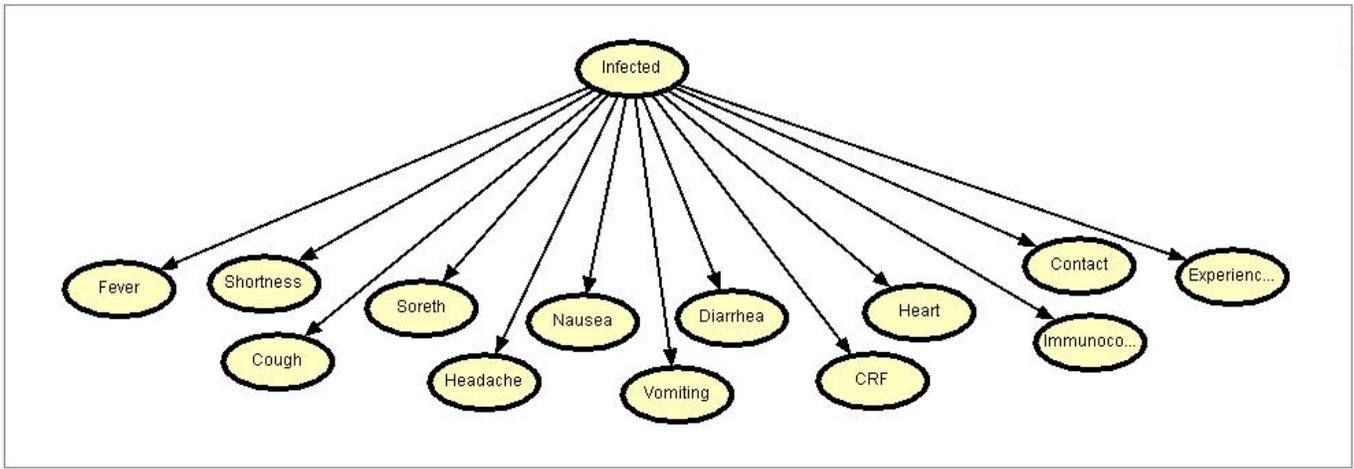


Fig. 4. The Result of using Chow-Liu Algorithm to Learn the Bayesian Network Model BN_2 from the Authentic Dataset DS_{covid} .

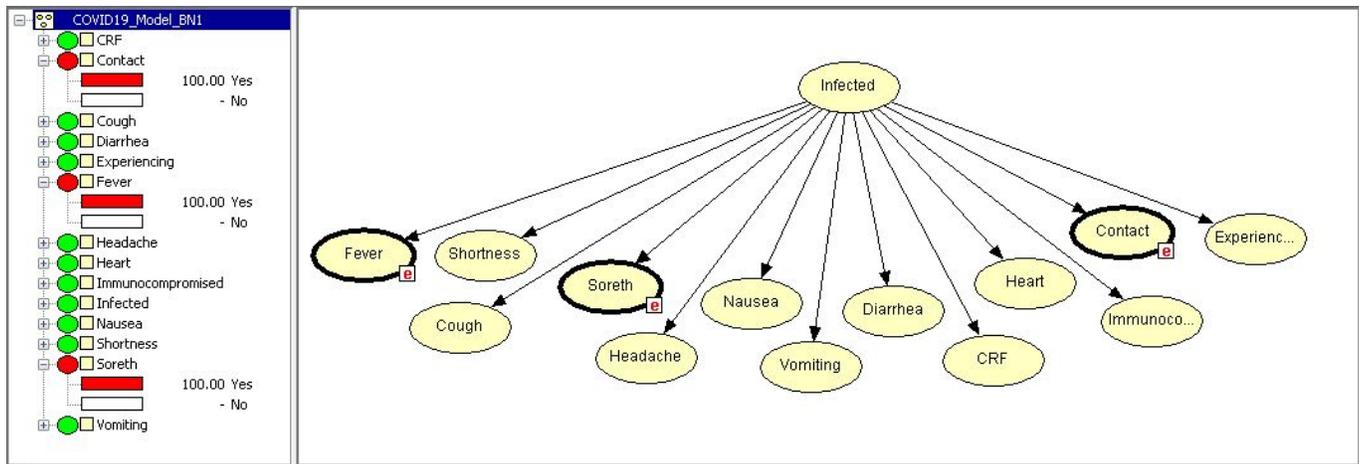


Fig. 5. The Bayesian Network Model BN_1 After Entering the Evidence on the Three Nodes: Fever, Soreth and Contact.

BN_1 also can be used to obtain the probability of the evidence as follows: $P(Fe = "Yes", ST = "Yes", COM = "Yes") = 1.99e - 6$. Thus, the conditional probability can be calculated as follows: $P(INF = "Yes" | Fe = "Yes", ST = "Yes", COM = "Yes") = \frac{8.5916e-7}{1.99e-6} = 0.432$.

Using BN_2 for Reasoning under Uncertainty: Assume that an arriving patient presented at a clinic suffers from the following symptoms: Fever and Shortness of Breath, and Cough. In addition, the patient has reported that there was a communication with a reported case of coronavirus in the last 14 days before onset of symptoms. The medical practitioner thus employs the Bayesian network model BN_2 to calculate the inferential probability $P(INF = "Yes" | Fe = "Yes", SB = "Yes", Co = "Yes", COM = "Yes")$, which can be calculated by dividing the joint probability $P(INF = "Yes", Fe = "Yes", SB = "Yes", Co = "Yes", COM = "Yes")$ by the probability of evidence $P(Fe = "Yes", SB = "Yes", Co = "Yes", COM = "Yes")$.

The medical practitioner can simply provide the evidence on the nodes "Fever", "Shortness", "Cough", and "Contact". That is, the medical practitioner selects the state "Yes" on these

four nodes as shown in Fig. 6.

These pieces of evidence are then propagated with the ultimate goal of calculating the conditional probability. The conditional probability is calculated as follows: 1) Employ the Bayesian network model BN_2 to obtain the joint probability. Using the model BN_2 , the joint probability $P(INF = "Yes", Fe = "Yes", SB = "Yes", Co = "Yes", COM = "Yes") = 4.0411e - 7$. 2) Employ the Bayesian model BN_2 to find the probability of the evidence. Using BN_2 , the probability of the evidence $P(Fe = "Yes", SB = "Yes", Co = "Yes", COM = "Yes") = 5.495e - 7$. 3) Calculate the conditional probability as follows: $P(INF = "Yes" | Fe = "Yes", SB = "Yes", Co = "Yes", COM = "Yes") = \frac{4.0411e-7}{5.495e-7} = 0.735$.

we have shown that our implemented causal models are efficient in detecting patients with COVID-19 disease given the set of symptoms they suffer from. It is important to point out that our causal models are an outstanding technique to control the disease. Moreover, these models are great as they enable medical practitioners to take decisions under uncertainty.

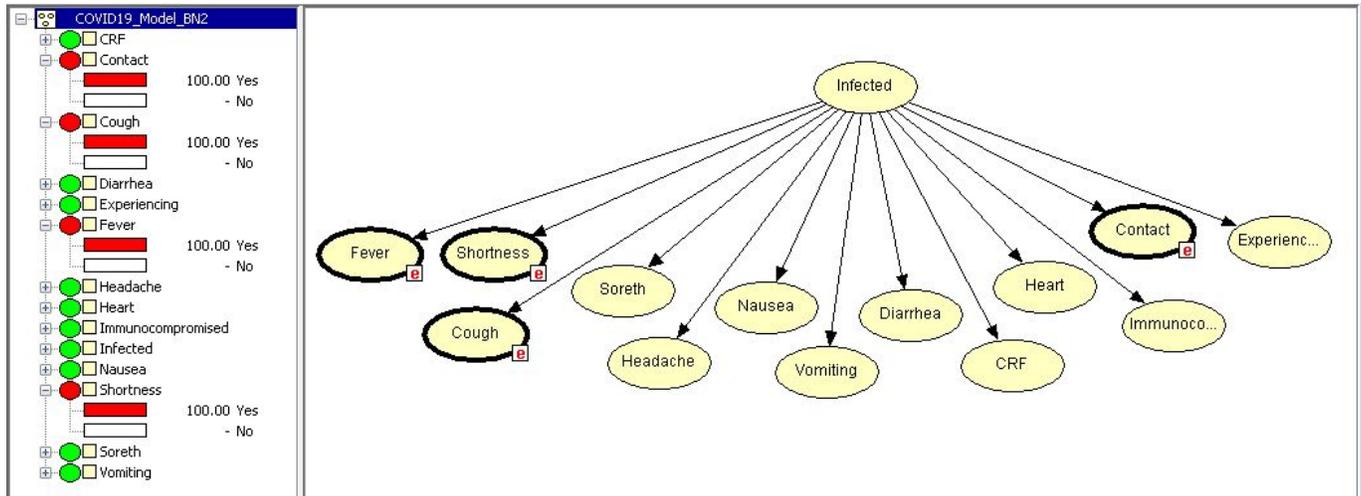


Fig. 6. The Bayesian Network Model BN_2 After Entering the Evidence on the Three Nodes: Fever, Shortness, Cough, and Contact.

V. BACKGROUND INFORMATION

In this section, we will give a brief overview of the use of artificial intelligence, namely Bayesian network models, a favorable tool for decision making. More detailed information about probabilistic graphical models and Bayesian network can be found in [4], [7], [8], [9].

Artificial intelligence is a technology that became prevalent in all different kinds of businesses. However, its use in healthcare has only started in recent years. This increased use is because the technologies have the transforming potential in different aspects of healthcare [10], [11], [12]. Machine learning is a statistical technique for learning the training models with data and for fitting models to data. It is one of the common forms of artificial intelligence. It can be evident from the fact that 63% of corporates that use artificial intelligence were employing machine learning in their corporates and businesses [10]. There are different applications of machine learning in healthcare. For instance, precision medicine- predicting the best treatment method on the basis of the patient attributes and treatment context [10], [13].

The use of probabilistic graphical models has started in 1985 when Pearl (1985) combines the probabilities theories and graphs for obtaining the comprehensible representation of the probability distribution. Further improvements in the specific models can be observed by the research of Lauritzen and Speigelhalter (1988) by Cooper and Herskovits (1992) [1], [14], [4]. The need to overcome difficult and complex problems that exist in the real world has increased the demand for using effective meta-heuristic algorithms that are able to get solutions by performing a momentous search for getting possible solutions.

The probabilistic graphical models have traditionally been used for different purposes. For instance, research indicates the use of probabilistic graphical models for dealing with age estimation of living people [15]. The importance of age estimation of the living individual has been raised because of the civil, criminal, and administrative judicial situations, and it has a significant role in the medicolegal and forensic services all around the globe. Other researchers used the Bayesian

approach for determination of clavicular epiphysis [15]. Moreover, it has been suggested to use probabilistic graphical models for evolutionary algorithm and solving complex problems. The evolutionary computation was performed perfectly using the probabilistic graphical models [15].

The diagnostic is considered as one of the major practices in artificial intelligence. Research shows several probabilistic diagnostic methods that can be implemented in the upcoming medicine research. We refer the readers to [16] for a complete list of possible uses of artificial intelligence in medicine. Another widely use of Bayesian networks other than medical research is detecting spam (a.k.a. spam filtering) [17]. Bayesian networks has also been employed and showed outstanding results in intrusion detection. Recent survey on intrusion detection using machine learning can be found in [18], [19].

Another significant application of Bayesian networks is predicting the type of hematological malignancies [20]. Bayesian networks have helped scientists to model Acute Myeloid Leukemia (AML), which is a subtype of blood cancer. Thus, research has shown that Bayesian network was not only able to model and predict many subtypes of cancer but also help physicians choose the most effective treatment. Bayesian network performance in prediction and reasoning under uncertainty was compared in the literature with other machine learning types. Indeed, Bayesian networks was proved to be a very efficient predictive model in finding probabilistic graphical relationships among random variables. Recently, researchers have employed Bayesian networks to improve screening of COVID-19 cases [21]. The above literature has given model examples that suggests the use of Bayesian networks in healthcare sector for detecting and differentiating between different and complex medical conditions. Previous work for detecting COVID-19 has focused on different machine learning tools except Bayesian networks even though Bayesian networks are very powerful in reasoning under uncertainty [22], [23], [24]. In this work, we explore the feasibility of using the probabilistic graphical models in the form of Bayesian networks to build an effective diagnosis system for controlling and detecting COVID-19 disease. The proposed

Bayesian models are useful in current situation for increasing the speed of diagnosis, which can be helpful in treating and coping with this disease efficaciously.

VI. CONCLUSION AND FUTURE WORK

In this research, we demonstrated the applicability of utilizing probabilistic graphical models in the form of Bayesian networks to detect patients with COVID-19 disease. We proposed two diagnostic Bayesian network models: BN_1 and BN_2 . We implemented the proposed Bayesian network models and conducted our experimental results. Our results indicate this research succeeded in designing and building two effective Bayesian network models to tackle the issue of dealing with arriving patients who may have COVID-19 disease. Our experimental results also show that our proposed Bayesian network models are effective in not only extracting the casual relations but also reducing uncertainty and increasing the effectiveness of causal reasoning and prediction.

In our future work, we aim to extend Bayesian network models we built in this research in order to address more diseases beside COVID-19. In particular, we aim to extend our Bayesian network models to handle rare subtypes of blood cancer as such diseases are very unlikely to get funded or find interested researchers.

ACKNOWLEDGMENT

This project was supported by the Deanship of Scientific Research at Taif University. The authors would like to thank Taif University for providing financial support for this research project (Project number: 1-441-64)

REFERENCES

- [1] P. Larrañaga, H. Karshenas, C. Bielza, and R. Santana, "A review on probabilistic graphical models in evolutionary computation," *Journal of Heuristics*, vol. 18, no. 5, pp. 795–819, 2012.
- [2] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.
- [3] T. D. Nielsen and F. V. Jensen, *Bayesian networks and decision graphs*. Springer Science & Business Media, 2009.
- [4] R. E. Neapolitan, "Learning bayesian networks, vol. 38 pearson prentice hall," *Upper Saddle River, NJ*, 2004.
- [5] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.
- [6] A. L. Madsen, F. Jensen, U. B. Kjaerulff, and M. Lang, "The hugin tool for probabilistic graphical models," *International Journal on Artificial Intelligence Tools*, vol. 14, no. 03, pp. 507–543, 2005.
- [7] A. Dawid, R. Cowell, S. Lauritzen, and D. Spiegelhalter, "Probabilistic networks and expert systems." Springer-Verlag, 1999.
- [8] P. Congdon, *Bayesian statistical modelling*. John Wiley & Sons, 2007, vol. 704.
- [9] R. G. Cowell, P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter, *Probabilistic networks and expert systems: Exact computational methods for Bayesian networks*. Springer Science & Business Media, 2006.
- [10] T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," *Future healthcare journal*, vol. 6, no. 2, p. 94, 2019.
- [11] G. Acampora, D. J. Cook, P. Rashidi, and A. V. Vasilakos, "A survey on ambient intelligence in healthcare," *Proceedings of the IEEE*, vol. 101, no. 12, pp. 2470–2494, 2013.
- [12] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, "Artificial intelligence in healthcare: past, present and future," *Stroke and vascular neurology*, vol. 2, no. 4, 2017.
- [13] R. Hamamoto, K. Suvarna, M. Yamada, K. Kobayashi, N. Shinkai, M. Miyake, M. Takahashi, S. Jinnai, R. Shimoyama, A. Sakai *et al.*, "Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine," *Cancers*, vol. 12, no. 12, p. 3532, 2020.
- [14] P. Spirtes, C. N. Glymour, R. Scheines, and D. Heckerman, *Causation, prediction, and search*. MIT press, 2000.
- [15] E. Sironi, M. Gallidabino, C. Weyermann, and F. Taroni, "Probabilistic graphical models to deal with age estimation of living persons," *International journal of legal medicine*, vol. 130, no. 2, pp. 475–488, 2016.
- [16] P. Szolovits, *Artificial intelligence in medicine*. Routledge, 2019.
- [17] B. Nelson, M. Barreno, F. J. Chi, A. D. Joseph, B. I. Rubinstein, U. Saini, C. Sutton, J. Tygar, and K. Xia, "Misleading learners: Co-opting your spam filter," in *Machine learning in cyber trust*. Springer, 2009, pp. 17–51.
- [18] C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, and W.-Y. Lin, "Intrusion detection by machine learning: A review," *expert systems with applications*, vol. 36, no. 10, pp. 11 994–12 000, 2009.
- [19] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications surveys & tutorials*, vol. 18, no. 2, pp. 1153–1176, 2015.
- [20] R. Agrahari, A. Foroushani, T. R. Docking, L. Chang, G. Duns, M. Hudoba, A. Karsan, and H. Zare, "Applications of bayesian network models in predicting types of hematological malignancies," *Scientific reports*, vol. 8, no. 1, pp. 1–12, 2018.
- [21] S. Eyheramendy, P. A. Saa, E. A. Undurraga, C. Valencia, C. Lopez, L. Mendez, J. Pizarro-Berdichevsky, A. Finkelstein-Kulka, S. Solari, N. Salas *et al.*, "Improved screening of covid-19 cases through a bayesian network symptoms model and psychophysical olfactory test," *medRxiv*, 2021.
- [22] I. Arpaci, M. Al-Emran, G. Marques, M. Al-Kabi *et al.*, "A survey of using machine learning algorithms during the covid-19 pandemic," *Emerging Technologies During the Era of COVID-19 Pandemic*, vol. 348, pp. 1–8, 2021.
- [23] S. Abir, S. N. Islam, A. Anwar, A. N. Mahmood, and A. M. T. Oo, "Building resilience against covid-19 pandemic using artificial intelligence, machine learning, and iot: A survey of recent progress," *IoT*, vol. 1, no. 2, pp. 506–528, 2020.
- [24] J. Nayak, B. Naik, P. Dinesh, K. Vakula, B. K. Rao, W. Ding, and D. Pelusi, "Intelligent system for covid-19 prognosis: A state-of-the-art survey," *Applied Intelligence*, vol. 51, no. 5, pp. 2908–2938, 2021.

An Optimized Artificial Neural Network Model using Genetic Algorithm for Prediction of Traffic Emission Concentrations

Akibu Mahmoud Abdullah¹, Raja Sher Afgun Usmani²,
Thulasyammal Ramiah Pillai³, Mohsen Marjani⁴
School of Computer Science and Engineering,
Taylor's University,
Selangor, Malaysia

Ibrahim Abaker Targio Hashem⁵
College of Computing and Informatics,
Department of Computer Science,
University of Sharjah,
27272 Sharjah, UAE

Abstract—Global warming and climate change have become universal issues recently. One of the leading sources of climate change is automobiles. Automobiles are the prime source of air pollution in urban areas globally. This has resulted in a problematic and chaotic state in the development of an automatic traffic management system for capturing and monitoring vehicles' hourly and daily passage. With the significant advancement of sensor technology, atmospheric information such as air pollution, meteorological, and motor vehicle data can be harvested and stored in databases. However, due to the complexity and non-linear associations between air quality, meteorological, and traffic variables, it is difficult for the traditional statistical and mathematical models to analyze them. Recently, machine learning algorithms in the field of traffic emissions prediction have become a popular tool. Meteorological and traffic variables influence the variation and the trend of the traffic pollutants. In this paper, an optimized artificial neural network (OANN) was developed to enhance the existing artificial neural network (ANN) model by updating the initial weights in the network using a Genetic Algorithm (GA). The OANN model was implemented to predict the concentration of CO , NO , NO_2 , and NO_x pollutants produced by motor vehicles in Kuala Lumpur, Malaysia. OANN was compared with Artificial Neural Network (ANN), Random Forest (RF), and Decision Tree (DT) models. The results show that the developed OANN model performed better than the ANN, RF, and DT models with the lowest MSE values of 0.0247 for CO , 0.0365 for NO , 0.0542 for NO_2 , and 0.1128 for NO_x . It can be concluded that the developed OANN model is a better choice in predicting traffic emission concentrations. The developed OANN model can help environmental agencies monitor traffic-related air pollution levels efficiently and take necessary measures to ensure the effectiveness of traffic management policy. The OANN model can also help decision-makers mitigate traffic emissions to protect citizens living in the neighborhood of highways.

Keywords—Optimized Artificial Neural Network (OANN); Genetic Algorithm; traffic emissions

I. INTRODUCTION

Global warming and climate change have become universal issues recently [1, 2]. One of the leading sources of climate change is emissions from motor vehicles. Carbon monoxide (CO), nitrogen dioxide (NO_2), carbon dioxide (CO_2), and nitrogen monoxide (NO) are among the significant risk to human health and the environment, which can be emitted by motorized vehicles [3]. Road transport emissions exposure

can increase the risk of lung cancer [4], respiratory and cardiovascular effects [5, 6], pulmonary, chronic diseases [7] and mortality [8, 9]. In 2015 air pollution, in general, are responsible for over 6 million deaths in the world [10, 11]. There is more than 400,000 premature death in Europe [12], and around 7 million worldwide [13].

Automobiles are the prime sources of air pollution in city areas universally. For example, 80.49% of emissions in Beijing, China, were produced by motor vehicles [14]. It was also found that 80% of air pollution in the Lima Metropolitan Area was produced from automobiles [13]. In China, 85% of emissions were from transport [15], while in the United Kingdom was 92% [16], and 75% in Malaysia [17], but in the United States, automobiles are responsible for emitting 57% of air pollution [7]. Many other studies also show that 92% of CO and 65% of hydrocarbon (HC) pollutants were emitted from the transportation activities in Shanghai [18], and 60% of NO_x and PM emissions were from heavy-duty trucks in China [19].

The variation and trend of air pollution strongly depend on meteorological parameters and traffic characteristics [20]. A study conducted in Shiraz, Iran, confirmed that meteorological conditions increase the air pollution level [21]; a similar result was also found in Karaj, Iran [22]. A relationship between air pollution and meteorological condition was studied in Beijing and Nanjing, China. The study reveals that air pollution concentrations depend on meteorological factors [23]. Research in Linfen city, China, shows negative correlations between air pollutants and meteorological parameters [24]. A study was conducted in Penang, Malaysia, to investigate the sources contributing to air pollution concentrations. Five air pollutants were investigated. The result shows that a negative correlation was found between relative humidity with CO , O_3 , SO_2 , PM_{10} , and NO_2 .

A negative correlation was found between wind speed and CO , SO_2 , and NO_2 , but positive relation with PM_{10} and O_3 . Whenever temperature increases, the O_3 , NO_2 , CO , and PM_{10} pollutants increase too, but SO_2 decrease [25]. A study on O_3 variability due to meteorological parameters was investigated in Selangor, Malaysia. The result shows that wind speed, temperature, relative humidity, and wind direction significantly impact O_3 concentrations [26]. [27]

studied air pollution variation due to meteorological in four areas in Malaysia, namely, Petaling Jaya, Cheras, Shah Alam, and Klang. The result reveals that meteorological parameters influenced the seasonal trend of air pollution. Association between air pollution and traffic characteristics has also been investigated. A study was conducted to compare the impact of traffic volume on air pollution levels during COVID-19 in Italy. Data from 2017 to 2018 before COVID-19 and 2020 data during COVID-19 were used. The result reveals that traffic volume significantly impacts PM_{10} , NO_x , NO , and NO_2 concentrations [28].

A study of [29] investigates the exposure of air pollution produced by vehicles on cyclists in Brazil. The research indicates an increase in motor vehicles during peak hours in the morning and evening. The expansion of the vehicle increases the level of air pollution. In addition, air pollution level due to traffic characteristics was studied in Japan [20]. The result indicates that the low speed of the vehicle increases the pollution level. Similarly, traffic volume and congestion increase the emission of air pollution in Kyoto, Japan. It was also found that trucks are the main contributor of PM and NO_x emissions. Furthermore, a study in Kuala Lumpur was conducted to investigate the effect of traffic characteristics on the air pollution level. The study reveals that air pollution level strongly depends on fuel consumption, traffic volume, vehicle speed, and waiting time on the road. The result also shows that lower traffic congestion reduces the level of air pollution in Kuala Lumpur [30].

With the rapid advancement of sensor technology, atmospheric information such as air pollution, meteorology, and motor vehicle data can be collected and stored in databases. Due to the complexity and non-linear associations that exist between air quality, meteorological, and traffic variables, it is difficult for traditional statistical and mathematical models to analyze them [31, 32]. Lately, the usage of machine learning algorithms such as long short term memory, random forest, support vector machine, decision tree, and artificial neural network (ANN) in traffic-related air pollution prediction has become popular [33]. ANN model appeared to be the most used model for predicting traffic emissions because it reduces time, cost, and complexity. It also provides fast and accurate prediction with less error and provides prediction values closer to the observed values [34]. ANN can solve complex multidimensional variables and non-linear problems related to traffic emission concentrations due to meteorological conditions and traffic features [34, 35].

In this paper, an optimized artificial neural network (OANN) was developed to enhance the existing artificial neural network (ANN) model by updating the weights in the network using a Genetic Algorithm (GA). The OANN model was implemented to predict the concentration of CO , NO , NO_2 , and NO_x pollutants produced by motor vehicles in Kuala Lumpur, Malaysia. The remaining structure of the paper is given as follows. Section II discusses the related work on vehicle emissions prediction using the ANN model. Section III presents the methodology used in this study. Section IV presents the result and the comparison with existing machine learning models for evaluation. Finally, the conclusion is discussed in Section V.

II. RELATED WORK

Motor vehicles are producing harmful pollutants that disperse to the atmosphere [33]. These pollutants have significant impact on human health [36]. Several statistical models have been developed for predicting the traffic emissions concentrations at intersection, canyon, street, near the school, and many other locations. However, these statistical models could not predict the emissions rate due to the variability and influence of meteorological variables and traffic parameters [31]. Machine learning models have recently been applied. These models were able to predict the concentrations of emitted pollutants from motor vehicles. ANN has become the most popular model for predicting traffic-related air pollution [34]. These models are highly dependent on the independent variables provided in the study. There is a lack of either meteorological data or traffic data in many studies [37]. Several studies suggested that meteorological and traffic variable influence the trend and variation of traffic pollutants [33], such as relative humidity, temperature, wind direction, and wind speed, [38], traffic volume, vehicle speed [39], types of the vehicle, etc. [40].

The variables mentioned above are needed to predict emission levels. Still, it is not always available [37], for example, prediction of carbon monoxide (CO) concentrations at Jiyan Ave and Shuanglong intersections was conducted by [41] using Gated Recurrent Unit (GRU) neural networks based in the absence of a meteorological dataset. The accuracy of the model's performance was found good with a root mean squared error (RMSE) of 0.088 and mean absolute error (MAE) of 0.056. [42] proposed machine learning algorithms by comparing and selecting the best model with good performance to reduce the effect of Greenhouse gas (GHG) emitted by passenger vehicles on climate change in Canada. Artificial Neural Network (ANN) shows better performance over the other machine learning algorithms with RMSE of 0.442 and MAE of 0.347. This study also lacks the meteorological dataset. ANN was developed by [43] to predict the emission of CO , CO_2 , NO_x , and HC from a liquefied natural gas bus in Zhenjiang, China, without considering meteorological features. The performance of the prediction of CO_2 was unsatisfactory. The MSE value is 52, but the remaining predictions of the contaminants were good. The MSE value for CO 2.23, HC 0.68, and NO_x 9.4. Carbon monoxide was predicted using a Non-linear Autoregressive Exogenous (NARX) based neural network in the absence of traffic data at Shiraz, Iran [44]. The proposed model performed well compared to the previous models with RMSE values of 0.43, R^2 0.31, and MAPE 51.

There are some studies used three datasets, namely, air quality, meteorological, and traffic datasets. ANN was applied to predict the level of NO , NO_2 , O_3 , NO_x , CO_2 , PN_{10} , NH_3 , PM_{10} , $PM_{2.5}$, and PM_1 pollutants from on-road vehicles at the street canyon in Germany. The model have the lowest RMSE for some pollutants, while others have the highest RMSE, which shows that the model has to be improved for predicting these pollutants. The RMSE for NO , NO_2 , O_3 , NO_x , CO_2 , PN_{10} , NH_3 , PM_{10} , $PM_{2.5}$, and PM_1 were 16.017, 5.092, 5.774, 32.820, 0.790, 12,872.74, 13.474, 0.050, 0.013, and 0.010 [45]. [46] proposed an ANN model to predict CO concentration at Subang Jaya Toll plaza, Selangor, Malaysia. Traffic and meteorological variables were used as an input to the model. The model shows good accuracy with

MAE 0.8925, RMSE 1.2736, RAE 21.99, and RRSE 19.40. A comparison with ANN and developed ResNetELF was conducted [47] to predict the CO , CO_2 , NO_x , and HC levels. The ANN performed well with RMSE 0.0930 for CO , 0.080 for CO_2 , 0.0856 for NO_x , and 0.0798 for HC . Furthermore, [48] predict NO , CO , and HC concentrations using the ANN model. The model's performance was found good with RMSE 1.89, 0.97, 1.09 for NO , CO , and CO_2 . Table I presents the variables and size of the dataset used by previous studies. Table II summarizes the performance of the models used in the previous researches.

TABLE I. SUMMARY OF THE VARIABLES AND DATASET SIZE USED BY PREVIOUS STUDIES

Author	Variables	Dataset
(Wang et al., 2020)	CO , traffic volume, population density.	One week
(Khan et al., 2019)	CO , car, truck, population, year, and GDP transportation.	N/A
(Zahoor et al., 2019)	HC , CO , CO_2 , NO_x , LNG, buses, speed, acceleration, passenger load, and road grade.	N/A
(Mohebbi et al., 2019)	CO , rainfall, temperature, wind direction, moisture, and wind speed.	Four years (2005-2008)
(Goulier et al., 2020)	NO , NO_2 , O_3 , NO_x , CO_2 , PN_{10} , NH_3 , PM_{10} , $PM_{2.5}$, PM_1 , sound, traffic, time, temperature, wind direction and speed, and relative humidity.	Three Month (2018)
(Azeez et al., 2019)	CO , heavy truck, buses, medium truck, and special duty-truck) cars (taxi and private cars) and motorbikes, wind direction, temperature, and wind speed.	One Month (April 2017)
(Xu et al., 2019)	NO_x , CO , CO_2 , HC , temperature, humidity, weather, pressure, wind speed, and road network.	Three Month (2017)
(Xu et al., 2017)	CO , HC , NO , traffic volume, vehicle speed and length, vehicle registration, wind speed, opacity, temperature, wind direction, pressure, and humidity.	N/A

III. METHODOLOGY

A. Data and Location

In this paper, air quality, meteorological, and traffic datasets were used. The traffic data was obtained from the Ministry of Works, Malaysia, while the air pollution and meteorological datasets are collected from the Department of Environment (DOE), Malaysia. These datasets are set of observations recorded at a specific time for sixteen hours daily for three years (2014-2016). The CO , NO , NO_2 , and NO_x features from air quality and meteorological features such as relative humidity, wind speed, and temperature were used [36]. The traffic dataset consists of traffic volume, the volume of the type of vehicle (taxi and car, bus, van, heavy and light lorries, and motorcycle), time spent on the road, and speed of the vehicle were used [49]. The traffic dataset was collected from the Ministry of Works Malaysia at Jalan Kepong traffic census station located in Kuala Lumpur, Malaysia.

TABLE II. SUMMARY OF THE MODEL'S PERFORMANCE OF THE PREVIOUS STUDIES

Author	Pollutant	MAE	MSE	RMSE
(Wang et al., 2020)	CO	0.056	—	0.088
(Khan et al., 2019)	CO	0.347	—	0.442
(Zahoor et al., 2019)	CO , CO_2 , NO_x , HC	—	2.23 52 9.4 0.68	—
(Azeez et al., 2019)	CO	—	—	0.43
(Goulier et al., 2020)	NO , NO_2 , O_3 , NO_x , CO_2 , PN_{10} , NH_3 , PM_{10} , $PM_{2.5}$, PM_1	—	—	16.017 5.092 5.774 32.820 0.790 12.872.27 13.474 0.050 0.013 0.010
(Mohebbi et al., 2019)	CO ,	0.8925	—	1.2736
(Xu et al., 2019)	CO , CO_2 , NO_x , HC	—	—	0.0930 0.080 0.0856 0.0798
(Xu et al., 2017)	NO , CO , HC	—	—	1.89 0.97 1.09

B. Design and Development of OANN Model

An optimized artificial neural network (OANN) was designed and developed to enhance the existing artificial neural network (ANN) model by updating the initial weights in the network using a Genetic Algorithm (GA). An artificial Neural Network (ANN) is an information processing system develop to imitate the human brain's learning and decision-making from experience and examples [50, 51]. The structure of the ANN model consists of input layer, hidden layers, and output layer(s). These layers consist of neurons. These neuron's connection is associated with weight and bias. The neurons have an activation function that determines the neuron's output [52]. The structure of the ANN equation is given below [51]:

$$y = f \left(\sum_{i=1}^n w_i v_i - b \right) \quad (1)$$

Where y is the output of the network, n is the number of neurons in hidden layers, w_i is the weights of the respective neurons, v_i is the input values of the neurons, f is the activation function, and b is the bias. The learning process of ANN improves the model's performance during training by updating the weights in the network. The weights of neurons in ANN define how much influence the input has on the output. The initial weights are randomly chosen [53]. The optimization method has been used to improve and update the weights to efficiently improve the model's accuracy.

Optimization is a process of making something better or finding the best solution, or making a good decision. In this study, a Genetic Algorithm (GA) was designed and developed to optimize the initial weights of ANN to improve its performance in this study. GA works as a process of making changes or finding optimal solutions for the problems. GA works on a

population with a chromosome, and each chromosome has a number of values known as genes [54, 55, 56].

Population in GA represents the entire network's weights (that is, the weights of the input layer to the hidden layer and the weights of the hidden layers to the output layer). Chromosomes represent the weights in one layer (for example, the weights of the input layer to the hidden layer). Finally, Gene represents each neuron's weight (for example, if we have five neurons in the input layer). Fig. 1 illustrates the representations of the GA in the ANN model.

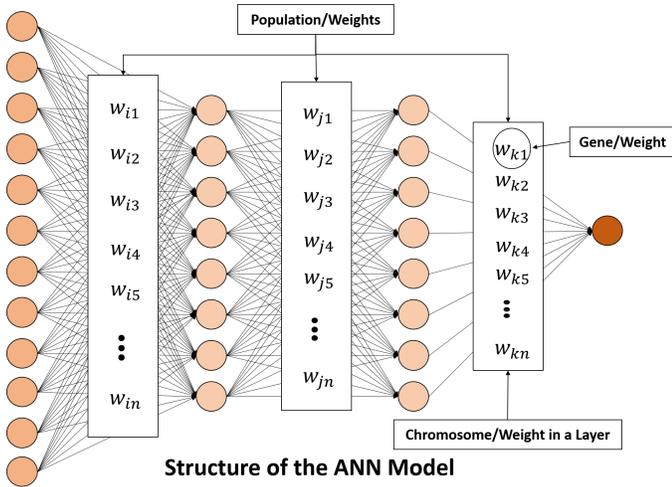


Fig. 1. Representation of GA in an ANN Model.

We will create three vectors. The first vector holds the number of chromosomes per population, the second vector holds the size of the population, and the last vector holds the initial population. There are few processes to generate a new and better population. The processes are given below:

- 1) Select the best parent based on their fitness function.
- 2) The selected parents will be used to produce offspring.
- 3) The production of offspring was performed using crossover and mutation.
- 4) Generated offspring (we create a vector to hold the generated offspring (new population)).
- 5) The algorithm generates new populations (based on how many generations needed).
- 6) One of the best population will be selected.
- 7) The algorithm stops when the optimal solution is found.

Ring pattern was used for selecting parent; for example, if we have ten chromosomes and each chromosome has six genes, the newly generated population will have the same number of chromosomes and genes. If seven chromosomes were found to have the highest fitness value, they would be selected. The remaining three chromosomes will be produced using the ring style. The algorithm will combine the seventh and the first chromosome to produce the eighth one. The first and second chromosomes will be combined to produce the ninth one; we apply a similar way until we get similar

chromosomes in the new generation. An example of initial weight (population) is given in Fig. 2.



Fig. 2. Population in Genetic Algorithm (GA)/Weights in ANN.

The selection of parents was done by calculating the fitness value or function. This was performed using the equation below [57, 58]:

$$f(c) = 2c + 2 \quad (2)$$

Where f is the fitness and c is the chromosome values (genes/weights). An example of a population given in Fig. 2 was calculated in Table III. Based on the table, the first, third, and last chromosomes have the highest fitness values. These parents will be selected to produce new offspring.

TABLE III. FITNESS VALUES CALCULATION

Chromosome	Fitness Function	Parent Selection
0.42 + 0.95 + 0.01 + 0.11 + 0.22 + 0.10 = 1.61	$f(1.72) = 2(1.72)+2 = 5.44$	5.44
0.12+ -0.51 + 0.05 + 0.19 + 0.25 + 0.26 = 0.36	$f(0.36) = 2(0.36)+2 = 2.72$	2.72
0.02 + 0.06 + 0.09 + 0.50 + 0.32 + 0.69 = 1.68	$f(1.68) = 2(1.68)+2 = 5.36$	5.36
0.31 + 0.60 + -0.02 + 0.52 + 0.16 + 0.01 = 1.58	$f(1.58) = 2(1.58)+2 = 5.16$	5.16

As we have discussed earlier, producing offspring was done through crossover and mutation. The crossover is the process of combining two selected parents to produce two offspring [56, 59]. The first half of the first parent and last half of the second parent are selected as the first half of the offspring, while the last half of the first parent and the first half of the second parent were chosen as the second half of the offspring. Fig. 3 shows the producing offspring using the crossover.

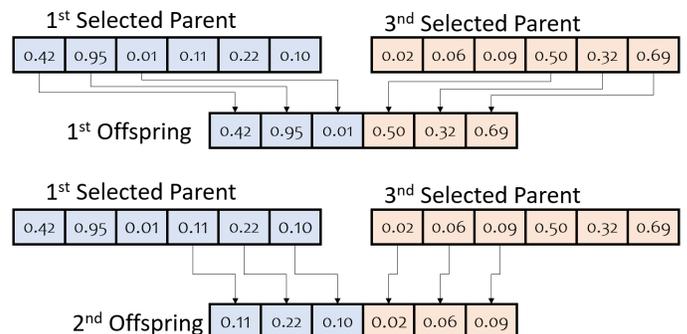


Fig. 3. Producing Offspring Using Crossover.

The mutation was achieved by changing the gene's (weight) value from the new offspring [57, 60]; the altered gene was called a mutant. The gene alteration was done randomly by changing the gene with a lower or higher number than the value of the previous gene range between -1 to 1. An example of mutation is presented in Fig. 4.

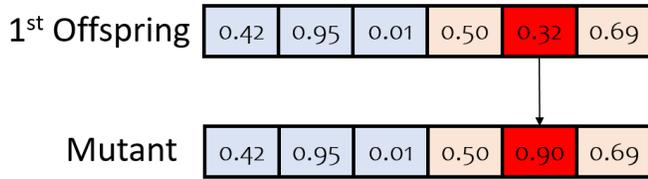


Fig. 4. Producing Offspring Using Mutation.

After the new population was generated, the algorithm produces new generations (based on how many generations are needed). One of the best generations (optimized weights) will be selected and used as the weights of the ANN model. The flow chart of the Genetic Algorithm (GA) for selecting or producing a new solution has been presented in Fig. 5.

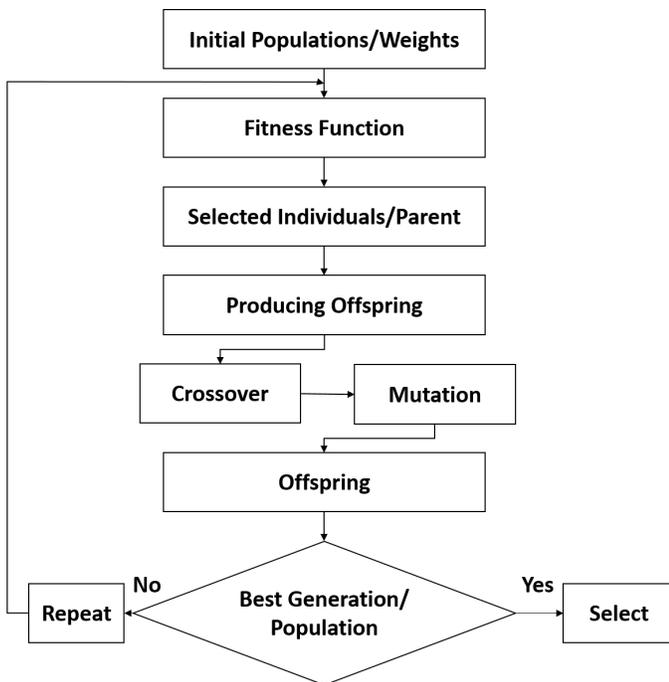


Fig. 5. The Flow Chart of GA for Finding Optimal Solution.

C. Implementation of the OANN Model

The developed OANN predictive model was implemented to predict the concentrations of pollutants produced by motor vehicles. The estimated hourly traffic volume, generated vehicle speed and time spent on the road, standardized pollutants values, meteorological and air quality variables were used to train and to test the developed predictive model for traffic emissions concentrations.

The developed OANN model consists of four layers: one input layer, two hidden layers, and one output layer. The input

layer consists of 12 inputs: traffic volume, taxi and car volume, bus volume, van volume, heavy lorries volume, light lorries volume, motorcycle volume, time spent on the road, vehicle speed, and relative humidity, wind speed, and temperature. The first hidden layer has ten neurons. The second hidden layer has 100 neurons. The output layer has only one output. The output is the predicted values of the pollutants. For OANN prediction model was created for all the four pollutants namely, CO, NO, NO₂, and NO_x. The linear activation function was used from the output layer. The ReLU activation functions were used for the hidden layers.

The weights of the model were optimized using a Genetic Algorithm. Fifty populations were generated, and one of the best populations was selected. The structure of the OANN model has been presented in Fig. 6 and applying the GA in the model was presented in the Fig. 7.

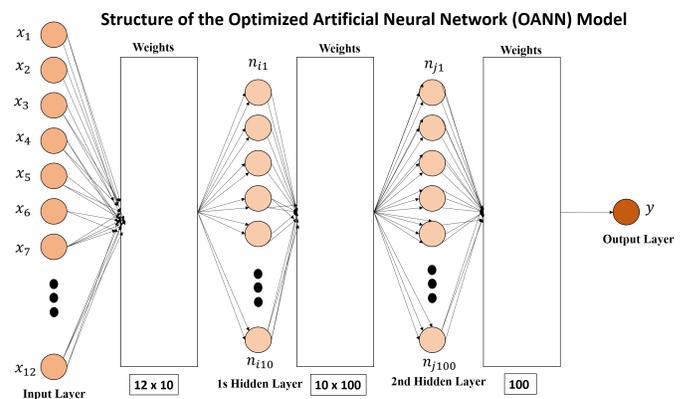


Fig. 6. Structure of the OANN Model.

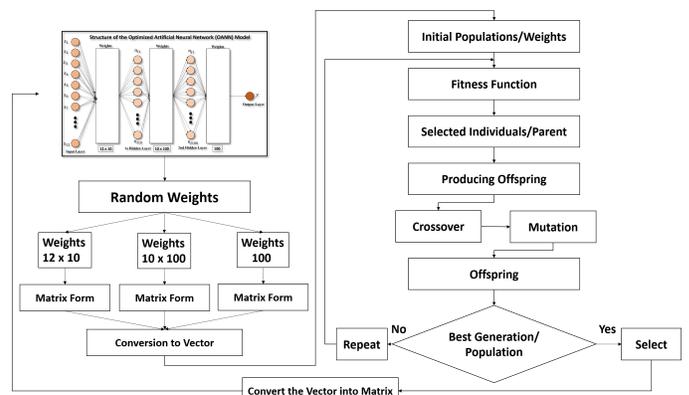


Fig. 7. GA in an OANN Model.

IV. RESULT AND DISCUSSION

The developed Optimized Artificial Neural Network (OANN) model was evaluated using two main regression metrics, namely, Mean Absolute Error (MAE), Mean Squared Error (MSE), and comparison with Artificial Neural Network (ANN), Decision Tree (DT), and Random Forest (RF) models. MSE is the difference between the actual and predicted values. The MEA calculates and finds the differences between the

measure or predicted value and the observed value. Equation 3 used in this study is given below [61]. MSE is the difference between the actual and predicted values. MSE was calculated using equation four [62].

The n is the number of prediction values, \sum represent the summation (adding all the values), \hat{x}_i is the predicted values, x_i represents the actual values, and $|\hat{x}_i - x_i|$ is the absolute errors.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (3)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x - \hat{x})^2 \quad (4)$$

An Optimized Artificial Neural Network (OANN) was developed to predict the concentration of traffic emission in Jalan Kepong area, Kuala Lumpur. The OANN was enhanced by updating the initial weights of the ANN model using GA to improve the accuracy of the existing ANN model. There is a total of 1120 initial weights generated by the model. These weights were from the input layer neurons, first hidden layer neurons, and second hidden layer neurons. There are 12 neurons in the input layer, and they are connected to the first hidden layer. There are ten neurons in the first hidden layer, and they are connected to the second hidden layer. Finally, there are 100 neurons in the second hidden layer, and they are attached to the neuron in the output layer. The generated weights were optimized to improve the ANN model. The GA generated 50 populations/optimized weights, and one of the best populations was selected. Fig. 8 presents the best solution or best population/optimized weights for predicting the concentrations of CO , NO , NO_2 , and NO_x emissions.

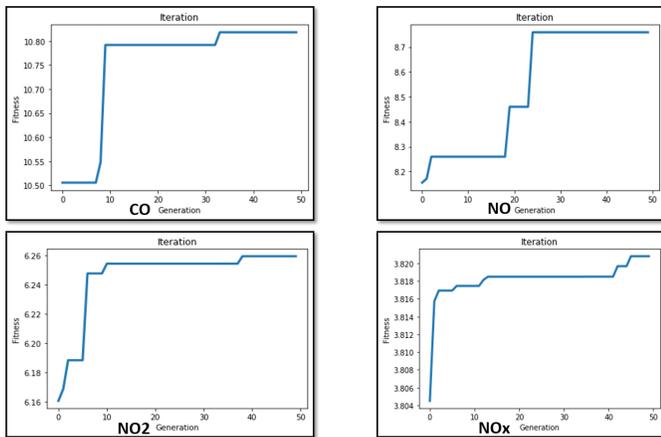


Fig. 8. Best Solution/Population/Weights for Predicting the CO , NO , NO_2 , and NO_x Pollutants.

The best set of weights will be selected from the generated 50 population. In Fig. 8 the set of weights was selected after 33 generation for predicting the CO pollutant.

The best set of weights for predicting NO pollutant was selected after 24 generations. Best optimized weights were

selected after 38 generations for NO_2 prediction. The best set of weights was selected after 45 generations for predicting the level of NO_x emissions. Sample of the optimized populations/weights was illustrated in Fig. 9.

best_solution_weights - List (6 elements)

Ind	Type	Size	Value
0	Array of float64	(12, 10)	[[0.93529863 0.47495294 -0.36123049 ... -1.04555445 0 ...
1	Array of float64	(10,)	[0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
2	Array of float64	(10, 100)	[[-1.06787533e-02 4.32580709e-03 4.50262090e-01 ... -5.35256118e-02 ...
3	Array of float64	(100,)	[0. 0. 0. ... 0. 0. 0.]
4	Array of float64	(100, 1)	[[0.13767141] [-0.17720821]
5	Array of float64	(1,)	[0.]

Fig. 9. The Optimized Weights in The Layers.

The highlighted in yellow colors in Fig. 9 were the optimized weights from the input layer to the first hidden layer, which consist of 12 neurons connected to the first hidden layer with ten neurons. The first hidden layer consists of 10 neurons connected to the second hidden layer with 100 neurons. The second or last hidden layer has 100 neurons connected to the output layer with one neuron. The population with the optimized weights were chosen for prediction of the CO , NO , NO_2 , and NO_x . The set of the optimized weights for the prediction of the CO pollutant was presented in Fig. 10.



Fig. 10. The Optimized Weights in the Network.

The highlighted green color in Fig. 10 were the number of neurons in the layer. While the highlighted in yellow colors were the number of neurons connected to the next layer.

A sample of the total fitness of each generation by GA for predicting CO pollutant has been given. The fitness values of the first set of weights were 10.505416290436793. There are no changes in the fitness values until the seventh generation. In the eighth generation, there is a change from 10.505416290436793 to 10.548116345442276. In the ninth generation, there is a change from 10.548116345442276 to 10.791954330972974. From the tenth generation until the thirty-second generation, there were no changes for the fitness functions. In the thirty-three generation, there is a change from 10.791954330972974 to 10.818411975096184. From thirty-fourth generation until fifty generations there is no change in the fitness value. The highest fitness value of the weights was found in the thirty-three generation.

The developed Optimized Artificial Neural Network model was used to predict the concentrations of CO , NO , NO_2 , and NO_x pollutants emitted by motor vehicles in Jalan Kepong, Kuala Lumpur. Mean absolute error (MAE) and mean squared error (MSE) regression metrics were used to evaluate the performance of the model. Fig. 11 shows the predicted and original values of the pollutants. The prediction of the CO pollutant shows that the model was able to follow the trend of the original values. Similarly, the prediction of NO also follows the trend of the original values with a slightly difference. The OANN model also followed the trends of the original values for predicting the NO_2 and NO_x pollutants.

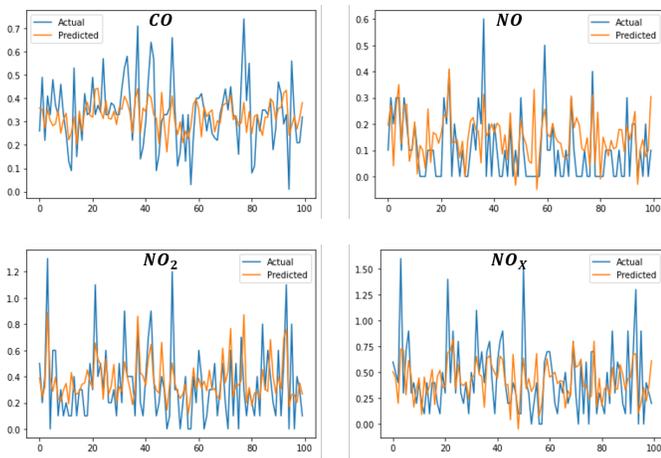


Fig. 11. Predicted Values and the Original Values of the Pollutants Using OANN Model.

The developed OANN model's performance was compared with ANN, DT, and RF models for evaluation by predicting the concentrations of CO , NO , NO_2 , and NO_x pollutants. The performance of the models was evaluated using mean squared error and mean absolute error. Fig. 12, 13, 14, and 15 present the original or real values with the predicted values for CO , NO , NO_2 , and NO_x predictions. It can be seen that the predicted values of the developed OANN model were able to follow the trend of the real values more than the DT, RF, and ANN models. The ANN model was not able to follow the trend of the real values except for predicting the NO_x pollutant.

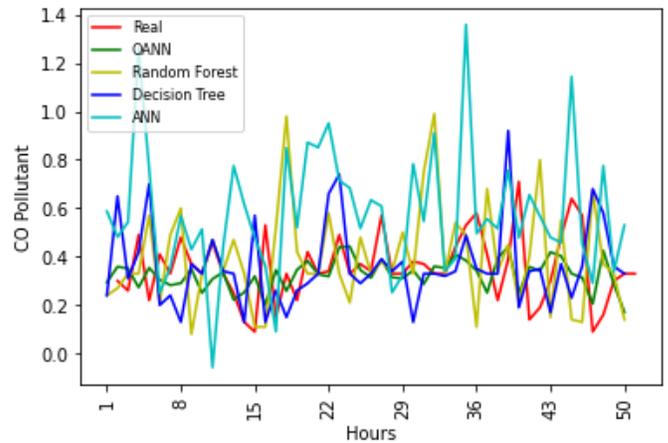


Fig. 12. Comparison of Original values and Predicted Values between OANN, ANN, RF, and DT Models for CO Prediction.

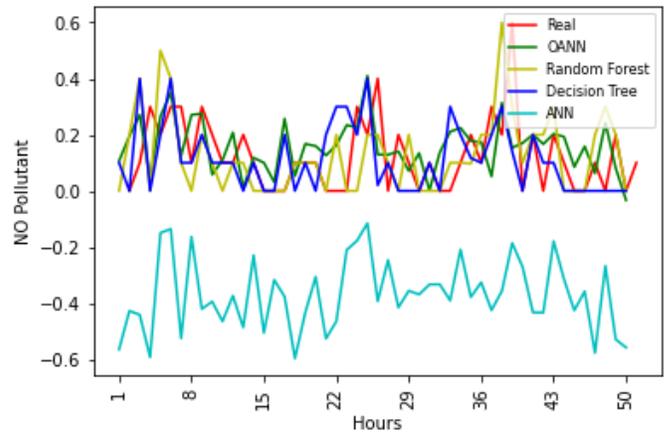


Fig. 13. Comparison of Original values and Predicted Values between OANN, ANN, RF, and DT Models for NO Prediction.

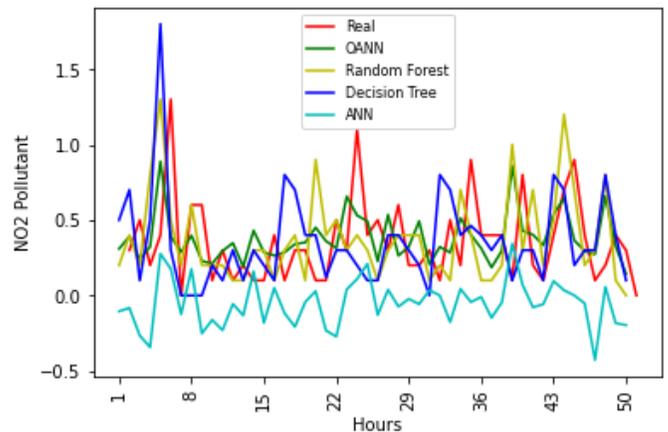


Fig. 14. Comparison of Original values and Predicted Values between OANN, ANN, RF, and DT Models for NO_2 Prediction.

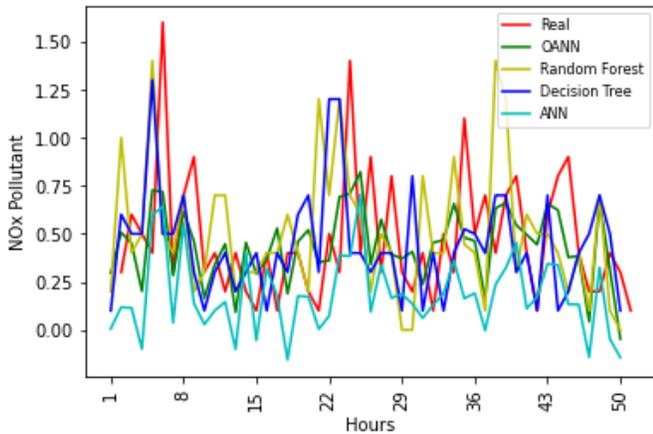


Fig. 15. Comparison of Original values and Predicted Values between OANN, ANN, RF, and DT Models for NO_x Prediction.

Table IV presents the comparison of MAE and MSE values between OANN, ANN, RF, and DT for predicting the concentrations of CO , NO , NO_2 , and NO_x pollutants. The result shows the performance of the OANN model for predicting the concentrations of CO was found better than the ANN, DT, and RF models with lowest MAE and MSE values of 0.1155 and 0.0248. The ANN model performed better than RF and DT with the lower MSE values of 0.0441 for the CO prediction. The DT model performed a little better than the RF model with lower MAE and MSE scores of 0.1522 and 0.0446, whilst the MSE and MAE values of RF were 0.0463 and 0.1551.

The comparison of the OANN, ANN, DT, and RF models shows the significant difference of results in terms of MSE and MAE values for predicting the level of NO pollutant. The MAE and MSE values of the developed model were found lowest compared to DT, RF, and ANN models. The MAE and MSE values were 0.1263 and 0.0365, and this shows that the enhanced OANN model was better than the RF, ANN, and DT models. The DT model performed better than the RF and ANN models with lower MAE and MSE scores of 0.1446 and 0.0642. The MAE and MSE values of the ANN model were higher than the RF model, which shows that the RF model was performed better than the ANN model. The MAE and MSE values of the RF model were 0.1450, 0.0644, and for the ANN model were 0.5042 and 0.3040.

The performance of the OANN model is better than the DT, RF, and ANN models with the lowest MAE and MSE values for predicting the concentrations of NO_2 emissions. The MAE and MSE values were 0.1731 and 0.0542. Moreover, the performance of the ANN model was found unsatisfactory with the highest MAE and MSE values of 0.3928, 0.2288 more than RF and DT models for the NO_2 prediction. The RF model performed better than the DT model for predicting the level of NO_2 emissions with a lower MSE value of 0.0810.

Table IV indicates that the OANN model has the lowest MAE and MSE values than the ANN, DT, and RF models for the prediction of NO_x pollutant., this shows that the OANN model performed better than ANN, DT, and RF models. The MAE and MSE scores of the OANN model were 0.2368 and

0.1128. Unfortunately, the DT model has the highest MAE and MSE scores than the ANN and RF models. The MAE and MSE values of the DT model were 0.3178 and 0.2003. However, the RF model performed better than the ANN model with the MAE and MSE values of 0.2872 and 0.1757 for predicting the concentrations of NO_x pollutant. The MAE and MSE values of the ANN model were 0.2928 and 0.1812.

TABLE IV. COMPARISON BETWEEN OANN, ANN, DT, AND RF FOR TRAFFIC EMISSIONS PREDICTION

AP	OANN		ANN		Random Forest		Decision Tree	
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
CO	0.1155	0.0247	0.1696	0.0441	0.1551	0.0463	0.1522	0.0446
NO	0.1263	0.0365	0.5042	0.3040	0.145	0.0644	0.1446	0.0642
NO_2	0.1731	0.0542	0.3928	0.2288	0.2024	0.0810	0.2007	0.0826
NO_x	0.2368	0.1128	0.2928	0.1812	0.2872	0.1757	0.3178	0.2003

It can be concluded that the developed OANN model achieved the best results in comparison with all models, especially the ANN model for predicting the concentrations of CO , NO , NO_2 , and NO_x pollutants.

V. CONCLUSION

A based Artificial Neural Network (OANN) model was developed to enhance the existing ANN model by updating the initial weights that connect the neurons in the network using a Genetic Algorithm (GA). The OANN model was used to predict the level of pollutants emitted by vehicles in Kuala Lumpur, Malaysia. The OANN model was evaluated using performance metrics and comparison with ANN, DT, and RF models. The result shows the developed OANN model performed better than the existing ANN, DT, and RF models with the lowest regression metrics when compared.

Based on the literature review, the study of [45] predicts the concentration of air pollutants at canyon street. Still, our study predicts the level of air pollutants at the roadside, which is an open space. [46] study was the closest study with our study; the difference is, the study focuses more on the truck vehicles and also used a one-month dataset. Additionally, they predict the concentrations of CO pollutants. Our study added three more pollutants namely, NO , NO_2 , and NO_x . The study of [48] used different variables such as length of the vehicle, vehicle registration, and opacity. The study also has a higher RMSE score than the RMSE values of our study.

The developed OANN model can help environmental agencies monitor traffic-related air pollution levels efficiently and take necessary actions to ensure the effectiveness of traffic management policy. Moreover, the model can help decision-makers mitigate traffic emissions to protect the health of the citizens who are inhabiting very close to highways.

Admittedly, there are some limitations to this study. Firstly, Some studies suggested that different types of fuel have different types of emissions, but the fuel type was not considered in this study. The Jalan Kepong traffic census station was selected, but the other stations were not considered. There are many pollutants emitted by motor vehicles. However, the developed OANN model was limited to predict the hourly CO , NO , NO_2 , and NO_x concentrations. It can be applied to the other traffic pollutants as well. Furthermore, daily, weekly,

monthly, or yearly predictions were not the focus of this research. The prediction using only one type of vehicle was not considered as well.

ACKNOWLEDGMENT

This research is funded by Taylor's University under the research grant application ID (TUFR/2017/004/04) entitled as "Modeling and Visualization of Air-Pollution and its Impacts on Health".

REFERENCES

- [1] M. Bilal, R. S. A. Usmani, M. Tayyab, A. A. Mahmoud, R. M. Abdalla, M. Marjani, T. R. Pillai, and I. A. Targio Hashem, "Smart cities data: Framework, applications, and challenges," *Handbook of Smart Cities*, pp. 1–29, 2020.
- [2] R. S. A. Usmani, I. A. T. Hashem, T. R. Pillai, A. Saeed, and A. M. Abdullahi, "Geographic information system and big spatial data: A review and challenges," *International Journal of Enterprise Information Systems (IJEIS)*, vol. 16, no. 4, pp. 101–145, 2020.
- [3] Z. Sun, C. Wang, Z. Ye, and H. Bi, "Long short-term memory network-based emission models for conventional and new energy buses," *International Journal of Sustainable Transportation*, vol. 15, no. 3, pp. 229–238, 2021.
- [4] C.-H. Lai, P.-K. Hsiao, Y.-T. Yang, S.-M. Lin, and S.-C. C. Lung, "Effects of the manual and electronic toll collection systems on the particulate pollutant levels on highways in taiwan," *Atmospheric Pollution Research*, vol. 12, no. 3, pp. 25–32, 2021.
- [5] L. Bertrand, L. Dawkins, R. Jayaratne, and L. Morawska, "How to choose healthier urban biking routes: Co as a proxy of traffic pollution," *Heliyon*, vol. 6, no. 6, p. e04195, 2020.
- [6] R. S. A. Usmani, T. R. Pillai, I. A. T. Hashem, N. Z. Jhanjhi, A. Saeed, and A. M. Abdullahi, "A spatial feature engineering algorithm for creating air pollution health datasets," *International Journal of Cognitive Computing in Engineering*, vol. 1, pp. 98–107, 2020.
- [7] J. Huang, G. Song, J. Zhang, C. Li, Q. Liu, and L. Yu, "The impact of violations of bicycles and pedestrians on vehicle emissions at signalized intersections," *Journal of advanced transportation*, vol. 2020, 2020.
- [8] J. Su, "Portable and sensitive air pollution monitoring," *Light, science & applications*, vol. 7, 2018.
- [9] R. S. A. Usmani, T. R. Pillai, I. A. T. Hashem, M. Marjani, R. Shaharudin, and M. T. Latif, "Air pollution and cardiorespiratory hospitalization, predictive modeling, and analysis using artificial intelligence techniques," *Environmental Science and Pollution Research*, pp. 1–13, 2021.
- [10] S. Munir, M. Mayfield, and D. Coca, "Understanding spatial variability of no₂ in urban areas using spatial modelling and data fusion approaches," *Atmosphere*, vol. 12, no. 2, p. 179, 2021.
- [11] R. S. A. Usmani, A. Saeed, A. M. Abdullahi, T. R. Pillai, N. Z. Jhanjhi, and I. A. T. Hashem, "Air pollution and its health impacts in malaysia: a review," *Air Quality, Atmosphere & Health*, vol. 13, no. 9, pp. 1093–1118, 2020.
- [12] S. Kim, C. Xiao, I. Platt, Z. Zafari, M. Bellanger, and P. Muennig, "Health and economic consequences of applying the united states' pm_{2.5} automobile emission standards to other nations: a case study of france and italy," *Public health*, vol. 183, pp. 81–87, 2020.
- [13] Y. Romero, C. Diaz, I. Meldrum, R. A. Velasquez, and J. Noel, "Temporal and spatial analysis of traffic-related pollutant under the influence of the seasonality and meteorological variables over an urban city in peru," *Heliyon*, vol. 6, no. 6, p. e04029, 2020.
- [14] J. Zheng, S. Dong, Y. Hu, and Y. Li, "Comparative analysis of the co₂ emissions of expressway and arterial road traffic: A case in beijing," *PloS one*, vol. 15, no. 4, p. e0231536, 2020.
- [15] X. Bing, Q. Wei, J. Lu, C. Li, and Y. Zhang, "Sustainable highway design: Disentangling the effects of geometric-related and traffic-related factors on urban highway traffic emissions," *Advances in Civil Engineering*, vol. 2018, 2018.
- [16] A. Sfyridis and P. Agnolucci, "Road emissions in london: Insights from geographically detailed classification and regression modelling," *Atmosphere*, vol. 12, no. 2, p. 188, 2021.
- [17] R. Akhtar and C. Palagiano, "Climate change and air pollution," *Switzerland: Springer International Publishing, AG*, 2018.
- [18] X. Shi, D. J. Sun, Y. Zhang, J. Xiong, and Z. Zhao, "Modeling emission flow pattern of a single cruising vehicle on urban streets with cfd simulation and wind tunnel validation," *International Journal of Environmental Research and Public Health*, vol. 17, no. 12, p. 4557, 2020.
- [19] D. Pan, L. Tao, K. Sun, L. M. Golston, D. J. Miller, T. Zhu, Y. Qin, Y. Zhang, D. L. Mauzerall, and M. A. Zondlo, "Methane emissions from natural gas vehicles in china," *Nature communications*, vol. 11, no. 1, pp. 1–10, 2020.
- [20] N. Abdull, M. Yoneda, and Y. Shimada, "Traffic characteristics and pollutant emission from road transport in urban area," *Air Quality, Atmosphere & Health*, vol. 13, no. 6, pp. 731–738, 2020.
- [21] M. Masoudi, N. Rajaipoor, and F. Ordibeheshti, "Status of pm₁₀ as an air pollutant and prediction using meteorological indexes in shiraz, iran," *Advances in environmental research*, vol. 7, no. 2, pp. 109–120, 2018.
- [22] M. Kermani, A. J. Jafari, M. Gholami, F. Fanaei, H. Arfaeinia *et al.*, "Association between meteorological parameter and pm_{2.5} concentration in karaj, iran," *International Journal of Environmental Health Engineering*, vol. 9, no. 1, p. 4, 2020.
- [23] H. Zhou, Y. Yu, X. Gu, Y. Wu, M. Wang, H. Yue, J. Gao, R. Lei, and X. Ge, "Characteristics of air pollution and their relationship with meteorological parameters: Northern versus southern cities of china," *Atmosphere*, vol. 11, no. 3, p. 253, 2020.
- [24] H. Cui, R. Ma, and F. Gao, "Relationship between meteorological factors and diffusion of atmospheric pollutants," *Chemical Engineering Transactions*, vol. 71, pp. 1417–1422, 2018.
- [25] M. S. Jamil, A. Z. Ul-Saufie, A. Abu Bakar, K. A. M. Ali, and H. Ahmat, "Identification of source contributions to air pollution in penang using factor analysis," *Int J Integr Eng*, vol. 11, no. 8, pp. 221–228, 2019.
- [26] A. Nazif, N. I. Mohammed, A. Malakahmad, and M. S. Abualqumboz, "Evaluating the influence of meteorological parameters on ozone concentration levels," *International Journal of Engineering & Technology*, vol. 7, no. 3.7, pp. 65–67, 2018.
- [27] A. A. A. Mohtar, M. T. Latif, N. H. Baharudin, F. Ahamad, J. X. Chung, M. Othman, and L. Juneng, "Variation of major air pollutants in different seasonal conditions in an urban environment in malaysia," *Geoscience Letters*, vol. 5, no. 1, p. 21, 2018.
- [28] R. Rossi, R. Ceccato, and M. Gastaldi, "Effect of road traffic on air pollution. experimental evidence from covid-19 lockdown," *Sustainability*, vol. 12, no. 21, p. 8984, 2020.
- [29] P. Krecil, Y. A. Cipoli, A. C. Targino, L. B. Castro, L. Gidhagen, F. Malucelli, and A. Wolf, "Cyclists' exposure to air pollution under different traffic management strategies," *Science of the Total Environment*, vol. 723, p. 138043, 2020.
- [30] S. S. Anjum, R. M. Noor, N. Aghamohammadi, I. Ahmedy, L. M. Kiah, N. Hussin, M. H. Anisi, and M. A. Qureshi, "Modeling traffic congestion based on air quality for greener environment: an empirical study," *IEEE Access*, vol. 7, pp. 57 100–57 119, 2019.
- [31] B. Choubin, M. Abdolshahnejad, E. Moradi, X. Querol, A. Mosavi, S. Shamshirband, and P. Ghamisi, "Spatial hazard assessment of the pm₁₀ using machine learning models in barcelona, spain," *Science of The Total Environment*, vol. 701, p. 134474, 2020.
- [32] A. Suleiman, M. Tight, and A. Quinn, "Applying machine learning methods in managing urban concentrations of traffic-related particulate matter (pm₁₀ and pm_{2.5})," *Atmospheric Pollution Research*, vol. 10, no. 1, pp. 134–144, 2019.
- [33] I. Šimić, M. Lovrić, R. Godec, M. Kröll, and I. Bešlić, "Applying machine learning methods to better understand, model and estimate mass concentrations of traffic-related pollutants at a typical street canyon," *Environmental Pollution*, vol. 263, p. 114587, 2020.
- [34] V. Gnanamoorthi, P. Purushothaman, A. Gurusamy, G. Devaradjane *et al.*, "Prediction efficiency of artificial neural network for crdi engine output parameters," *Transportation Engineering*, vol. 3, p. 100041, 2021.
- [35] M. Nandal, N. Mor, and H. Sood, "An overview of use of artificial neural network in sustainable transport system," *Computational Methods and Data Engineering*, pp. 83–91, 2021.
- [36] I. A. T. Hashem and T. R. Pillai, "A novel feature engineering algorithm for air quality datasets," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 3, pp. 1444–1451, 2020.
- [37] E. López-Pérez, T. Hermosilla, J.-M. Carot-Sierra, and G. Palau-Salvador, "Spatial determination of traffic co emissions within street canyons using inverse modelling," *Atmospheric Pollution Research*, vol. 10, no. 4, pp. 1140–1147, 2019.
- [38] M. Hilpert, M. Johnson, M.-A. Kioumourtzoglou, A. Domingo-Relloso, A. Peters, B. Adria-Mora, D. Hernández, J. Ross, and S. N. Chillrud,

- "A new approach for inferring traffic-related air pollution: Use of radar-calibrated crowd-sourced traffic data," *Environment international*, vol. 127, pp. 142–159, 2019.
- [39] M. Gustafsson, G. Blomqvist, I. Järskog, J. Lundberg, S. Janhäll, M. Elmgren, C. Johansson, M. Norman, and S. Silvergren, "Road dust load dynamics and influencing factors for six winter seasons in stockholm, sweden," *Atmospheric Environment: X*, vol. 2, p. 100014, 2019.
- [40] C. Xu, J. Zhao, and P. Liu, "A geographically weighted regression approach to investigate the effects of traffic conditions and road characteristics on air pollutant emissions," *Journal of Cleaner Production*, vol. 239, p. 118084, 2019.
- [41] Y. Wang, P. Liu, C. Xu, C. Peng, and J. Wu, "A deep learning approach to real-time co concentration prediction at signalized intersection," *Atmospheric Pollution Research*, vol. 11, no. 8, pp. 1370–1378, 2020.
- [42] M. J. U. R. Khan and A. Awasthi, "Machine learning model development for predicting road transport ghg emissions in canada," *WSB Journal of Business and Finance*, vol. 53, no. 2, pp. 55–72, 2019.
- [43] O. Zahoor, M. Usama, Q. Bao, Z. Abbas, Y. Shen, and S. Chen, "Lng bus emissions prediction using neural network," in *CICTP 2019*. ASCE, 2019, pp. 4156–4168.
- [44] M. R. Mohebbi, A. K. Jashni, M. Dehghani, and K. Hadad, "Short-term prediction of carbon monoxide concentration using artificial neural network (nrx) without traffic data: Case study: Shiraz city," *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, vol. 43, no. 3, pp. 533–540, 2019.
- [45] L. Goulier, B. Paas, L. Ehrnsperger, and O. Klemm, "Modelling of urban air pollutant concentrations with artificial neural networks using novel input variables," *International journal of environmental research and public health*, vol. 17, no. 6, p. 2025, 2020.
- [46] O. S. Azeez, B. Pradhan, H. Z. Shafri, N. Shukla, C.-W. Lee, and H. M. Rizeei, "Modeling of co emissions from traffic vehicles using artificial neural networks," *Applied Sciences*, vol. 9, no. 2, p. 313, 2019.
- [47] Z. Xu, Y. Cao, and Y. Kang, "Deep spatiotemporal residual early-late fusion network for city region vehicle emission pollution prediction," *Neurocomputing*, vol. 355, pp. 183–199, 2019.
- [48] Z. Xu, Y. Kang, and W. Lv, "Analysis and prediction of vehicle exhaust emission using ann," in *2017 36th Chinese Control Conference (CCC)*. IEEE, 2017, pp. 4029–4033.
- [49] A. M. Abdullah, R. S. A. Usmani, T. R. Pillai, I. A. T. Hashem, and M. Marjani, "Feature engineering algorithms for traffic dataset," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 4, 2021. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2021.0120435>
- [50] B. Liu, Q. Zhao, Y. Jin, J. Shen, and C. Li, "Application of combined model of stepwise regression analysis and artificial neural network in data calibration of miniature air quality detector," *Scientific Reports*, vol. 11, no. 1, pp. 1–12, 2021.
- [51] M. Nandal, N. Mor, and H. Sood, "An overview of use of artificial neural network in sustainable transport system," *Computational Methods and Data Engineering*, pp. 83–91, 2020.
- [52] R. Kenanoğlu, M. K. Baltacıoğlu, M. H. Demir, and M. E. Özdemir, "Performance & emission analysis of hho enriched dual-fuelled diesel engine with artificial neural network prediction approaches," *International Journal of Hydrogen Energy*, vol. 45, no. 49, pp. 26 357–26 369, 2020.
- [53] R. E. Neapolitan and R. E. Neapolitan, "Neural networks and deep learning," *Artificial Intelligence*, pp. 389–411, 2018.
- [54] S. Mirjalili, "Genetic algorithm," in *Evolutionary algorithms and neural networks*. Springer, 2019, pp. 43–55.
- [55] N. Razmjoo, M. Ashourian, and Z. Foroozandeh, *Metaheuristics and Optimization in Computer and Electrical Engineering*. Springer, 2019.
- [56] G. Zhang, C. Xiao, and N. Razmjoo, "Optimal parameter extraction of pem fuel cells by meta-heuristics," *International Journal of Ambient Energy*, pp. 1–10, 2020.
- [57] A. F. Gad, A. F. Gad, and S. John, *Practical computer vision applications using deep learning with CNNs*. Springer, 2018.
- [58] N. Razmjoo, V. V. Estrela, H. J. Loschi, and W. Fanfan, "A comprehensive survey of new meta-heuristic algorithms," *Recent Advances in Hybrid Metaheuristics for Data Clustering*, Wiley Publishing, 2019.
- [59] N. Razmjoo, V. V. Estrela, and H. J. Loschi, "Entropy-based breast cancer detection in digital mammograms using world cup optimization algorithm," *International Journal of Swarm Intelligence Research (IJSIR)*, vol. 11, no. 3, pp. 1–18, 2020.
- [60] N. Razmjoo, M. Khalilpour, and M. Ramezani, "A new meta-heuristic optimization algorithm inspired by fifa world cup competitions: theory and its application in pid designing for avr system," *Journal of Control, Automation and Electrical Systems*, vol. 27, no. 4, pp. 419–440, 2016.
- [61] T. Chai and R. R. Draxler, "Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature," *Geoscientific model development*, vol. 7, no. 3, pp. 1247–1250, 2014.
- [62] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009.

Discrete Time-Space Stochastic Mathematical Modelling for Quantitative Description of Software Imperfect Fault-Debugging with Change-Point

‘Mohd Taib’ Shatnawi¹
Al-Huson University College
Al-Balqa’ Applied University
Irbid 21510, Jordan

Omar Shatnawi²
Computer Science Department
Al al-Bayt University
Mafraq 25113, Jordan

Abstract—Statistics and stochastic-process theories, along with the mathematical modelling and the respective empirical evidence support, describe the software fault-debugging phenomenon. In software-reliability engineering literature, stochastic mathematical models based on the non-homogeneous Poisson process (NHPP) are employed to measure and boost reliability too. Since reliability evolves on account of the running of computer test-run, NHPP type of discrete time-space models, or difference-equation, is superior to their continuous time-space counterparts. The majority of these models assume either a constant, monotonically increasing, or decreasing fault-debugging rate under an imperfect fault-debugging environment. However, in the most debugging scenario, a sudden change may occur to the fault-debugging rate due to an addition to, deletion from, or modification of the source code. Thus, the fault-debugging rate may not always be smooth and is subject to change at some point in time called change-point. Significantly few studies have addressed the problem of change-point in discrete-time modelling approach. The paper examines the combined effects of change-point and imperfect fault-debugging with the learning process on software-reliability growth phenomena based on the NHPP type of discrete time-space modelling approach. The performance of the proposed modelling approach is compared with other existing approaches on an actual software-reliability dataset cited in literature. The findings reveal that incorporating the effect of change-point in software-reliability growth modelling enhances the accuracy of software-reliability assessment because the stochastic characteristics of the software fault-debugging phenomenon alter at the change-point.

Keywords—Stochastic mathematical modelling; discrete time-space; non-homogenous poisson process; change-point; imperfect fault-debugging; software-reliability

I. INTRODUCTION

Today, computer-based software systems are indispensable, and their successful operation depends mainly on software. A fault can be introduced at any point during the software development life cycle (abbreviated as SDLC) due to the deficiency of human-being. A failure is a consequence of a fault. The software development process’s testing phase aims at debugging faults. Software reliability is typically defined as a statistical measure of a software system’s ability to operate failure-free. To quantify software reliability, a number of analytical approaches to mathematical modelling the stochastic-behavior of software debugging phenomenon have been proposed. Non-homogeneous Poisson process (abbreviated as NHPP) models

are widely used in software-reliability engineering to quantitatively express reliability. They are broadly divided into discrete time-space and continuous time-space groups. However, discrete time-space models which adopt the number of executed computer test-run as a unit of fault-debugging period are more suited than their continuous time-space counterparts. Despite the difficulties in formulating them, many studies have highlighted their usefulness [1]–[4].

Several existing NHPP stochastic software-reliability models are based on constant or monotonically increasing fault-debugging rates under perfect and imperfect fault-debugging environments [1], [3], [5]–[16]. In practice, as the debugging grows, a sudden change may take place to the fault-debugging rate as a result of an addition to, deletion from, or modification of the source code. Thus, the fault-debugging rate may not always be smooth and is subject to change at some point in time called change-point [1], [17]–[23]. Many studies argue that considering the change-point concept in imperfect fault-debugging is expected to enhance the software-reliability assessment accuracy due to the stochastic characteristics of software fault-debugging phenomenon changed at the change-point. Thus, the incorporation of change-point provides a notable enhancement in the reliability assessment. However, most of the endeavors are in continuous time-space. Lately, this area has received little attention and few studies have incorporated the change-point concept in developing discrete time-space software-reliability modelling [1], [24], [25]. However, the study of the change-point problem is still very limited in the discrete-time modelling approach. As a result, the paper studies the effect of incorporating a change-point concept in modelling the imperfect fault-debugging phenomenon.

The rest of the paper is structured as follows: Section II reviews related work and models the change-point problem into an imperfect fault-debugging environment through a discrete time-space NHPP based approach. Data analyses and parameter estimation techniques, model validation, comparison criteria, and software-reliability measures are discussed in Sections III and IV, respectively. The descriptive-performance and predictive-capability, and software-reliability evaluation measures based on the proposed modelling approach are shown in Section V. Finally, Section VI concludes the paper and presents future works.

II. SOFTWARE-RELIABILITY MODELLING: A DISCRETE TIME-SPACE APPROACH

In this approach, we model the debugged faults by n computer test-run as a pure-birth counting process $(N_n; n \geq 0)$, [1]–[3] subject to

- No fault debugged at $n = 0$, that is, $N(0) = 0$.
- For any number of computer test-run n_1, n_2, \dots, n_k where $(0 < n_1 < n_2 < \dots < n_k)$. The k random variables $N(n_1), N(n_2) - N(n_1), \dots, N(n_k) - N(n_{k-1})$ are statistically independent.
- For any number of computer test-run n_i and n_j , where $(0 \leq n_i \leq n_j)$, we have

$$Pr[N(n_j) - N(n_i) = k] = \frac{(\lambda_n)^k}{k!} e^{-\lambda_n} \quad (1)$$

here λ be the fault-debugging intensity function.

Accordingly, the NHPP software-reliability model can be formulated

$$Pr[N_n = k] = \frac{(m_n)^k}{k!} e^{-m_n} \quad (2)$$

here N_n be the expected value number of faults whose mean-value function (abbreviated as MVF) is known as m_n .

A. Model Development

In this section, we review some of the well-documented NHPP type of discrete time-space models

1) *Exponential Model* [3], [26]: This model considers that the average number of faults debugged between n^{th} and the $(n + 1)^{\text{th}}$ computer test-run, is corresponding to the current number of undebugged faults after the execution of the n^{th} computer test-run, meets the resulting difference equation:

$$\lambda_n = \frac{m_{n+1} - m_n}{\delta} = b(a - m_n) \quad (3)$$

here a be a fault-content, and the constant b represents a fault-debugging per undebugged faults per test case.

To solve the difference equation (3), we employ the probability generating function (abbreviated as PGF) computational technique, multiply both sides by w^n , and sum over n from 0 to ∞ , we get:

$$\begin{aligned} \sum_{n=0}^{\infty} w^n m_{n+1} - \sum_{n=0}^{\infty} w^n m_n &= \delta b \left(a \sum_{n=0}^{\infty} w^n - \sum_{n=0}^{\infty} w^n m_n \right) \\ \frac{1}{w} \sum_{n=0}^{\infty} w^{n+1} m_{n+1} - \sum_{n=0}^{\infty} w^n m_n &= \delta ab \sum_{n=0}^{\infty} w^n - \delta b \sum_{i=0}^{\infty} w^n m_n \\ \sum_{n=0}^{\infty} w^{n+1} m_{n+1} - \sum_{n=0}^{\infty} w^{n+1} m_n &= \delta ab \sum_{n=0}^{\infty} w^{n+1} - \delta b \sum_{n=0}^{\infty} w^{n+1} m_n \\ \sum_{n=0}^{\infty} w^{n+1} m_{n+1} &= \delta ab \sum_{n=0}^{\infty} w^{n+1} - (1 - \delta b) \sum_{n=0}^{\infty} w^{n+1} m_n \\ (w^1 m_1 + w^2 m_2 + w^3 m_3 + \dots) &= \delta ab (w^1 + w^2 + w^3 + \dots) + \\ (1 - \delta b) (w^1 m_0 + w^2 m_1 + w^3 m_2 + \dots) & \quad (4) \end{aligned}$$

Comparing the coefficients of like powers of w on both sides in (4) and under the initial-condition $m_{n=0} = 0$, yields

$$\begin{aligned} m_1 &= \delta ab + (1 - \delta b) m_0 = a(1 - (1 - \delta b)^1) \\ m_2 &= \delta ab + (1 - \delta b) m_1 = a(1 - (1 - \delta b)^2) \\ m_3 &= \delta ab + (1 - \delta b) m_2 = a(1 - (1 - \delta b)^3) \end{aligned} \quad (5)$$

The closed-form solution, by mathematical-induction, obtained as

$$m_n = a(1 - (1 - \delta b)^n) \quad (6)$$

Accordingly, the fault-debugging intensity function is

$$\lambda_n = \frac{m_{n+1} - m_n}{\delta} = ab(1 - \delta b)^n \quad (7)$$

The equivalent continuous time-space model [27] corresponding to (6) is

$$m_t = a(1 - e^{-bt}) \quad (8)$$

which derived as a limiting case of discrete time-space model substituting $t = n\delta$, $\lim_{x \rightarrow 0} (1 + x)^{1/x} = e$, and taking limit $\delta \rightarrow 0$.

2) *Delayed S-shaped Model* [1], [16]: This model considers that the fault-debugging phenomenon as a two-stage process namely, detection and correction. Accordingly, we have

$$\begin{aligned} \frac{m_{d(n+1)} - m_{d(n)}}{\delta} &= b(a - m_{d(n)}) \\ \frac{m_{c(n+1)} - m_{c(n)}}{\delta} &= b(m_{d(n)} - m_{c(n)}) \end{aligned} \quad (9)$$

Solving the system of difference equation (9), using the PGF computational technique in terms of the initial condition that at $n = 0$, $m_{d(n)} = m_{c(n)} = 0$, we can obtain the closed-form exact solution as

$$m_n \equiv m_{c(n)} = a((1 + \delta bn)(1 - \delta b)^n) \quad (10)$$

An alternate formulation of (9), to obtain it in single stage, is

$$\frac{m_{n+1} - m_n}{\delta} = b_{n+1}(a - m_n)$$

where

$$b_{n+1} = \frac{\delta b^2 n}{1 + \delta bn} \quad (11)$$

Solving (11), using the PGF computational technique in terms of the initial condition that at $n = 0$, $m_n = 0$, we can obtain the closed-form exact solution as

$$m_n = a((1 + \delta bn)(1 - \delta b)^n) \quad (12)$$

It can be noticed that (12) and (11) are same.

Accordingly, the fault-debugging intensity function is

$$\lambda_n = \frac{m_{n+1} - m_n}{\delta} = \frac{\delta ab^2 n}{1 + \delta bn} ((1 + \delta bn)(1 - \delta b)^n) \quad (13)$$

The equivalent continuous time-space model [3], [28] corresponding to (10) or (12) is

$$m_t = a(1 - (1 + bt)e^{-bt}) \quad (14)$$

which derived as a limiting case of discrete time-space model substituting $t = n\delta$, $\lim_{x \rightarrow 0} (1 + x)^{1/x} = e$, and taking limit $\delta \rightarrow 0$.

3) *Inflection S-shaped Model [1], [29], [30]*: The primary assumption of the models mentioned above is that the fault-debugging rate depends linearly upon the number of undebugged faults. In practice, it has been observed that as debugging progress, the fault debugging rate has three possible trends: decreasing, constant, and increasing. To analyze these trends, we interpret the fault debugging rate as a function of the number of executed computer test-run. Accordingly, we have the following difference equation:

$$\frac{m_{n+1} - m_n}{\delta} = b_{n+1}(a - m_n)$$

where

$$b_{n+1} = b_i + (b_f - b_i) \frac{m_{n+1}}{a} \quad (15)$$

Solving (15), using the PGF computational technique in terms of the initial condition that at $n = 0$, $m_n = 0$, we can obtain the closed-form exact solution as

$$m_n = a \frac{1 - (1 - \delta b_f)^n}{1 + \frac{b_f - b_i}{b_i} (1 - \delta b_f)^n} \quad (16)$$

The growth curve's shape is determined by the parameters b_i and b_f and can be either exponential or S-shaped. However, (16) can be rewritten to define a constant fault-debugging rate, as

$$m_n = a(1 - (1 - \delta b_f)^n) \quad (17)$$

An alternate formulation of the model, where the fault-debugging rate per undebugged fault, b_{n+1} , is a non-decreasing S-shape curve that captures the learning-process phenomenon of the debugging-team, as

$$b_{n+1} = \frac{b}{1 + \beta(1 - \delta b)^{n+1}} \quad (18)$$

Substituting (18) in (15) and solving using the PGF computational technique in terms of the initial condition that at $n = 0$, $m_n = 0$, we can obtain the closed-form exact solution as

$$m_n = a \frac{1 - (1 - \delta b)^n}{1 + \beta(1 - \delta b)^n} \quad (19)$$

Setting $b = b_f$ and $\beta = \frac{b_f - b_i}{b_i}$, it noticed that (19) and (16) are identical.

The fault-debugging rate for the above-specified model is a logistic function and $b_{n+1} \rightarrow b$ as $n \rightarrow \infty$. here β is the learning factor, and it represents the skill and experience gained by the debuggers during debugging. If $\beta = 0$, then $b_{n+1} = b$, that is, constant. The fault-debugging intensity function can be obtained as

$$\lambda_n = \frac{m_{n+1} - m_n}{\delta} = \frac{ab}{1 + \beta(1 - \delta b)^n} \frac{(1 + \beta)(1 - \delta b)^n}{1 + \beta(1 - \delta b)^n} \quad (20)$$

The equivalent continuous time-space model [1], [16], [31] corresponding to (19) or (16) is

$$m_t = a \frac{1 - (1 - e^{-bt})}{1 + \beta(1 - e^{-bt})} \quad (21)$$

which derived as a limiting case of discrete time-space model substituting $t = n\delta$, $\lim_{x \rightarrow 0} (1 + x)^{1/x} = e$, and taking limit $\delta \rightarrow 0$.

4) *Imperfect Fault-Debugging Inflection S-shaped Model [29], [30]*: The primary assumption of the models mentioned above is that the debugged fault is perfectly-debugged with certainty. However, due to human-imperfection and software-complexity, the debugging-team incapable to debug the fault-perfectly, and the debugged fault persist, resulting in a phenomenon known as imperfect fault-debugging. As a result, we write

$$\frac{m_{n+1} - m_n}{\delta} = b_{n+1}(a - m_n)$$

where

$$b_{n+1} = \frac{bp}{1 + \beta(1 - \delta bp)^{n+1}} \quad (22)$$

here p is the probability of perfectly debugging the fault.

Solving (22), using the PGF computational technique in terms of the initial condition that at $n = 0$, $m_n = 0$, we can obtain the closed-form exact solution as

$$m_n = a \frac{1 - (1 - \delta bp)^n}{1 + \beta(1 - \delta bp)^n} \quad (23)$$

The equivalent continuous time-space model [1], [32] corresponding to (23) is

$$m_t = a \frac{1 - e^{-bpt}}{1 + \beta(1 - e^{-bpt})} \quad (24)$$

which derived as a limiting case of discrete time-space model substituting $t = n\delta$, $\lim_{x \rightarrow 0} (1 + x)^{1/x} = e$, and taking limit $\delta \rightarrow 0$.

In the following section, the approach is extended to address the concept of the change-point.

B. Model Formulation

In addition to the software-reliability models above, several models assume a constant or monotonically increasing fault-debugging rate. However, in reality, as the debugging progress, the fault-debugging rate function may be changed at some point in time. The change point's position need not be estimated; it can be judged from the plotted graph of the actual software-reliability data or using a change point analyzer tool. To incorporate the change-point concept in (26), both fault-debugging rate and probability of perfect fault-debugging before and after the change-point τ , are subject to change. Accordingly, we write

$$b_{n+1} = \begin{cases} \frac{b_1 p_1}{1 + \beta(1 - \delta b_1 p_1)^{n+1}} & \text{when } 0 \leq n < \tau \\ \frac{b_2 p_2}{1 + \beta(1 - \delta b_2 p_2)^{n+1}} & \text{when } n \geq \tau \end{cases} \quad (25)$$

here $b_1(b_2)$ represent the fault-debugging rates before(after) the change-point, $p_1(p_2)$ represent the probability of perfect fault-debugging before(after) the change-point, and the change-point τ represents the computer test-run number from whose execution onwards change in the fault-debugging rate is noticed.

Substituting (25) in (22) and solving using the PGF computational technique in terms of the initial conditions at $n = 0$, $m_n = 0$ and at $n = \tau$, $m_n = m_\tau$ respectively, we can obtain the closed-form exact solution as

$$m_n = \begin{cases} a \frac{1-\alpha_1^n}{1+\beta\alpha_1^n} & \text{when } 0 \leq n < \tau \\ a \frac{1-(1+\beta)(1+\beta\alpha_2^\tau)(\alpha_1^\tau\alpha_2^{n-\tau})}{(1+\beta\alpha_1^\tau)(1+\beta\alpha_2^n)} & \text{when } n \geq \tau \end{cases} \quad (26)$$

here $\alpha_1 = 1 - \delta b_1 p_1$ and $\alpha_2 = 1 - \delta b_2 p_2$

The proposed modelling approach given by (26) integrates the joint effects of change-point concept and learning of the debugging personnel in an imperfect fault-debugging environment into software-reliability modelling.

III. DATA ANALYSES AND PARAMETER ESTIMATION TECHNIQUES

To evaluate the progress of the debugging-process during the testing phase of the SDLC, reliability trend analysis is employed. Therefore, before applying the model, it is reasonable to decide whether the dataset exhibits software-reliability growth. In other words, assessing reliability is pointless if the software-reliability dataset during debugging does not exhibit growth. The trend tests that are widely adopted with f fault debugging are [2], [11], [33]:

- Arithmetic Average Test.

$$\rho(f) = \frac{1}{f} \sum_{i=1}^f n_i \quad (27)$$

Decreasing values promote reliability growth.

- Laplace Test.

$$v(f) = \frac{\sum_{i=1}^f (i-1)n_i - \frac{f-1}{2} \sum_{i=1}^f n_i}{\sqrt{\frac{f^2-1}{12} \sum_{i=1}^f n_i}} \quad (28)$$

Negative values promote reliability growth.

To estimate the unknown parameters of the software-reliability models, either the maximum likelihood estimate (abbreviated as MLE) or the least-square estimate (abbreviated as LSE) methods is employed. Software reliability dataset can be collected during debugging in the form of ordered pairs $(n_i, y_i), (i = 0, 1, 2, \dots, f)$ where y_i is the cumulative number of faults debugged by n_i computer test-run ($0 < n_1 < n_2 < \dots < n_f$). Table I tabulated the software-reliability dataset used. The dataset had been obtained in twenty weeks of debugging release 1 of Tandem computer project, and 100 faults were debugged [34]. In other words, the dataset consisted of 20 data pairs $(n_i, y_i), (i = 0, 1, 2, \dots, 20; n_{20} = 20; y_{20} = 100)$. To fit and estimate the parameters of the models, all data points have been utilized.

TABLE I. SOFTWARE-RELIABILITY DATASET [34]

n_i	y_i	n_i	y_i	n_i	y_i	n_i	y_i
1	16	6	49	11	81	16	98
2	24	7	54	12	86	17	99
3	27	8	58	13	90	18	100
4	33	9	69	14	93	19	100
5	41	10	75	15	96	20	100

The graphical plot of the dataset reveals a change-point occurs in the 11th week of debugging, that is, $\tau = 11$.

The Likelihood function L for the unknown parameters with the MVF, that is, m_n takes on the form:

$$L = \prod_{i=1}^f \frac{(m_{n(i)} - m_{n(i-1)})^{x_i - x_{i-1}}}{(x_i - x_{i-1})!} e^{-(m_{n(i)} - m_{n(i-1)})} \quad (29)$$

The statistical package for social sciences (abbreviated as SPSS) is applied to estimate the software-reliability model's parameters for quicker and more precise computations.

IV. MODEL VALIDATION AND COMPARISON CRITERIA

The credibility of the stochastic software-reliability model is decided by its descriptive and predictive power capabilities [1]–[3], [5], [8], [16], [35].

A. Goodness-of-Fit Criteria

To test the validity of the proposed modelling approach in (26), we evaluate three goodness-of-fit test indices based on a real software-reliability dataset [34]. The goodness-of-fit test indices adopted for the purpose are [1]–[3]:

- Sum of squared error (abbreviated as SSE). It assesses the dispersion between estimated values \widehat{m}_{n_i} and the actual data y_i of the dependent variable.

$$SSE = \sum_{i=1}^f (\widehat{m}_{n_i} - y_i)^2 \quad (30)$$

- The Akaike information criterion (abbreviated as AIC).

$$AIC = -2\log L + 2N \quad (31)$$

here N is the number of parameters of a model and the likelihood function L for the parameters with the MVF in (29).

- Root mean square prediction error (abbreviated as RMSPE).

$$RMSPE = \sqrt{\sum_{i=1}^f (Bias^2 + Variation^2)} \quad (32)$$

where

$$Bias = \frac{1}{f} \sum_{i=1}^f (\widehat{m}_{n_i} - y_i)$$

$$Variation = \sqrt{\frac{1}{f-1} \sum_{i=1}^f ((\widehat{m}_{n_i} - y_i) - Bias)^2}$$

- Coefficient of multiple determination (abbreviated as R^2).

$$R^2 = 1 - \frac{\sum_{i=1}^f (\widehat{m}_{n_i} - y_i)^2}{\sum_{i=1}^f (y_i - \sum_{i=1}^f y_i)^2} \quad (33)$$

It should be noted that the smaller value of SSE, AIC and RMSE, indicates the better the fit of the model. On the contrary, the larger the value of R^2 indicates better the fit of the model.

B. Predictive Validity Criterion

The predictive validity of the software-reliability model is defined as its ability to predict the future behavior of the debugging process based on its current and past behavior [35]. Suppose we have corrected x_f faults after the execution of the last computer test-run n_f . First, we utilize the software-reliability dataset up to a computer test-run say $n_e (\leq n_f)$ to obtain the MVF, that is, m_{n_f} . Then, put estimated values in m_n to obtain the number of faults corrected by n_f . Next, we compare it with the actually corrected number x_f . Last, his method is iterated for different values of n_e . The predicted relative error (abbreviated as PRE) is given as

$$PRE = \frac{(\widehat{m}_{n_f} - x_f)}{x_f} \quad (34)$$

When the PRE value is positive/negative, the model overestimates or underestimates the future debugging phenomenon. Acceptable values are within the range $\pm 10\%$ [1], [2], [16].

C. Software-Reliability Evaluation Measures

Based on Section II, the following quantitative assessment are derived [1]–[3], [26]:

- Number of residual faults. Let $\xi(t)$ denotes the number of residual faults after the executing of the computer test-run n^{th} , then

$$\xi_n = a_\infty - a_n \quad (35)$$

here a_∞ denotes the number of faults eventually corrected.

- Software-reliability. The probability of no faults debugged between the n^{th} and $(n+n_o)^{th}$ computer test-run, given that x_d faults have been debugged by the n^{th} computer test-run, is

$$R(n_o | n) = e^{-(m(n+n_o)-m(n))} \quad (36)$$

here $n_o (n_o \geq 0)$ is the mission time.

V. DATASET ANALYSES AND MODEL COMPARISON

To validate the proposed modelling approach, we perform the following dataset analyses and model comparisons. First, we use both of the trend tests presented in Section III to determine whether or not the given software-reliability dataset promotes reliability growth. The parameters of the models under comparison [1], [3], [16], [26], [30] including the proposed modelling approach given in (26) are then estimated using SPSS. Then we perform the goodness-of-fit test described in Section IV-A. Based on the obtained results, we conduct a comparative assessments of the models under comparison. Finally, for the proposed modelling approach, we perform the predictive-validity test described in Section IV-B and estimate the software-reliability evaluation measures defined in Section IV-C.

It is worth noting that, while formulating the proposed modelling approach, we set $\delta = 1$ for simplification and $\tau = 11$ by following the stochastic-behavior of debugging phenomenon on the basis of the adopted dataset.

The models under comparison are given below:

- Exponential model [3], [26] given in (6).
- Delayed S-shaped model [1], [16] given in (10).
- Inflection S-shaped model [1], [30] given in (19).
- Imperfect fault-debugging inflection S-shaped model [1], [30] given in (23).
- The proposed modelling approach given in (26).

A. Software-Reliability Trend Analysis

The fitting result of the reliability dataset for arithmetic average and Laplace trend tests are demonstrated in Fig. 1 and 2, respectively. It is quite clear that the arithmetic average and the Laplace trend factor values are decreasing and entirely negative from the beginning, respectively. Negative Laplace factor values indicate that more faults were debugged in the first half of the debugging time, indicating an increase in reliability. Thus, they point to growth in reliability. Consequently, the dataset is suitable for applying software-reliability models.

B. Goodness-of-Fit Analysis

Table II shows the estimated parameters of the models under comparison using the statistical package statistical package for social sciences (abbreviated as SPSS) and their goodness-of-fit test indices. It can be seen that, when compared to all other models using the SSE, AIC, RMSPE, and R^2 indices, the proposed modelling approach has the lowest SSE, AIC and RMSPE values and the highest R^2 value. Therefore, among the models under comparison, the proposed modelling approach outperformed the others. The enhancement in the performance of the proposed modelling approach is attributed to the inclusion of the change-point concept in modelling software-reliability. These results agree with the findings of previously published works [1], [25].

It is worth noting that the proposed modelling approach's estimated value of a is the closest to the number of debugged faults actually present in the software, and the fault-debugging rates $b_2 < b_1$ clearly show a varying trend that first increases and then decreases following the change-point with an S-shaped varying trend. The latter finding is attributed to the presence of imperfect fault-debugging and learning-process phenomena.

Fig. 3 and 4 show the fitting results of the noncumulative and cumulative reliability dataset for the proposed modelling approach. It is quite clear that the proposed modelling approach fits the dataset perfectly.

C. Predictive Validity Analysis

To estimate the proposed modelling approach parameters, the software-reliability dataset is truncated into several parts. It is noticed that the PRE values differ from one truncation to the next. As a result, Fig. 5 depicts the fitting result of the predictive-validity of the proposed modelling approach.

It is noticed that the proposed modelling approach overestimates the debugging-process from the truncated n_e (70% approx.) onwards. Therefore, the dataset is truncated at t_e

TABLE II. PARAMETER-ESTIMATION AND COMPARISON-CRITERIA RESULTS

Models Under Comparison	Parameter-Estimation						Comparison-Criteria			
	a	$b_1(b)$	$p_1(p)$	b_2	p_2	β	SSE	AIC	$RMSPE$	R^2
Model due to [3], [26]	130	.080	—	—	—	—	233	1789	3.449	.986
Model due to [1], [16]	106	.217	—	—	—	—	357	1796	4.329	.978
Model due to [1], [30]	111	.158	—	—	—	1.21	180	1787	3.073	.989
Model due to [1], [30]	111	.158	1	—	—	1.21	180	1787	3.073	.989
Proposed given in (26)	105	.474	1	.357	.675	5.92	35	1780	1.352	.998

— The parameter is not part of the corresponding model.

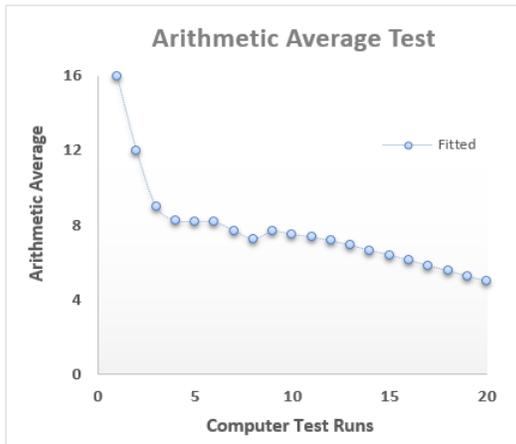


Fig. 1. The Fitting Result of the Reliability Dataset for Arithmetic Average Test Data.

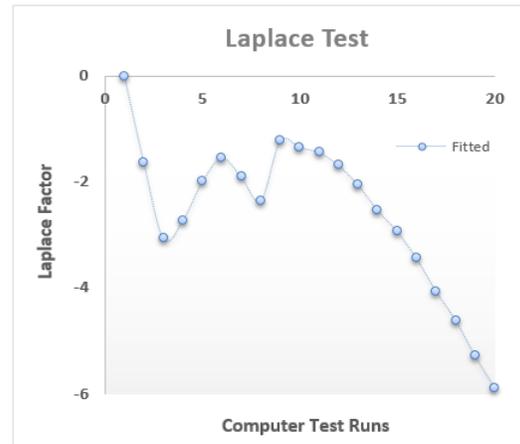


Fig. 2. The Fitting Result of the Reliability Dataset for Laplace Test Data.

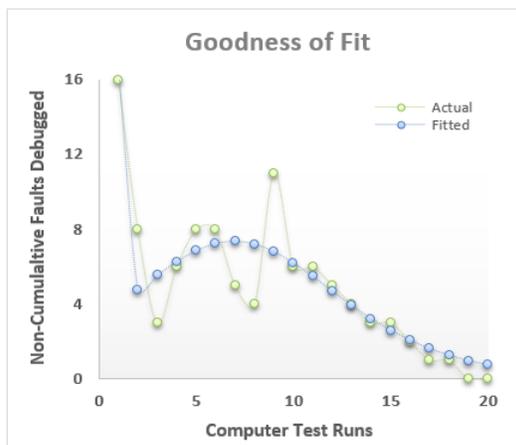


Fig. 3. The Fitting Result of the Noncumulative Reliability Dataset for the Proposed Modelling Approach.

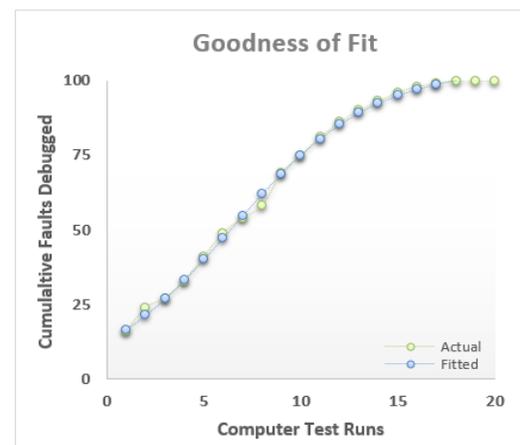


Fig. 4. The The Fitting Result of the Cumulative Reliability Dataset for the Proposed Modelling Approach.

(60% approx.) to obtain the whole dataset. The fitting result of the retrodictive and predictive capabilities is shown in Fig. 6. The points below n_e (marked by the intersection of the horizontal line with the curve) show the retrodictive-capability while the points above n_e show the predictive-capability of the proposed modelling approach.

The proposed modelling approach reveals that about 60% of the debugging-time is enough to predict the debugging-future process's behavior satisfactorily.

D. Software-Reliability Quantitative Measures Analysis

Fig. 7 depicts the fitting result of the number of residual faults for the proposed modelling approach. It is quite clear that the proposed modelling approach fits the dataset perfectly. Following the change-point, every three faults were debugged, one of which was imperfectly debugged according to the parameters' estimated values in Table II. Fig. 8 depicts the fitting result of the estimated software-reliability (three cases: when the mission time is one, two, and three days respectively) for the proposed modelling approach. It is noticeable that the proposed modelling approach reveals that software-reliability

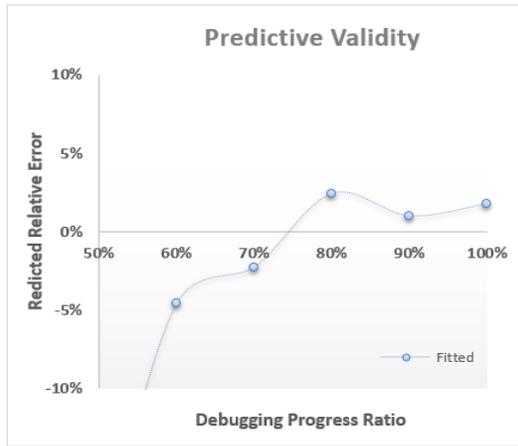


Fig. 5. The Fitting Result of the Predictive Validity for the Proposed Modelling Approach.

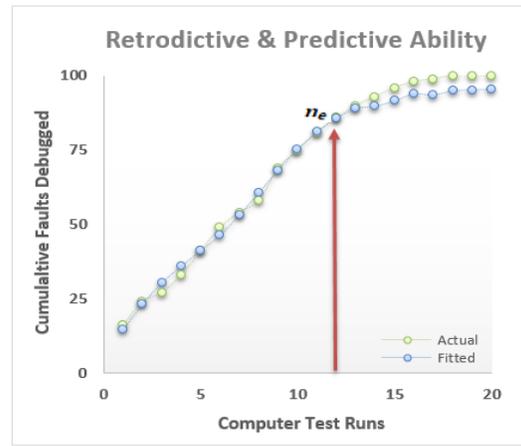


Fig. 6. The Fitting Result of the Retrodictive and Predictive Capabilities for the Proposed Modelling Approach.

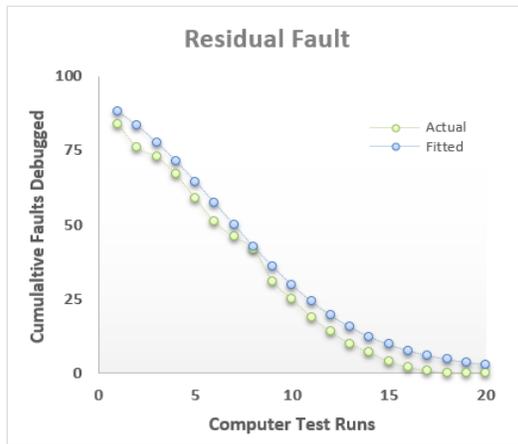


Fig. 7. The Fitting Result of the Number of Residual Faults for the Proposed Modelling Approach.

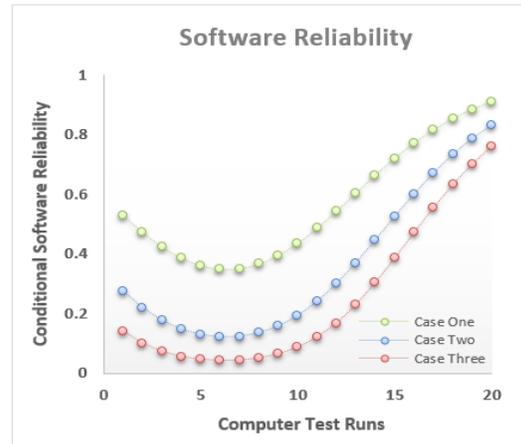


Fig. 8. The Fitting Result of the Estimated Software-Reliability (three cases) for the Proposed Modelling Approach.

monotonically increases after the change-point in all cases.

These types of information provided by the proposed modelling approach enable the developer to decide how best to allocate resources during the software testing and maintenance phases of the SDLC.

VI. CONCLUSIONS

Quantitative measures provided by software-reliability models play a pivotal role in the decision-making process during testing and fault debugging-process. To some extent, the concept of change-point is novel in the discrete time-space software-reliability modelling approach. The NHPP type of discrete time-space software-reliability modelling is well-documented with well-established concepts in the software-reliability engineering literature and can be used to describe the stochastic-behavior of the fault-debugging process.

The proposed discrete time-space NHPP based stochastic software-reliability model integrates the debugger's learning phenomenon with a single change-point under an imperfect fault-debugging environment. The results explained through

a numerical example in the tables and figures provided in Section V are encouraging compared with other well-established models built under imperfect and perfect fault-debugging environments. The conducted numerical example concludes that incorporating the concept of change-point into software-reliability modelling yields the most promising results concerning the descriptive and predictive performances of the proposed modelling approach and software-reliability quantitative measures. As a result, the paper demonstrates a high level of numerical agreement between the results of the proposed modelling approach obtained here and other results presented in the literature.

Finally, the paper shows how starting with fundamental assumptions, the discrete time-space NHPP based modelling approach is becoming more useful with the inclusion of the learning process in debugging and introducing the change-point concept in an imperfect fault-debugging environment. However, we confined ourselves to introducing just a single change-point, and we did not consider the debugging effort expenditures. Hence, it is essential to track and account for the reliability growth concerning the costs spent on software

debugging. Studying the effect of multiple change-point occurrences and software debugging efforts on reliability growth is an ongoing and future research effort.

ACKNOWLEDGMENT

We thank the two anonymous reviewers whose valuable, constructive and insightful comments helped improve and clarify the paper.

REFERENCES

- [1] P. K. Kapur, H. Pham, A. Gupta and P. C. Jha, *Software Reliability Assessment with OR Applications*, Springer, 2011.
- [2] O. Shatnawi, "An integrated framework for developing discrete-time modelling in software reliability engineering," *Quality and Reliability Engineering International*, vol. 32, pp. 2925–2943, 2016.
- [3] S. Yamada, *Software Reliability Modeling: Fundamentals and Applications*, Springer, 2014.
- [4] P. K. Kapur, S. K. Khatri, A. Tickoo and O. Shatnawi, "Release time determination depending on number of test runs using multi attribute utility theory," *International Journal of System Assurance Engineering and Management*, vol. 5, pp. 186–194, 2014.
- [5] M. Xie, *Software Reliability Modeling*, World Scientific, 1991.
- [6] D. H. Lee, I. H. Chang and H. Pham, "Software reliability model with dependent failures and SPRT," *Mathematics*, vol. 8, 1366 pages, 2020.
- [7] N. Qallab and O. Shatnawi, "Modelling software reliability growth phenomenon in distributed development environment," *International Journal of Scientific & Technology Research*, vol. 8, pp. 2176–2184, 2019.
- [8] H. Pham, *System Software Reliability*, Springer, 2007.
- [9] P. K. Kapur, A. G. Aggarwal, O. Shatnawi and R. Kumar, "On the development of unified scheme for discrete software reliability growth modeling," *International Journal of Reliability, Quality and Safety Engineering*, vol. 17, pp. 245–260, 2010.
- [10] T. Pham and H. Pham, "A generalized software reliability model with stochastic fault-detection rate," *Annals of Operations Research*, vol. 227, pp. 83–93, 2019.
- [11] O. Shatnawi, "Measuring commercial software operational reliability: an interdisciplinary modelling approach," *Eksploracja i Niezawodność – Maintenance & Reliability*, vol. 16, pp. 585–594, 2014.
- [12] P. K. Kapur, O. Singh, S. K. Khatri and A.K. Verma, *Strategic System Assurance and Business Analytics*, Springer, 2020.
- [13] K. Deep and M. Jain, S. Salhi, *Performance Prediction and Analytics of Fuzzy, Reliability and Queuing Models: Theory and Applications*, Springer, 2018.
- [14] D. D. Hanagal and N. N. Bhalerao, *Software Reliability Growth Models*, Springer, 2021.
- [15] O. Shatnawi, "Testing-effort dependent software reliability model for distributed systems," *International Journal of Distributed Systems and Technologies*, vol. 4, pp. 1–14, 2013.
- [16] P. K. Kapur, S. Kumar and R. B. Garg, *Contributions to Hardware and Software Reliability*, World Scientific, 1999.
- [17] M. Zhao, "Change-point problems In software and hardware reliability," *Communications in Statistics Theory and Methods*, vol. 22, pp. 757–768, 1993.
- [18] V. Nagaraju, L. Fiondella and T. Wandji, "A heterogeneous single changepoint software reliability growth model framework," *Software Testing, Verification and Reliability*, vol. 29, e1717, 2019.
- [19] Y. Minamino, S. Inoue and S. Yamada, "NHPP-based change-point modeling for software reliability assessment and its application to software development management," *Annals of Operations Research*, vol. 244, pp. 85–101, 2016.
- [20] P. K. Kapur, A. Gupta, O. Shatnawi and V.S.S. Yadavalli, "Testing effort control using flexible software reliability growth Model with change point," *International Journal of Performability Engineering*, vol. 2, pp. 245–262, 2006.
- [21] S. Inoue, J. Ikeda and S. Yamada, "Bivariate change-point modeling for software reliability assessment with uncertainty of testing-environment factor," *Annals of Operations Research*, vol. 244, pp. 209–220, 2016.
- [22] S. Chatterjee and A. Shukla, "Change point-based software reliability model under imperfect debugging with revised concept of fault dependency," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 230, pp. 579–597, 2016.
- [23] V. Nagaraju, L. Fiondella and T. Wandji, "A heterogeneous single changepoint software reliability growth model framework," *Software Testing, Verification and Reliability*, vol. 29, e1717, 2019.
- [24] O. Shatnawi, P. K. Kapur and M. T. Shatnawi, "Release policy, change-point concept, and effort control through discrete-time imperfect software reliability modelling," *International Journal of Computer Applications*, vol. 137, pp. 17–25, 2016.
- [25] D. N. Goswami, S. K. Khatri, R. Kapur, "Discrete software reliability growth modeling for errors of different severity incorporating change-point concept," *International Journal of Automation and computing*, vol. 4, pp. 396–405, 2007.
- [26] S. Yamada and S. Osaki, "Discrete software reliability growth models," *Applied stochastic models and data analysis*, vol. 1, pp. 65–77, 1985.
- [27] A. L. Goel and K. Okumoto, "Time-dependent error-detection rate model for software reliability and other performance measures," *IEEE transactions on Reliability*, vol. 34, pp. 206–211, 1979.
- [28] S. Yamada, M. Ohba and S. Osaki, "S-shaped reliability growth modeling for software error detection," *IEEE Transactions on reliability*, vol. 32, 475–484, 1983.
- [29] O. Shatnawi and M. T. Shatnawi, "Mathematical modelling in software reliability," *Proceedings of the International Conference on Computational & Information Science*, University of Houston, USA, 2009, pp. 425–430.
- [30] O. Shatnawi, "Discrete time NHPP models for software reliability growth phenomenon," *International Arab Journal of Information Technology*, vol. 6, pp. 124–131, 2009.
- [31] M. Ohba, "Software reliability analysis models," *IBM Journal of research and Development*, vol. 28, pp. 428–443, 1984.
- [32] K. Edris and O. Shatnawi, "The Pham Nordmann Zhang (PNZ) software reliability model revisited," *Proceedings of the 10th IASTED International Conference on Software Engineering*, Innsbruck, Austria, 2011, pp.15–17.
- [33] K. Kanoun, M. Kaniche and J. P. Laprie, "Qualitative and quantitative reliability assessment," *IEEE Software*, vol. 14, pp. 77–87, 1997.
- [34] A. Wood, "Predicting software reliability," *IEEE Computers*, vol. 29, pp. 69–77, 1996.
- [35] J. D. Musa, *Software Reliability Engineering: More Reliable Software, Faster and Cheaper*, McGraw-Hill, 2004.

The Role of Data Pre-processing Techniques in Improving Machine Learning Accuracy for Predicting Coronary Heart Disease

Osamah Sami¹, Yousef Elsheikh², Fadi Almasalha³
Faculty of Information Technology
Applied Science Private University
Amman, Jordan 11931

Abstract—These days, in light of the rapid developments, people work day and night to live at a good level. This often causes them to not pay much attention to a healthy lifestyle, such as what they eat or even what physical activities they do. These people are often the most likely to suffer from coronary heart disease. The heart is a small organ responsible for pumping oxygen-rich blood to the rest of the human body through the coronary arteries. Accordingly, any blockage or narrowing in one of these coronary arteries may cause blood not to be pumped to the heart and from it to the rest of the body, and thus cause what is known as heart attacks. From here, the importance of early prediction of coronary heart disease has emerged, as it can help these people change their lifestyle and eating habits to become healthier and thus prevent coronary heart disease and avoid death. This paper improve the accuracy of machine learning techniques in predicting coronary heart disease using data preprocessing techniques. Data preprocessing is a technique used to improve the efficiency of a machine learning model by improving the quality of the feature. The popular Framingham Heart Study dataset was used for validation purposes. The results of the research paper indicate that the use of data preprocessing techniques had a role in improving the predictive accuracy of poorly efficient classifiers, and shows satisfactory performance in determining the risk of coronary heart disease. For example, the Decision Tree classifier led to a predictive accuracy of coronary heart disease of 91.39% with an increase of 1.39% over the previous work, the Random Forest classifier led to a predictive accuracy of 92.80% with an increase of 2.7% over the previous work, the K-Nearest Neighbor classifier led to a predictive accuracy of 92.68% with an increase of 2.58% over the previous work, the Multilayer Perceptron Neural Network (MLP) classifier led to a predictive accuracy of 92.64% with an increase of 2.64% over the previous work, and the Naïve Bayes classifier led to a predictive accuracy of 90.56% with an increase of 0.66% over the previous work.

Keywords—Coronary heart disease; heart; machine learning; data preprocessing; classification technique

I. INTRODUCTION

The heart is one of the most important organs in the human body. It is a small, muscular pumping organ responsible for supplying other organs in the body with oxygen and other important nutrients [1]. This means that a person's life depends on the efficiency of heart function. Therefore, if the heart does not function well, other organs also cannot function well [2].

People, in light of the difficult economic conditions, seek to secure their basic needs by working long hours daily. This

lifestyle often does not take into account the diet and health of these people to ensure their safety [3]. This type often leads to a risk of diseases such as diabetes, high cholesterol and blood pressure at an early age, and all of these diseases, if not controlled, can lead to coronary heart disease [3].

Heart disease is a term that refers to any problem that can affect the heart and blood vessels [2], such as coronary heart disease, congenital heart disease, and rheumatic heart disease [4], which, according to the National Heart, Lung, and Blood Institute ranks among the most dangerous and common diseases in the world.

In coronary heart disease, a complete or partial blockage of the coronary arteries usually occurs due to blood clotting or the accumulation of fatty plaques on the walls, which leads to the inability of the heart to get enough oxygen [5] and thus it is difficult for the heart to function as efficiently as required.

There are two risk factors for coronary heart disease. The first type is stable and cannot be changed, such as age, gender and family history, while the other type depends on lifestyle such as diabetes, smoking, high cholesterol, high blood pressure, high body mass index, and low exercise [6]. However, the second type of risk factors can usually be controlled, according to experts, by changing our lifestyle and diet, and using certain medications if needed.

In recent years, artificial intelligence techniques have been used extensively in the medical fields in order to improve the efficiency of disease diagnosis/classification in its early stages [7]. Among those techniques stand out machine learning techniques, which are a set of statistical models that help the machine learn from past data [8]. In spite of this, it is often difficult to deal with patient data for diagnosis in the early stages due to reasons such as data volume, missing values and noise in the data. But machine learning techniques and their capabilities have helped process such data [9].

Also, it is noticeable regarding data features that they may be incomplete and huge. The range of some data features is small while the range is large for other data features. The type of data features is combined between categorical and numerical; all of this will affect the accuracy of machine learning techniques in diagnosing and classifying diseases in their early stages, including coronary heart disease. Using different techniques to manipulate the features under the so-called data preprocessing techniques and thus improve the

accuracy of machine learning techniques in early prediction of the disease [10]. This paper is organized as follows: The second section is a review of some relevant work. The third section presents the methodology for this research paper. The fourth section is for presenting, evaluating and discussing the results of the research paper. The fifth section is for conclusion and the sixth section is the future work.

II. RELATED WORK

Recently, there has been an increase in the number of papers dealing with the use of machine learning techniques in predicting serious diseases that may affect people's lives, including coronary heart disease. In [11], the researchers applied a logistic regression technique on the Framingham Heart Study dataset to predict the ten-year risk of coronary heart disease. The researchers used 65% of the dataset for the training set. The accuracy obtained was 84.8%.

The researchers in [12] had a contribution by implementing four machine learning algorithms, namely support vector machine (SVM), neural network, XGBoost, and random forest to predict the ten-year risk of coronary heart disease. The researchers also used the Framingham Heart Study dataset to validate the results. The accuracy obtained was 84.8% for support vector machine, 85.4% for neural network, 86.99% for XGBoost, and 84.9% for random forest.

Also, the researchers in [4] contributed to the literature of this field by using boosting adaptive algorithm on four datasets, namely (UCI Cleveland, UCI Switzerland, UCI Long Beach, and UCI Hungarian) to diagnose coronary heart disease. This approach obtained accuracy (97.16% and 80.14% for Cleveland, 98.63% and 89.12% for Hungarian, 93.15% and 77.78% for Long Beach, 100% and 96.72% for Switzerland) for training and testing set respectively.

In [13], the researchers applied three machine learning algorithms, namely support vector machine, neural network, and Hybrid-SVM on the Framingham Heart Study dataset to predict the ten-year risk of heart attack. The accuracy obtained was 86.03% for support vector machine, 84.7% for neural network, and 94% for Hybrid-SVM. However, these results were better for some of the machine learning techniques used than those used for [12].

In [14], the researchers applied six algorithms, namely decision tree, boosted decision tree, random forest, support vector machine, neural network, and logistic regression on the Framingham Heart Study dataset to predict the ten-year risk of coronary heart disease. The data was divided into 80% training and 20% testing. The researchers used R Studio and Rapid-Miner in their work. The researchers used three techniques to deal with missing values. The first technique is to ignore missing values, and obtained accuracy of 85% for the decision tree, 63% for the boosted decision tree, and 63% for logistic regression. All this while using the Rapid-Miner tool. Whereas, the R studio tool enabled the researchers to obtain the accuracy of 84% for the decision Tree, 85% for the boosted decision tree, and 84% for logistic regression. Analysis of complete case is the second technique used, as the Rapid-Miner tool enabled the researchers to obtain accuracy of 54% for the decision tree, 64% for the boosted decision tree, 65% for the random forest, 69% for the support vector machine, 69%

for the neural network, and 68% for logistic regression. R studio tool obtained accuracy 67%, 81%, 79%, 69%, 67%, and 68% for the decision tree, boosted decision tree, random forest, support vector machine, neural network, and logistic regression respectively. The final technique is to be replaced with the average, and the accuracy obtained while using the Rapid-Miner tool was 62% for the decision tree, 62% for the boosted decision tree, 63% for the random forest, 68% for the support vector machine, 68% for the neural network, and 67% for logistic regression. Whereas, the R Studio tool enabled the researchers to obtain an accuracy of 84% for the decision tree, 84% for the boosted decision tree, 78% for the random forest, 68% for the support vector machine, 71% for the neural network, and 66% for logistic regression.

However, other researchers such as those in [15] applied only one algorithm which is the logistic regression on the Framingham Heart Study dataset to predict the ten-year risk of coronary heart disease. This approach obtained better accuracy of 86.6% than ever.

In [16], the researchers applied the same previous method of logistic regression to the Framingham Heart Study dataset to predict a heart attack. This approach obtained an accuracy of 87%.

Other researchers such as those in [17] applied the neural network algorithm to real data from patient of Paris Hôtel-Dieu University Hospital to diagnose coronary heart disease. Their approach used a different number of input factors (6 to 14). The approach obtained 63% for features (age, diabetes, hypertension, obesity, smoking, family anamnesis of CHD), 76% for features (age, sex, diabetes, hypertension, obesity, smoking, family anamnesis of CHD), 77% for features (age, sex, diabetes, hypertension, obesity, smoking, family anamnesis of CHD, glycaemia, cholesterol total), 81% for features (age, sex, diabetes, hypertension, obesity, smoking, family anamnesis of CHD, TG, cholesterol 0.81 69 79 total, HDL, LDL, glycaemia), 83% for features (age, sex, diabetes, hypertension, obesity, smoking, family anamnesis of CHD, carotid plaque), 87% for features (diabetes, hypertension, obesity, smoking, family anamnesis of CHD, PWV index), 91% for features (diabetes, hypertension, obesity, smoking, family anamnesis of CHD, carotid plaque, PWV index), 93% for features (diabetes, hypertension, obesity, smoking, family anamnesis of CHD, TG, cholesterol, HDL, 0.93 80 92 LDL, glycaemia, carotid plaque, PWV index), 77% for features (age, sex, diabetes, hypertension, obesity, smoking, family anamnesis of CHD, glycaemia, 0.77 53 87 cholesterol total, cGFR), and 77% for features (age, sex, diabetes, hypertension, obesity, smoking, family anamnesis of CHD, glycaemia, cholesterol total, left ventricular hypertrophy)

Those in [18] applied the deep belief algorithm to the KNHANES-6 dataset to predict the risk of coronary heart disease and obtained an accuracy of 82%. However, the researchers applied the genetic algorithm to improve the deep belief network and the obtained accuracy was 74%.

In [19], the researchers applied a logistic regression and neural network to the KNHANES-VI dataset to predict the risk of coronary heart disease. However, this approach obtained accuracy 86.11% for the logistic regression and 87.04% for the neural network. The researchers used a distinct correlation

analysis to improve the accuracy of the neural network to become 87.63%.

In other research such as [20], the researchers applied Naïve Bayes, KNN, random forest, decision tree, SVM, logistic regression, and the ensemble classification approach to the NHANES and Framingham Heart Study dataset, to monitor the risk of chronic diseases. For the NHANES dataset, the decision tree algorithm obtained an accuracy of 97.6%, 96.5% for the ensemble approach, 80.8% for the KNN, 96.4% for logistic regression, 95.7% for Naïve Bayes, 98.5% for random forest, 95.4% for SVM. Whereas, the results for Framingham Heart Study dataset were as follows: The decision tree obtained an accuracy of 90%, 89.3% for the ensemble approach, 90.1% for the KNN, 90% for the logistic regression, 89.9% for Naïve Bayes, 90.1% for random forest, and 90.2% for SVM.

Similarly, the researchers of [21] applied Naïve Bayes, KNN, random forest, decision tree, SVM, logistic regression, neural network, and the ensemble classification approach to the NHANES and Framingham Heart Study dataset to predict Cardiovascular disease. For the NHANES dataset, the decision tree algorithm obtained an accuracy of 97.6%, 96.5% for the ensemble approach, 80.8% for the KNN, 96.4% for logistic regression, 95.7% for Naïve Bayes, 98.5% for random forest, 95.4% for SVM, 98.8% neural network. Whereas, the results for Framingham Heart Study dataset were as follows: The decision tree obtained an accuracy of 90 89.3% for the ensemble approach, 90.1% for the KNN, 90% for logistic regression, 89.9% for Naïve Bayes, 90.1% for random forest, 90.2% for SVM, and 89% for neural network.

In [22], the researchers applied neural network algorithm on the Framingham Heart Study dataset to predict the heart disease. The accuracy obtained was 90% .

Other researchers such as those in [23] applied the k-nearest neighbor (KNN), Logistic regression (LR), linear discriminant analysis (LDA), support vector machine (SVM), classification and regression tree (CART), gradient boosting (GB), and random forest (RF) the Framingham Heart Study dataset to detect the heart disease. The accuracy obtained was 81% for KNN, 83% for LR, 83% for LDA, 82% for SVM, 75% for CART, 83% for GB, and 83% for RF. After that some ensemble techniques were applied and the accuracy was improvement to 86%.

Those in [24] applied k-nearest neighbor, decision tree, random forest logistic regression, and neural network on the Framingham Heart Study dataset to predict the heart disease. The accuracy obtained was 86% for k-nearest neighbor, 77% for decision tree, 86% for random forest, 85% for logistic regression, and 85% for neural network.

Most of previous researchers using either the UCI dataset or Framingham Heart Study dataset, UCI dataset is a good dataset for diagnosis, and prediction heart disease, but this data has some limitations, first limitation is the size of instance of the data is bit small, second limitation the dataset does not include some important features for predict and diagnose heart disease such as LDL cholesterol, HDL cholesterol, smoking or not smoking, diastolic blood pressure, systolic blood presume, number of cigarettes per day, body mass index, and family history of any type of heart disease. This means this data does not fit to diagnose or predict heart disease for smoking patients,

patient with history of blood pressure, obesity patients, and patients with a family history of heart disease. also, Framingham Heart Study dataset is good data for predict heart disease, this data does not contain feature for family history of any type of heart disease. This means this data specific for patient with no family history of any type of heart disease.

Despite this and many other researches, the field is still open for researchers to conduct their experiments in order to improve the accuracy of the machine learning techniques for predicting diseases that pose a risk to human life, including coronary heart disease.

III. RESEARCH METHOD

It is unfortunate to hear that there is an increase in the number of patients diagnosed with coronary heart disease (angina or heart attack) day after day. High blood pressure, high cholesterol, uncontrolled diabetes, smoking, and a diagnosis of cardiovascular impairment and other risks, all increase the chance of diagnosis with coronary heart disease in the future. Therefore, an accurate system needed to help the patient protect him/herself from the risk of coronary heart disease, relying in this on the patient's demographic information, medical history, medical examination, behavior, and laboratory examination.

Many researchers have developed machine learning models using different classification algorithms such as decision tree, Naïve Bayes, SVM, KNN, and neural network. Most of these models were utilizing the Cleveland Heart Diseases dataset to predict coronary heart diseases, but few were using the Framingham Study dataset. This paper uses the Framingham Study dataset to validate the resulting model since it includes features for most of the potential risk factors for coronary heart disease and some of these features are not found in the most common dataset of heart disease namely, Cleveland Heart Disease dataset. In this paper, five machine learning classification algorithms were used such as decision tree, Naïve Bayes, neural network, random forest, and KNN. These five algorithms used the Framingham Heart Study dataset with two events for target (output) features to predict coronary heart disease, as a number of different Data Preprocessing techniques will be used to improve the accuracy of machine learning models for predicting coronary heart disease.

A. Dataset

The Framingham Heart Study dataset is the first long-term epidemiological study concerned with the possible causes of cardiovascular disease that began in 1948 in Framingham, Massachusetts [20]. The Framingham Heart Study dataset identified the prospective risk factors of cardiovascular diseases and their effects [20], [25].

The dataset contains 19 input features divided into demographic features(Age, Gender), behavioral features(Current Smoker, Cigarettes Per Day, Body Mass Index), medical history features(Prevalent Coronary Heart Diseases, Prevalent Angina Pectoris, Prevalent Myocardial Infarction, Prevalent Stroke, Prevalent Hypertensive, Use Blood Pressure Drugs,

Diabetes), medical examination features(Systolic Blood Pressure, Diastolic Blood Pressure, Heart Rate) and laboratory testing features(Glucose, High-Density Lipoprotein, Low-Density Lipoprotein, Total Cholesterol), and two features for prediction (Angina Pectoris, Myocardial Infarction).

B. Data Preprocessing

Data preprocessing is a group of techniques that are applied on the data to improve the quality of the data, such as handling missing values, convert the type of feature and many other techniques [10].

1) *Impute Missing Values By Knn*: knn for missing values working by calculate the distance or similarity to find the most similar case in the dataset and change the missing value with it [26], by applying (1).

$$Dist(X, Y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (1)$$

Where X_i some known values, and Y_i some values that should predict their values.

2) *Min_Max Normalization*: This method is convert each numerical feature value into new value depending on the minimum and maximum values of the feature [27], by applying (2).

$$\bar{X} = \frac{X - Min}{Max - Min} \quad (2)$$

Where Min is the smallest value in the selected feature, Max is the biggest value in the selected feature, \bar{X} is a new select value after applying normalization, X is a selected value from a numerical feature.

3) *Z-Score Standardization*: This method is convert each numerical feature value into new value depending on the standard deviation and Mean of the feature [28], by applying (3).

$$\bar{X} = \frac{X - \mu}{\sigma} \quad (3)$$

\bar{X} is a new select value after applying standardization, X is a selected value from a numerical feature.

4) *One Hot Encoding*: One Hot Encoding splits the categorical feature into a separate number of features depending on the number of the cases in the original categorical feature, and give 0 for absence and 1 for presence in each new feature [29].

5) *Ordinal Encoding*: In this technique, each case in the categorical feature is converted into integer value [29].

6) *Equal Width Discretization*: This is an easy method that sorting the values of numerical feature and split the range of sorting values into predefined equal-width bins [30] by applying (4) and (5).

$$W = \frac{V_{Max} - V_{Min}}{K} \quad (4)$$

$$Boundaries = V_{Min} + (i * W) \quad (5)$$

Where W is the width of the bin, V_Max is the maximum value in the selected numerical feature, V_Min is the minimum in the selected numerical feature, $i = 1, \dots, k-1$.

7) *Equal Frequency Discretization*: In this method, firstly sorting the values in ascending order. Split the range of sorting values into predefined number of equal-frequency bins by applying $\frac{N}{K}$, each bin has the same number of values [30].

C. Classification Algorithms

Classification is a supervised machine learning model used with a label's output to determine the result of the model from many labels or categorical input data [31]. The classifier model is built for training depending on many known labelled or categorical feature of input data [31]. In the next step, the model tested by using the test set to identify the number of the known target for the model and try to correct the unknown target for the model [31].

1) *ID3 Decision Tree*: Each decision tree contains a root node, leaf node, internal node and branches. In ID3 decision tree, all features set as root node, and after that the features are divided by finding the entropy which it utilizes the measure of the harmony in the data; the values of entropy is between 0 and 1 [7], and information gain is the difference between the feature and the subsets of this feature [7]. Entropy and information gain can be found by applying (6) and (7), and the feature which has the highest information gain value is selected as the root node of the tree [7].

$$Entropy(F) = \sum_{i=1}^C (-P_i \log_2 P_i) \quad (6)$$

$$Gain(F, A) = Entropy(F) - \sum_{i=1}^K \left(\frac{|F_i|}{|F|} Entropy(F_i) \right) \quad (7)$$

Where C is number of outputs, P_i is probability of occurrences each output from all output, K number of spilt data, F feature with some data, F_i spilt data from feature F.

2) *Random Forest*: Random forest is a classification algorithm [32] works by creating many decision trees from the dataset [32]. The features are selected randomly from the training set to build the trees in the random forest [32]. After building each decision tree and find the result of each the tree, applying majority voting to decide the final result of the random forest [32]. In the process of building each decision tree, the randomization is applied to find the value the split node.

3) *K-Nearest Neighbours*: KNN is a lazy supervised machine learning algorithm that used to predict and classify unknown data from known data by measuring the distance between them [33]. The distance metric is using to measure the distance between point from testing data with all the point in training data [33], [34], the distance can calculate by applying (8).

$$Cosine(X_i, Y_i) = \frac{\sum_{i=1}^n X_i Y_i}{\sqrt{\sum_{i=1}^n (X_i)^2} \sqrt{\sum_{j=1}^n (Y_i)^2}} \quad (8)$$

Where X_i some values belong to known output class, and Y_i some values that should predict their output class.

4) *Multilayer Perceptron Neural Network*: Artificial Neural network structure is the same as the brain of human [35]. Multilayer perceptron (MLP) that contains more than one layer(**input layer, hidden layer(s), output layer**) [36].

First, in the neural network before start training from the dataset, the value of weight (w) is randomly assigned [36]. After that, the neural network begin the training [36]. Sigmoid is a non-linear activation function commonly use in feedforward neural networks to find the output [37]. Sigmoid function can be calculated by applying (9).

$$F(X) = \frac{1}{1 + \exp^{-X}} \quad (9)$$

Back Propagation algorithm is commonly used to train Multilayer Perceptron Neural Network In the first step of this algorithm is to compare between predict output (\hat{Y}) and actual output (Y) to find the error between them, this error return to neural network and the weight change depending on this error, and the weight numerical change until the value of (\hat{Y}) become closer to (Y) [36].

5) *Naïve Bayes*: Naïve Bayes is a statistical classification algorithm that works on the basis of Bayes' theory, and Naïve Bayes assumes that each feature is separate, and each variable is distinct in prediction and occurrence [3]. Naïve Bayes uses the prior probability of Bayes theorem to calculate the likelihood of the relationship between each feature in the test data with each target, the target with the highest probability is selected as the result of the model [38]. The probability can be found using (10):

$$P(C_i|F_j) = \frac{P(F_j|C_i)P(C_i)}{P(F_j)} \quad (10)$$

Where $P(C_i|F_j)$ probability of specific class (C_i) appear with specific feature (F_j) from the total of all Features F and Classes C, $P(C_i)$ probability of specific class (C_i) from the total of all classes (C), $P(F_j|C_i)$ probability of specific feature (F_j) appear with specific class (Ci) from the total of all features (F) and classes(C), $p(F_j)$ probability of specific feature (F_j) from the total of all features (F).

D. Stratified KFold Cross Validation

Cross validation is a static method used to test an algorithm by dividing the data set into a training set used to train the model and the test set used to evaluate the model performance [39]. In cross-validation, every point has the same chance of being used in the test [39]. In kfold, the dataset is evenly divided into k number of fields [39]. Stratified KFold means that each fold has the same class naming distribution in the original dataset [40]. For each iteration, one test folds and others are used for training [39].

E. Tool

RapidMiner is a data science software platform developed by the company of the same name that provides an integrated environment for data preparation, machine learning, deep learning, text mining, and predictive analytics [41]. In machine learning, RapidMiner can be used for feature processing, dataset segmentation, model training, model testing, network research, and performance evaluation [41].

IV. RESULTS AND DISCUSSION

In this paper, five machine learning classification techniques used to predict two primary CHD events, namely, angina pectoris (528 yes, 2735 no) and myocardial infarction (308 yes, 2955 no).

A. Performance Evaluation

Performance evaluation is a group of equations used to measure the effectiveness of the classifier or the model [42]. Below is the definition of some essential terms used in the equations of performance evaluation:

1) *True Positive (TP)*: The person is healthy and also predict as healthy [42]

2) *False Positive (FP)*: The person is healthy, but predict as sick [42]

3) *True Negative (TN)*: The person is sick and predict as sick [42]

4) *False Negative (FN)*: The person is sick, but predict as healthy [42]

B. Confusion Matrix

The confusion matrix is used to analyze the ability of classifier or model to identify the classes of the dataset [42]. TN and TP are referred to correct classification, while FN and FP are referred to wrong classification [42]. For the accurate classifier or model, TP and TN are classified more than FN and FP [42], as shown in Table I

TABLE I. CONFUSION MATRIX

	Negative(Actual)	Positive(Actual)
Negative(Predict)	TN	FN
Positive(Predict)	FP	TP

C. Performance Metrics

1) *Accuracy*: Accuracy is an evaluation metric of the total number of predictions the model or the classifier gets right [43]. The accuracy can be calculated by applying (11).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (11)$$

2) *Precision*: Precision is used to identified is the diagnosis or the predicted result is close to the real result [43]. Precision can be calculated by apply (12).

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

3) *F-Measure*: F-Measure refers to the mean of consistency between Precision and Recall [43]. F-Measure can be calculated by apply (13).

$$F - Measure = 2 * \left(\frac{Recall * Precision}{Recall + Precision} \right) \quad (13)$$

4) *Sensitivity(Recall)*: Sensitivity is true positive rate measure. In other words, the rate of healthy person diagnosis or predict as healthy [43]. Sensitivity can be calculated by apply (14).

$$Sensitivity = \frac{TP}{TP + FN} \quad (14)$$

5) *Specificity*: Specificity is true negative rate measure. In other words, the rate of sick person diagnosis or predict as sick [43]. Specificity can be calculated by apply (15).

$$Specificity = \frac{TN}{TN + FP} \quad (15)$$

D. Algorithms Confusion Matrix

Below Table II, Table III, Table IV, Table V, and Table VI, shown the number of correct predict (**True Positive and True Negative**) and wrong predict (**False Positive and False Negative**) for each algorithm.

TABLE II. ID3 DECISION TREE CONFUSION MATRIX

	No(Actual)	Yes(Actual)
No(Predict)	2879	205
Yes(Predict)	76	103

TABLE III. RANDOM FOREST CONFUSION MATRIX

	No(Actual)	Yes(Actual)
No(Predict)	2921	201
Yes(Predict)	107	34

TABLE IV. K-NEAREST NEIGHBORS CONFUSION MATRIX

	No(Actual)	Yes(Actual)
No(Predict)	2930	214
Yes(Predict)	94	25

TABLE V. NEURAL NETWORK CONFUSION MATRIX

	No(Actual)	Yes(Actual)
No(Predict)	2923	208
Yes(Predict)	100	32

TABLE VI. NAÏVE BAYES CONFUSION MATRIX

	No(Actual)	Yes(Actual)
No(Predict)	2666	239
Yes(Predict)	69	289

E. Accuracy without Data Preprocessing

TABLE VII. ACCURACY WITHOUT DATA PREPROCESSING

Algorithms	Accuracy (%)
Decision Tree	87.19
Random Forest	92.68
MLP	90.56
KNN	90.50
Naïve Bayes	89

F. Algorithms Evaluation Result

TABLE VIII. MODEL EVALUATION RESULT

Algorithms	Accuracy	Precision	F-Measure	Sensitivity	Specificity
ID3 Decision Tree	92.8%0	93.57%	96.13%	98.85%	34.80%
Random Forest	91.39%	93.36%	95.35%	97.43%	33.50%
K-Nearest Neighbors	92.68%	93.20%	96.08%	99.15%	30.53%
Neural Network	92.64%	93.36%	96.06%	98.92%	32.51%
Naïve Bayes	90.56%	91.79%	94.54%	97.48%	54.77%

G. Accuracy Comparison

TABLE IX. ACCURACY COMPARISON

Algorithms	Previous Accuracy	Proposed Accuracy	Dataset Event
Decision Tree	90% [20], [21]	91.39%	Myocardial Infraction
Random Forest	90.1% [20], [21]	92.80%	Myocardial Infraction
K-Nearest Neighbors	90.1% [20], [21]	92.68%	Myocardial Infraction
Neural Network	90% [22]	92.64%	Myocardial Infraction
Naïve Bayes	89.9% [20], [21]	90.56%	Angina Pectoris

H. Discussion

In this research paper, a set of machine learning techniques used to predict two events of coronary heart disease namely, Angina Pectoris (528 Yes, 2735 No), and Myocardial Infarction (308 Yes, 2955 No). Despite the previous researchers used many data preprocessing techniques, the results obtained from this work were very encouraging compared to other studies that use the same data set to calculate accuracy as shown in Table IX.

It is noted that the techniques that have been used to improve the accuracy of machine learning models or classifiers in predicting coronary heart disease have proven effective and thus have achieved better results than previous research.

For example, [20] and [21] used the same data set and obtained by applying the decision tree algorithm a predictive accuracy of 90% to predict coronary heart disease (CHD), while this research paper obtained an accuracy of 91.39%, with a positive increase of 1.39% as shown in Table IX.

Also, this research paper and through the application of the random forest algorithm obtained a predictive accuracy of CHD 92.80%, shown in Table IX, which is higher than the result obtained in the decision tree algorithm in this research paper on the one hand, and on the other hand, higher and better than the results obtained by [20] and [21] and that was 90.10%, with a positive increase of 2.7%.

As for the use of the MLP algorithm in predicting CHD, researchers in [21] obtained an accuracy of predicting the

disease 90%, while this research paper obtained a better accuracy of 92.64%, with a positive increase of 2.64% shown in Table IX.

Regarding the use of the KNN algorithm, researchers in [20] and [21] obtained a prediction accuracy of 90.10%, which is less than the prediction accuracy of the disease obtained in this research paper, which is 92.68%, which was applied to calculate the missing values and equal width discretization, with a positive increase of 2.58% as shown in Table IX.

The application of the Naïve Bayes in this research paper obtained a predictive accuracy of coronary heart disease 90.56% as shown in Table IX, which is better than the predictive accuracy of 89.90% obtained in [20].

After applied data preprocessing techniques, this proposed work obtained accuracy better than previous researches used the same dataset and same techniques, such as, [13] that published in **2018** was obtained accuracy 84.7% for neural network; decision tree was obtained 85%, random forest was obtained 79%, and neural network was obtained 71% in [14] that published in **2017**; [20] that published in **2017** was obtained accuracy 90.1% for KNN, 90.1% for random forest, 89.9% for Naïve Bayes, and 90% for decision tree; the accuracy in [21] that published in **2018** was obtained 90.1% for KNN, 90.1% for random forest, 89.9% for Naïve Bayes, and 90% for decision tree; in **2020** the [22] was obtained accuracy 90% for neural network; [23] that published in **2021** was obtained accuracy 81% for KNN, 75% for decision tree, and 83% for random forest; decision tree was obtained 77%, random forest was obtained 86%, KNN was obtained 86%, and neural network was obtained 85% in [24] that was published in **2021**.

Although the results obtained in predicting coronary heart disease in terms of accuracy were not as significant as it should be, it may contribute to an increase in the number of cases with the correct diagnosis of the disease and at the same time reduce the number of cases that are incorrectly diagnosed with coronary heart disease and thus save lives

V. CONCLUSION

The heart is among the most important organs of the human body, as any problem with it can damage other important organs in the body, such as the brain. All doctors around the world warn of the sharp increase in the number of heart patients, being a serious disease that may lead to serious complications such as heart failure and cardiac arrest, both of which often lead to death if not diagnosed early.

In this paper, the researchers contributed to improving the accuracy of machine learning classification models in predicting two primary coronary heart disease events, namely, angina pectoris and myocardial infarction through the use of a number of feature processing techniques such as normalization, standardization, and discretization. For the purpose of validating the results obtained, the data set of the Framingham Heart Study was used with two main events (angina pectoris and myocardial infarction (heart attack)), due to its containment and after consulting with cardiologists about the most common factors causing coronary heart disease.

After using data preprocessing techniques on the dataset, the accuracy of machine learning algorithms for predicting coronary heart disease improved unevenly. For example, the improvement in accuracy prediction of CHD was 4.2% when using the ID3 decision tree algorithm, 0.14% when using the random forest algorithm, 3.18% when using the KNN algorithm, 2.08% when using the MLP algorithm, and 1.36% when using the Naive Bayes algorithm as shown in Table VII and Table VIII.

However, the best prediction accuracy obtained for the ID3 decision tree algorithm is at 91.39% when applied the equal width discretization method. Whereas, the random forest algorithm achieved a prediction accuracy of 92.80% when applied the equal width discretization and applied normalization methods. The MLP algorithm achieved an improvement in accuracy prediction by 92.64% when using one hot encoding technique. 92.68% represents the predictive accuracy obtained with the KNN algorithm when applied the ordinal coding and standardization techniques. However, all of the predicted values obtained were in the case of a myocardial infarction event. Whereas, the value obtained from Naive Bayes algorithm was 90.65% in the case of angina pectoris and when applied equal frequency discretization. The results obtained confirm the importance of using data preprocessing techniques in improving the accuracy performance of machine learning algorithms for predicting coronary heart disease compared to previous published research with the same objectives.

In the end, the presence of a correlation between some serious diseases such as the occurrence of stroke, high blood pressure, cardiovascular disease and coronary heart disease leads us in the future to predict such diseases and the effect of each of them on the occurrence of coronary heart disease on the one hand, and on the other hand the effect of the occurrence of coronary heart disease, on these diseases, to prevent death. This is because the patient in such cases does not have enough time to go to the doctor to see him and save his life.

VI. FUTURE WORK

In the future work, more data preprocessing techniques and more machine learning classification algorithms can apply to get better results than the ones that obtained in this proposed work.

Machine learning algorithms can be used to analyze big data to forecast coronary heart disease. This means that a huge amount of data means that the prediction will get better because more data means that the result is more accurate.

Sometimes the patient does not have enough time to go to the doctor, so develop a website or smartphones application for the graphical user interface solve this problem, and this site makes the prediction process easier and from the patient's place where the user only enters his risk factors information and the result is presented to him immediately.

ACKNOWLEDGMENT

The authors are grateful to the Applied Science Private University, Amman, Jordan, for the financial support granted to this research paper.

REFERENCES

- [1] S. Nashif, R. Raihan, R. Islam, and M. H. Imam, "Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System," *World Journal of Engineering and Technology*, vol. 6, no. 4, pp. 854-873, Nov, 2018. doi: 10.4236/wjet.2018.64057.
- [2] J. S. Sonawane and D. R. Patil, "Prediction of heart disease using multilayer perceptron neural network," *International Conference on Information Communication and Embedded Systems (ICICES2014)*, pp. 1-6, Feb, 2014. doi: 10.1109/ICICES.2014.7033860.
- [3] S. A. Pattekari and A. Parveen, "Prediction system for heart disease using Naive Bayes," *International Journal of Advanced Computer and Mathematical Sciences*, vol. 3, no. 3, pp. 290-294, 2012.
- [4] K. H. Miao, J. H. Miao, and G. J. Miao, "Diagnosing Coronary Heart Disease Using Ensemble Machine Learning," (IJACSA) *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 10, pp. 30-39, 2016. doi: 10.14569/IJACSA.2016.071004.
- [5] K. H. Miao and J. H. Miao, "Coronary Heart Disease Diagnosis using Deep Neural Networks," (IJACSA) *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, pp. 1-8, 2018. doi: 10.14569/IJACSA.2018.091001.
- [6] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7675-7680, May, 2009. doi: <https://doi.org/10.1016/j.eswa.2008.09.013>.
- [7] S. Bashir, U. Qamar, F. H. Khan, and M. Y. Javed, "An Efficient Rule-Based Classification of Diabetes Using ID3, C4.5, & CART Ensembles," in *2014 12th International Conference on Frontiers of Information Technology*, pp. 226-231, Dec 2014. doi: 10.1109/FIT.2014.50.
- [8] T. Panch, P. Szolovits, and R. Atun, "Artificial intelligence, machine learning and health systems," *Journal of Global Health*, vol. 8, December. 2018. doi: 10.7189/jogh.08.020303.
- [9] G. D. Magoulas and A. Prentza, "Machine Learning in Medical Applications," *Machine Learning and Its Applications*, vol. 2049, pp. 300-307, 2001. doi: https://doi.org/10.1007/3-540-44673-7_19
- [10] S. A. Alasadi and W. Bhaya, "Review of data preprocessing techniques in data mining," *Journal of Engineering and Applied Sciences*, vol. 12, no. 16, pp. 4102-4107, Sep 2017.
- [11] A. Gupta and V. Khathuria, "Framingham heart study," *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 4, no. 11, pp. 55-58, 2018.
- [12] K.V. Nagendra and M. Ussenaiah, "Analysis of classification algorithms on heart diseases data using association rule mining," *International Journal of Computational Engineering Research(IJCER)*, vol. 08, no. 6, pp. 39-46, 2018.
- [13] K. V. Nagendra and M. Ussenaiah, "Support vector machine and neural network classification improved by bagging," *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 4, no. 2, pp. 125-130, 2018.
- [14] J. Beunza, E. Puertas, E. García-Ovejero, G. Villalba, E. Condes, G. Condes, and M. F. Landecho "Comparison of machine learning algorithms for clinical event prediction(risk of coronary heart disease)," *Journal of Biomedical Informatics*, vol. 97, p. 103257, Sep.2017. doi:<https://doi.org/10.1016/j.jbi.2019.103257>.
- [15] A. S. T. Nishadi, "Predicting heart diseases in logistic regression of machine learning algorithms by python jupyterlab," *International Journal of Advanced Research and Publications*, vol. 3,no. 8, pp. 69-74, Aug 2019.
- [16] A. Bhardwaj, A. Kundra, B. Gandhi, S. Kumar, A. Rehalia, and M. Gupta, "Prediction of heart attack using machine learning," *IITM Journal of Management and IT*, vol. 10, no. 1, pp. 20-24, 2019.
- [17] A. Valle, A. Cinaud, V. Blachier, H. Lelong, M. E. Safar, and J. Blacher "Coronary heart disease diagnosis by artificial neural networks including aortic pulse wave velocity index and clinical parameters," *Journal of Hypertension*, vol. 37, no. 8, pp. 1682-1688, Aug. 2019. doi:10.1097/HJH.0000000000002075.
- [18] K. Lim, B. M. Lee, U. Kang, and Y. Lee "An optimized DBN-based coronary heart disease risk prediction," *International Journal of Computers Communications & Control*, vol. 13,no. 4, pp. 492-502, Jul 2018. doi: <https://doi.org/10.15837/ijccc.2018.4.3269> .
- [19] J.K. Kim and S. Kang, "Neural network-based coronary heart disease risk prediction using feature correlation analysis," *Journal of Healthcare Engineering*, vol. 2017, pp. 1-13, 2017. doi: <https://doi.org/10.1155/2017/2780501>.
- [20] N.S. Rajliwall, G. Chetty, and R. Davey, "Chronic disease risk monitoring based on an innovative predictive modelling framework," *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1-8, 2017. doi: 10.1109/SSCI.2017.8285257, 1-8, 2017.
- [21] N. S. Rajliwall, R. Davey, and G. Chetty, "Machine learning based models for cardiovascular risk prediction," *2018 International Conference on Machine Learning and Data Engineering(iCMLDE)*, pp. 142-148, 2018. doi: 10.1109/iCMLDE.2018.00034, 142- 148, 2018.
- [22] I. D. Mienye, Y. Sun, and Z Wang, "Improved sparse autoencoder based artificial neural network approach for prediction of heart disease," *Informatics in Medicine Unlocked*, vol. 18, 100307, 2020.
- [23] P. Puvar, N. Patel, A. Shah, R. Solanki, and D. Rana, "Heart Disease Detection using Ensemble Learning Approach," *International Research Journal of Engineering and Technology (IRJET)*, vol. 8, no. 5, pp. 1414-1418, May, 2021.
- [24] N. K. Sharma, M. Vemula, and V. Tadiboyina, "An Experimental Study of Heart Disease Prediction Using Different Supervised Machine Learning Algorithms", *International Journal of Engineering Research and Technology*, vol. 14, no. 3, pp. 227-240, 2021.
- [25] H. A. G. Elsayed and L. Syed, "An automatic early risk classification of hard coronary heart diseases using framingham scoring model," in *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing*. ACM, pp. 1-8, Mar 2017. doi: <https://doi.org/10.1145/3018896.3036384>
- [26] P. Jonsson and C. Wohlin, "An evaluation of k-nearest neighbour imputation using likertdata," *10th International Symposium on Software Metrics*, 2004. *Proceedings.IEEE*, pp. 108-118, 2004. doi: 10.1109/METRIC.2004.1357895.
- [27] G. Aksu, C. O. Güzeller, and M.T Eser, "The Effect of the Normalization Method Used in Different Sample Sizes on the Success of Artificial Neural Network Model", *International Journal of Assessment Tools in Education*, vol. 6, no. 2, , pp. 170-192, 2019. doi: <https://doi.org/10.21449/ijate.479404>.
- [28] S. Prasad, "Some notes on z- scores and t- scores," *International Journal of scientific research and management (IJSRM)*, vol. 3, no. 4, pp. 2608-2610, 2015.
- [29] K. Potdar, T. S. pardawala, and C. D. pai, "A comparative study of categorical variable encoding techniques for neural network classifiers," *International Journal of Computer Applications*, vol. 175 ,no.4, pp. 7-9, Oct, 2017.
- [30] H. LIU, F. Hussain, C. L.TAN, and M. DASH, "Discretization: An Enabling Technique," *Data Mining and Knowledge Discovery*, vol.6, pp. 393-423, 2002. doi: <https://doi.org/10.1023/A:1016304305535>.
- [31] D. Ramesh, P. Suraj, and L. Saini, "Big data analytics in healthcare: A survey approach," in *2016 International Conference on Microelectronics, Computing and Communications (Mi-croCom)*, pp. 1-6, Jan, 2016.
- [32] C. B. C. Latha and S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," *Informatics in Medicine Unlocked*, vol. 16, p.100203, 2019. doi: <https://doi.org/10.1016/j.imu.2019.100203>.
- [33] A. H. Khaleel, G. A. Al-Suhail, and B. M. Hussan, "A weighted voting of k-nearest neighbor algorithm for diabetes mellitus," *International Journal of Computer Science and Mobile Computing*, vol. 6, no. 1, pp. 43-51, 2017.
- [34] N. Ali, D. Neagu, and P. Trundle, "Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets," *SN Applied Sciences*, vol 1, no. 12, Nov, 2019.
- [35] Y. Zhang, Z. Lin, Y. Kang, R. Ning, and Y. Meng, "A feed-forward neural network model for the accurate prediction of diabetes mellitus," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol.7, no.8, pp. 151-155, Aug, 2018.
- [36] H. Yan, et al., "A multilayer perceptron-based medical decision support system for heart disease diagnosis," *Expert Systems with Applications*, vol.30, no.2, pp. 272 - 281, 2006.
- [37] A. Olgac and B. Karlik, "Performance analysis of various activation

- functions in generalized mlp architectures of neural networks," *International Journal of Artificial Intelligence And Expert Systems*, vol. 1, pp. 111–122, Feb, 2011.
- [38] A. Smith, F. Gu, and A.D. Ball, "An Approach to Reducing Input Parameter Volume for Fault Classifiers," *International Journal of Automation and Computing*, vol. 16, no. 2, pp. 199-212, Apr, 2019. doi: 10.1007/s11633-018-1162-7.
- [39] P. RefaeiNzadeh, L. Tang, and H. Liu, *Cross-Validation*. Boston, MA: Springer US, pp.532–538, 1927-2010, 2009. doi: https://doi.org/10.1007/978-0-387-39940-9_565.
- [40] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Vol. 2, pp. 1137–1143, Aug, 1995.
- [41] A. Kori, "Comparative study of data classifiers using rapidminer," *International Journal of Engineering Development and Research*, vol. 5, pp. 1041–1043, 2017.
- [42] W. Chen, S. Chen, H. Zhang, and T. Wu, "A hybrid prediction model for type 2 diabetes using k-means and decision tree," in *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pp. 386–390, Nov 2017. doi: 10.1109/ICSESS.2017.8342938.
- [43] M.Hossin and N.Sulaiman, "A review on evaluation metrics for data classification evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, pp. 01–11, Mar 2015.

Numerical Investigation on System of Ordinary Differential Equations Absolute Time Inference with Mathematica®

Adeniji Adejimi¹, Surulere Samuel², Mkolesia Andrew³, Shatalov Michael⁴
Department of Mathematics and Statistics
Tshwane University of Technology
Pretoria, South Africa

Abstract—The purpose of this research is to perform a comparative numerical analysis of an efficient numerical methods for second-order ordinary differential equations, by reducing the second-order ODE to a system of first-order differential equations. Then we obtain approximate solutions to the system of ODE. To validate the accuracy of the algorithm, a comparison between Euler's method and the Runge-Kutta method of order four was carried out and an exact solution was found to verify the efficiency, accuracy of the methods. Graphical representations of the parametric plots were also presented. Time inference analysis is taken to check the time taken to executes the algorithm in Mathematica®12.2.0. The obtained approximate solution using the algorithm shows that the Runge-Kutta method of order four is more efficient for solving system of linear ordinary differential equations.

Keywords—Euler's method; Runge-Kutta method; System of ODE; Mathematica®; AbsoluteTiming

I. INTRODUCTION

In recent times, computer algebraic system (CAS) have been employed to investigate the use of built-in functions and construction of algorithm to obtain solutions to initial value problems. Scientific problems arises in stem fields as Biology, Physics, Chemistry ,and Engineering [1], Ecology with respect to inverse problems [2]. They also arise in non-stem fields as humanities, virtual art, designs and films [3], [4]. Differential equations recently have been understood, that its application model life realities and numerical methods are used to investigate ODEs. Some problems have been modeled to explore dynamics of music, literature, poetry, prey-predator models, decay radiation, and numerical methods [5], have been used to understand its dynamics in terms of approximation. Some mathematical real life situations modeled by system of ODEs do not have exact solutions, hence, approximation and/or numerical techniques must be used. In this paper, we consider second-order ODEs, convert them to a system of ODEs and compare the exact and numerical solution. This exploration is investigated through evaluation using Mathematica®[6], [7], [8]. Euler's and Runge-Kutta [9], [5], [10] algorithm are implemented with its built-in function. Other researchers have implemented several methods to solve initial value problems and the analysis of the accuracy in areas such as Epidemiology [11], stability and efficiency in system of ODEs [12] but not with the algorithm implemented in this paper and the use of time inference obtained with Mathematica®. The Euler's method has large errors which

is illustrated using Taylor's series expansion. Taylor's series converges within a small range and as a result, if the step size is not relatively small, it diverges. From Taylor's expansion, the first two terms represent the Euler's method. Through this, its imperative enough to know if h is not small enough, the method is not accurate and will not be a good use for practical implementation. The Runge-Kutta method of order four gives a better approximation than the Euler's method but its complicated in computing. However, it converges faster to the exact solution. This procedure explore the comparative analysis between the exact and numerical solution of the ODEs [13],[14]. The algorithm works robustly for obtaining solution to system of first order ODEs [15] and comparing the solution by computation and investigating the absolute timing it takes for each algorithm to compute, hence the algorithm can be applied to biological models. This algorithm will be applied to investigate nonlinear system of ODEs for further studies.

In this paper, we are more concerned about the accuracy of the algorithm and its reliability after comparing the exact solution, Euler's method and Runge-Kutta method on the system of first-order linear ODEs.

II. METHODOLOGY

In this section, we will consider a second-order ODEs by recasting it to system of first-order ODEs. In considering the Euler's method and Runge-Kutta method of order four, this methods investigated by coding an algorithm using the CAS, Mathematica and its built-in function. The approximate solution of both methods would be obtained and compared by the result of algorithm and the built-in function of Mathematica®. For Equation (1) and Equation (2), the explicit formula for the solution can't be obtained for an initial value problem of such form but can approximate the solution using the numerical methods which is based on the tangent line approximation. In this paper, the Euler's method would be used to approximate the system of ODEs with initial conditions. This technique will investigate numerically with the aid of the Mathematica® software.

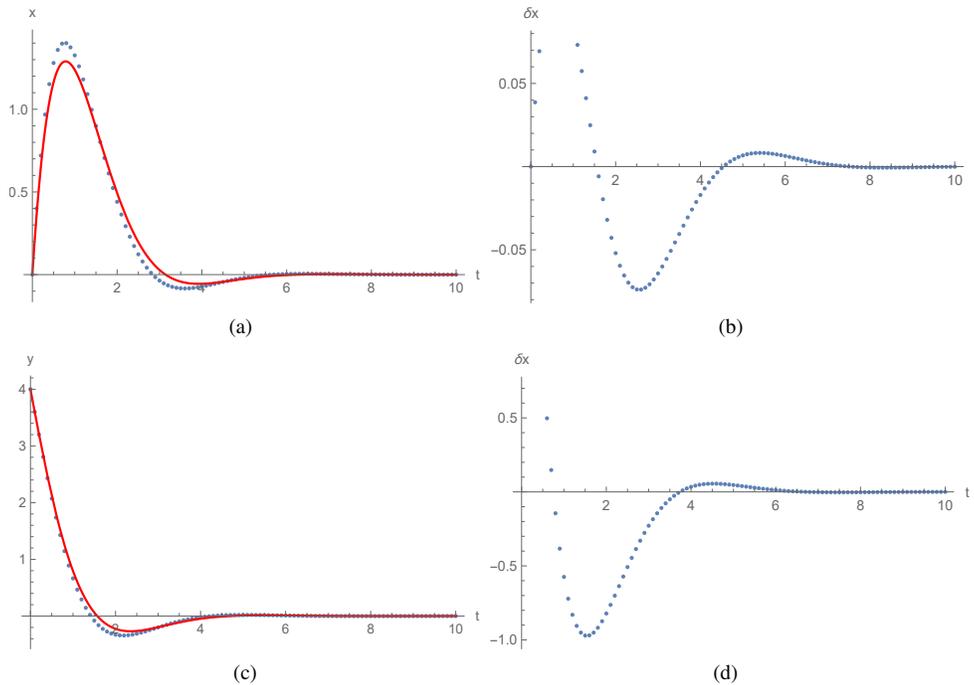


Fig. 1. (a) Graphical Representation: Exact and Numerical Solution of $x(t)$ (b) Error Plot of $x(t)$ (c) Graphical Representation: Exact and Numerical Solution of $y(t)$ (d) Error Plot of $y(t)$, using the Euler's Algorithm for Example 1.

III. EULER'S METHOD

Consider the initial value problem in

$$\frac{dx}{dt} = f(t, x, y), \quad (1)$$

$$\frac{dy}{dt} = g(t, x, y), \quad (2)$$

with initial condition $x(0) = 0$, and $y(0) = 0$ valid for $a \leq t \leq b$

Let

$$m_k = f(x_k, y_k), \quad (3)$$

$$n_k = g(x_k, y_k), \quad (4)$$

Such that

$$x_{k+1} = x_k + m_k \Delta t \Rightarrow x_{k+1} = x_k + f(x_k, y_k) \Delta t, \quad (5)$$

$$y_{k+1} = y_k + n_k \Delta t \Rightarrow y_{k+1} = y_k + g(x_k, y_k) \Delta t. \quad (6)$$

where $\Delta t = \frac{t_n - t_0}{n}$ is known as the step size. The rule of the thumb to the above algorithm is the smaller the step size (over the increased number of length interval), the more accurate the approximate solution. However, even when extremely small step sizes are used, over a large number of steps the error starts to accumulate and the estimate diverges from the actual functional value which denotes its limitation and the requires more time to compute the approximate solution using the CAS.

IV. RUNGE-KUTTA METHOD OF ORDER 4(RK4)

The Runge-Kutta method are an important family of methods for approximate solutions of ODEs which were developed by mathematics duo C Runge (1856-1927) and M.W Kutta (1867-1944). In this paper, the Runge-Kutta method (RK4)

was considered for a initial value problem (IVP), for a system of the first order ODEs, such that

$$x'(t) = f(t, x) \text{ where } x = x(t) = [(x_1(t), x_2(t), \dots, x_n(t))]^T, \quad (7)$$

$$f \in [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

$$y'(t) = g(t, y) \text{ where } y = y(t) = [(y_1(t), y_2(t), \dots, y_n(t))]^T, \quad (8)$$

$$g \in [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

with initial condition $x(0) = x_0$. To implement the RK4 method of solution of $x(t)$ and $y(t)$ of the IVP over the time interval $t \in [a, b]$. The time interval was subdivided into n equal intervals and we selected the step size such that $h = \frac{(b-a)}{n}$ where h is called the step size.

Consider a problem for the implementation of RK4

$$x' = f(t, x, y); x(t_0) = x_0, \quad (9)$$

$$y' = g(t, x, y); y(t_0) = y_0. \quad (10)$$

To obtain the solution through investigation of built-in algorithm of Mathematica[®] 12.2.0, it imperative to note that k 's obtains the solution for $x(t)$ and l 's for solutions to $y(t)$.

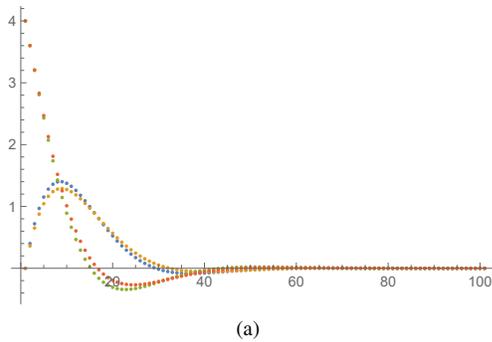


Fig. 2. (a) Graph of Exact and Numerical Solution, Euler's Algorithm for Example 1.

V. NUMERICAL EXAMPLES

A. Using Euler's Algorithm

Consider the system of ODE

$$x' = -x + y, \quad (11)$$

$$y' = -x - y, \quad (12)$$

$$x(0) = 0, y(0) = 4, h = 0.1.$$

The exact solution was obtained as $x(t) = 4e^{-t} \sin(t)$, $y(t) = 4e^{-t} \cos(t)$. Results generated using the Euler's algorithm were represented on the Table I

Let $x_1 = 4e^{-t} \sin(t)$, $y_1 = 4e^{-t} \cos(t)$ be the exact solution, and $X_n \Rightarrow x' = -x + y$, $Y_n \Rightarrow y' = -x - y$ be the numerical solution.

TABLE I. RESULTS OBTAINED USING EULER'S ALGORITHM

n	x1	X _n	y1	Y _n
0	0	0	4	4
1	0.361332	0.4	3.60127	3.6
2	0.650627	0.72	3.20964	3.2
3	0.875707	0.968	2.83092	2.808
4	1.04414	1.152	2.46962	2.4304
5	1.16315	1.27984	2.12912	2.07216
6	1.23953	1.359072	1.81182	1.7369
7	1.27964	1.3968608	1.51924	1.42736
8	1.28932	1.3999104	1.2522	1.14494
9	1.27391	1.374412864	1.01091	0.89045
10	1.23824	1.3260166272	0.795064	0.663964

The graphs in Figure 1 represents the exact solution, numerical solution of $x(t)$ & $y(t)$ and the error plots.

Figure 1 (b) and (d) shows that the order of errors observed in the graphs for Figure 1 (a) and (c) are 10^{-2} and 10^0 respectively. For example 1, we observed that the solutions for $x(t)$ yielded better approximations than the solutions for $y(t)$. Although, from a graphical perspective, Figure 1 (c) shows better agreement (between the exact and numerical solutions) as compared to Figure 1

The graph in Figure 2 represents the exact solution and numerical solution of $x(t)$ & $y(t)$

B. Using Runge-Kutta Algorithm

Now, let us consider the system of ODEs from equations in Equation (11)-Equation (12). The exact solution was obtained as $x(t) = 4e^{-t} \sin(t)$, $y(t) = 4e^{-t} \cos(t)$. The results generated by the Runge-Kutta algorithm are represented in the

Table II.

Let $L = 4e^{-t} \sin(t)$, $P = 4e^{-t} \cos(t)$ be the exact solution, and $X_n \Rightarrow x' = -x + y$, $Y_n \Rightarrow y' = -x - y$ be the numerical solution

TABLE II. TABLE II REPRESENT THE RESULT OBTAINED USING RUNGE-KUTTA ALGORITHM.

n	L	X _n	P	Y _n
0	0	0	4	4
1	0.361332	0.318733	3.60127	3.6
2	0.650627	0.579032	3.20964	3.21121
3	0.875707	0.78666	2.83092	2.8378
4	1.04414	0.94723	2.46962	2.48297
5	1.16315	1.06615	2.12912	2.14912
6	1.23953	1.14855	1.81182	1.83792
7	1.27964	1.19929	1.51924	1.55039
8	1.28932	1.22289	1.2522	1.28704
9	1.27391	1.22354	1.01091	1.04789
10	1.23824	1.20507	0.795064	0.832592

The graph in Figure 3 represents the exact solution and numerical solution of $x(t)$ & $y(t)$ and error plots.

The same phenomenon we observed under the Euler's algorithm section is also observed in this section. Figure 3 (b) and (d) are error plots having order of magnitudes 10^{-1} and 10^{-1} respectively. However, Figure 3 (a) and (c) shows that the agreement between the exact and numerical solutions for $y(t)$ is better compared to the agreements between the exact and numerical solutions for $x(t)$.

The graph in Figure 4 represents the exact solution and numerical solution of $x(t)$

C. Example 2

Consider the IVP of the form

$$x'' + 2x' + x = 0, \quad (13)$$

$$x(0) = 1, x'(0) = 0.$$

In what follows, we reduce Equation (13) to a system of first-order ODEs

$$x' = y, \quad (14)$$

$$y' = -x - 2y \quad (15)$$

$$x(0) = 1, y(0) = 0$$

The exact solution of the system was obtained as

$$x(t) = e^{-t}(1 + t), \quad (16)$$

$$y(t) = -e^{-t} \cdot t, \quad (17)$$

$$x(0) = 1, y(0) = 0,$$

$$\text{Analytical solution } y(t) = e^{-t}(1 + t). \quad (18)$$

The result generated using Euler's algorithm by comparative analysis is presented in Table III.

We define the following

$$\left. \begin{array}{l} x_2 \rightarrow e^{-t}(1 + t), \\ y_2 \rightarrow -e^{-t} \cdot t. \end{array} \right\} = \text{Exact solution} \quad (19)$$

and

$$\left. \begin{array}{l} X_n \rightarrow y, \\ Y_n \rightarrow -x - 2y. \end{array} \right\} = \text{Numerical solution} \quad (20)$$

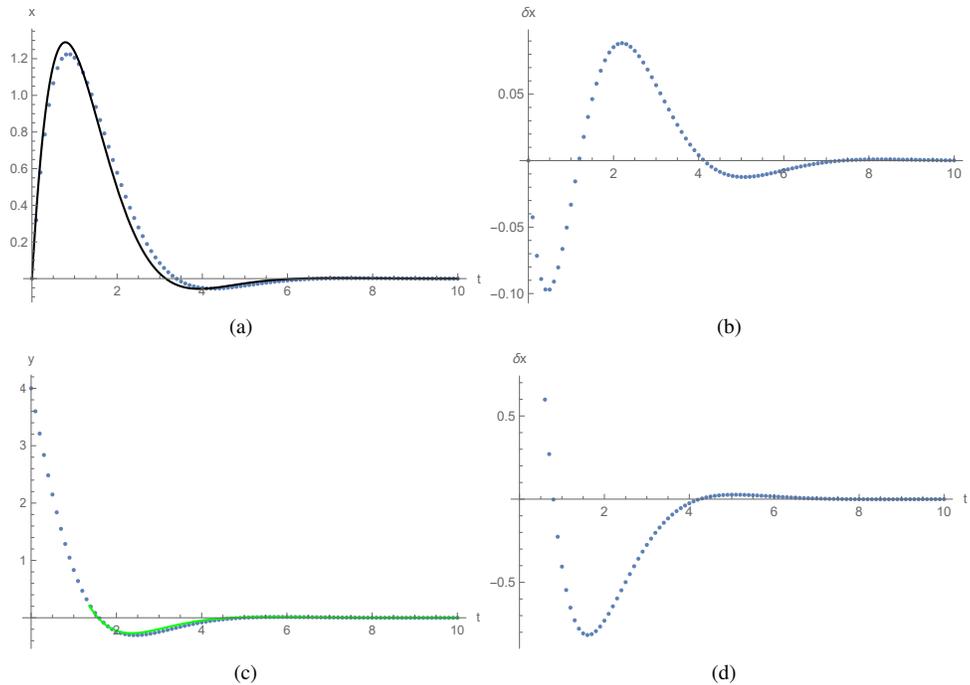


Fig. 3. (a) Graphical Representation: Exact and Numerical Solution of $x(t)$ (b) Error Plot of $x(t)$ (c) Graphical Representation: Exact and Numerical Solution of $y(t)$ (d) Error Plot of $y(t)$, using the Runge-Kutta Algorithm for Example 1.

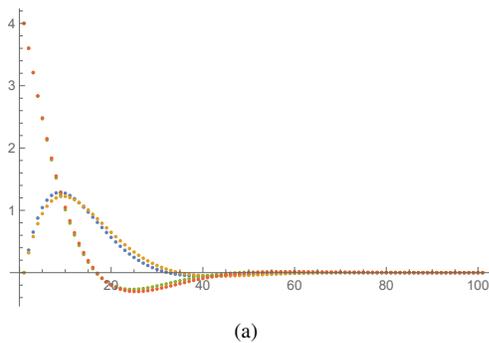


Fig. 4. (a) Graph of Exact and Numerical Solution, Runge-Kutta Algorithm for Example 1.

TABLE III. RESULTS OBTAINED USING EULER’S ALGORITHM

n	x_2	X_n	y_2	Y_n
0	1	1	0	0
1	0.995321	1	-0.0904837	0
2	0.982477	0.99	-0.163746	-0.1
3	0.963064	0.972	-0.222245	-0.18
4	0.938448	0.9477	-0.268128	-0.243
5	0.909796	0.885735	-0.303265	-0.2916
6	0.878099	0.850306	-0.329287	-0.32805
7	0.844195	0.813105	-0.34761	-0.354294
8	0.808792	0.774841	-0.359463	-0.372009
9	0.772482	0.736099	-0.365913	-0.382638
10	0.735759	0.697357	-0.367879	-0.38742

The graph in Figure 5 represents the exact solution, numerical solution of $x(t)$ & $y(t)$ and error plots.

For example 2, we can observe a stark contrast to example 1. Figure 5 (a) and (c) shows that the agreement between the exact and numerical solutions for $x(t)$ is better compared to the agreements between the exact and numerical solutions for

$y(t)$. The error plots for the graph in Figure 5 (b) shows that the approximation is quite good, as the error has order of magnitude 10^{-2} . The error plots for the solutions of $y(t)$ is not quite good (it has dimensions of order 10^0).

The graph in Figure 6 represents the exact solution and numerical solution of $x(t)$

D. Using Runge-Kutta Algorithm

Consider the system of ODEs in Equation (14)-Equation (15), result generated by RK4. The numerical solution is presented in Equation (20). The comparative Runge-Kutta algorithm is presented in the Table IV

TABLE IV. RESULTS OBTAINED USING RUNGE-KUTTA ALGORITHM

n	L	X_n	P	Y_n
0	1	1	0	0
1	0.995321	0.996738	-0.0904837	-0.0905542
2	0.982477	0.984877	-0.163746	-0.163981
3	0.963064	0.966075	-0.222245	-0.222686
4	0.938448	0.941753	-0.268128	-0.268776
5	0.909796	0.913129	-0.303265	-0.304097
6	0.878099	0.88124	-0.329287	-0.330261
7	0.844195	0.846968	-0.34761	-0.348673
8	0.808792	0.811058	-0.359463	-0.36056
9	0.772482	0.774135	-0.365913	-0.366986
10	0.735759	0.736721	-0.367879	-0.368873

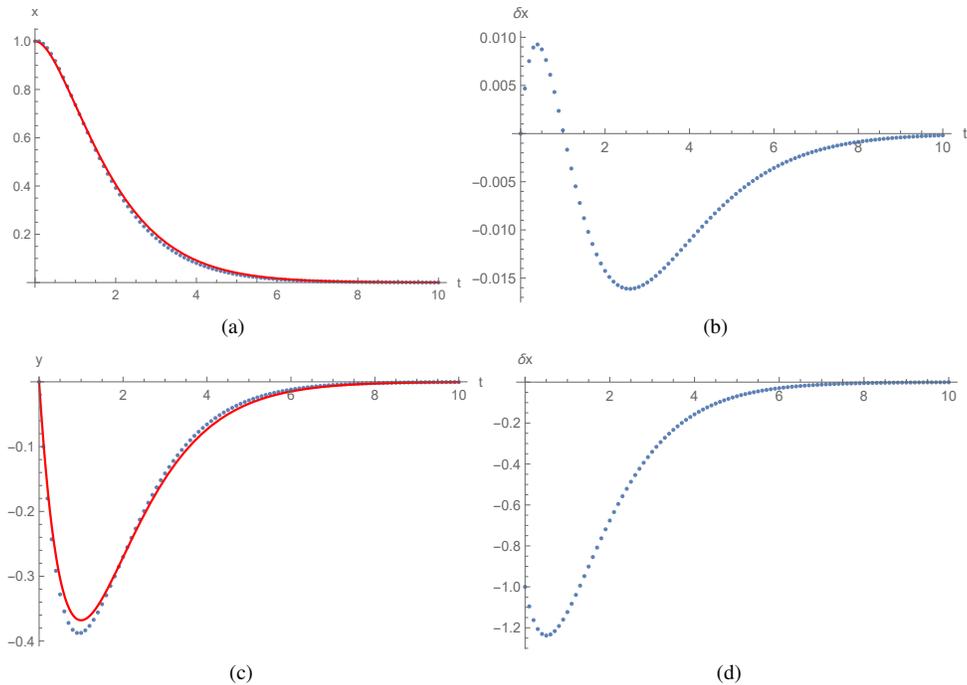


Fig. 5. (a) Graphical Representation: Exact and Numerical Solution of $x(t)$ (b) Error Plot of $x(t)$ (c) Graphical Representation: Exact and Numerical Solution of $y(t)$ (d) Error Plot of $y(t)$ using the Euler's Algorithm for Example 2.

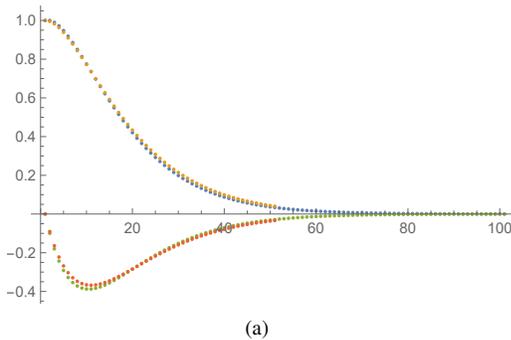


Fig. 6. (a) Graph of Exact and Numerical Solution, Euler's Algorithm for Example 2.

The graph in Figure 7 represents the exact solution, numerical solution of $x(t)$ & $y(t)$, and error plots.

For this section, the error plots for $x(t)$ has order of magnitude 10^{-3} (see Figure 7 (b)). The magnitude of order for the error plots of $y(t)$ still maintains the same value as it was in the section for Euler's algorithm (see Figure 7 (d)).

The graph in Figure 8 represents the exact solution and numerical solution of $x(t)$

E. Example 3

In what follows, we will consider the well-known model of a vibrating particle attached to a stationary wall through a spring

$$mx''(t) + dx'(t) + kx(t) = 0, \quad (21)$$

where m is the mass of the particle, d is the damping factor, and k is the stiffness constant of the spring. Equation (21) is

rewritten as

$$x''(t) + 2\delta x'(t) + \omega^2 x(t) = 0, \quad (22)$$

where $2\delta = d/m$ and $\omega^2 = k/m$. Equation (22) is reduced to the system of first-order ODEs

$$x' = y \quad (23)$$

$$y' = -2\delta y - \omega^2 x \quad (24)$$

$$x(0) = 0.1, y(0) = 0, \delta = 0, \omega = 2\pi$$

From Equation (23)-Equation (24), the exact solution was obtained as

$$x(t) = (0.1 - 2.6963 \times 10^{-34i} \cos[6.28319t] + (3.15475 \times 10^{-17} + 8.22996 \times 10^{-34i} \sin[6.28319t]), \quad (25)$$

$$y(t) = (-0.628319 + 2.38256^{-18i}) ((0 + 0.i) + (1 + 0.i) \sin[6.28319t]), \quad (26)$$

$$x(0) = 1, y(0) = 0.$$

The results generated from the Euler's algorithm is presented in Table V by defining the following: $x3$ to be Equation (25) and $y3$ represented as Equation (26), while the numerical solution of Equation (23) and Equation (24) are represented as (X_n, Y_n) (which were obtained using the Euler's algorithm).

TABLE V. RESULTS OBTAINED USING EULER'S ALGORITHM.

n	x3	X_n	y3	Y_n
0	$0.1 - 2.6963 \times 10^{-34i}$	0.1	$0. + 0i$	0
1	$0.0998027 - 2.63931 \times 10^{-34i}$	0.0998027	$-0.0394524 + 1.49602 \times 10^{-19i}$	-0.0394784
2	$0.0992115 - 2.57189 \times 10^{-34i}$	0.0996052	$-0.0787492 + 2.98614 \times 10^{-19i}$	-0.0789568
3	$0.0982287 - 2.49433 \times 10^{-34i}$	0.0988156	$-0.117735 + 4.46447 \times 10^{-19i}$	-0.118279
4	$0.0968583 - 2.40692 \times 10^{-34i}$	0.0976329	$-0.156256 + 5.92518 \times 10^{-19i}$	-0.15729
5	$0.0951057 - 2.31002 \times 10^{-34i}$	0.09606	$-0.194161 + 7.36251 \times 10^{-19i}$	-0.195834
6	$0.0929776 - 2.20399 \times 10^{-34i}$	0.0941016	$-0.231299 + 8.77078 \times 10^{-19i}$	-0.233757
7	$0.0904827 - 2.08927 \times 10^{-34i}$	0.091764	$-0.267525 + 1.01444 \times 10^{-18i}$	-0.270907
8	$0.0876307 - 1.96631 \times 10^{-34i}$	0.089055	$-0.302695 + 1.14781 \times 10^{-18i}$	-0.307134
9	$0.0844328 - 1.83558 \times 10^{-34i}$	0.0859836	$-0.33667 + 1.27664 \times 10^{-18i}$	-0.342291
10	$0.0809017 - 1.69761 \times 10^{-34i}$	0.0825607	$-0.369316 + 1.40043 \times 10^{-18i}$	-0.376236

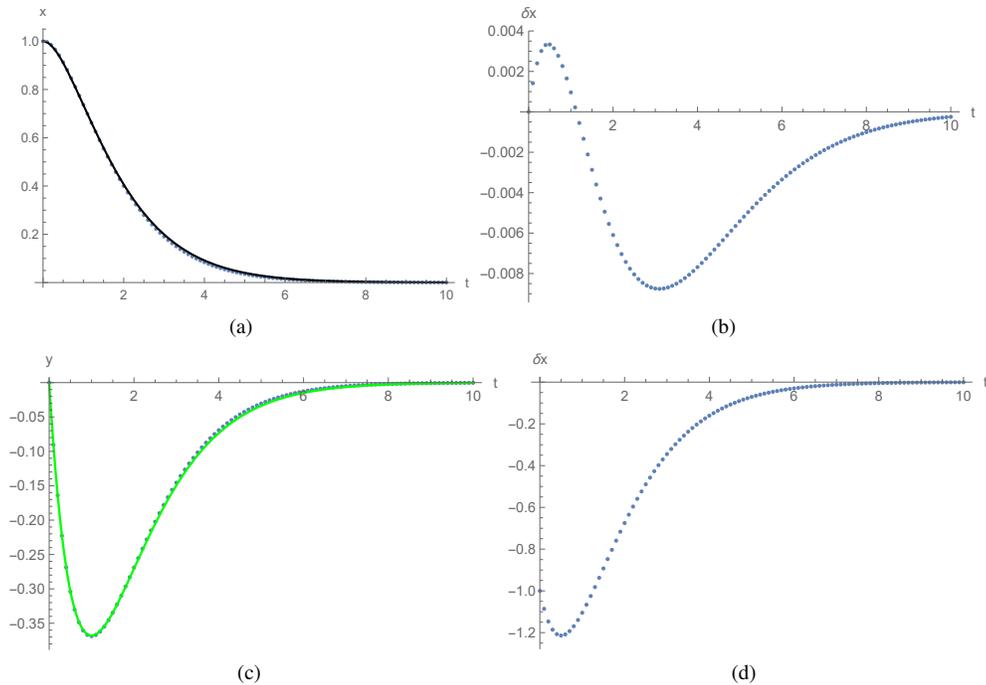


Fig. 7. (a) Graphical Representation: Exact and Numerical Solution of $x(t)$ (b) Error Plot of $x(t)$ (c) Graphical Representation: Exact and Numerical Solution of $y(t)$ (d) Error Plot of $y(t)$, using the Runge-Kutta Algorithm for Example 2.

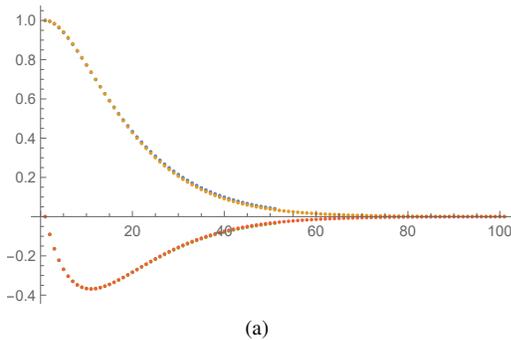


Fig. 8. (a) Graph of Exact and Numerical Solution, Runge-Kutta Algorithm for Example 2.

The graph in Figure 9 represents the exact solution, numerical solution of $x(t)$ & $y(t)$ and error plots. The graph in Figure 11 represents the exact solution and numerical solution of $x(t)$

F. Using Runge-Kutta Algorithm

Now, let us consider the system of ODEs in Equation (23)-Equation (24), for the results generated using RK4 and its numerical solution. The comparative analysis of RK4 algorithm is presented in the Table VI.

The graph in Figure 10 represents the exact solution, numerical solution of $x(t)$ & $y(t)$, and error plots.

The graph in Figure 12 represents the exact solution and numerical solution of $x(t)$

TABLE VI. RESULTS OBTAINED USING RUNGE-KUTTA ALGORITHM

n	L	X_n	P	Y_n
0	$0.1 - 2.6963 \times 10^{-34}i$	0.1	$0 + 0i$	0
1	$0.0998027 - 2.63931 \times 10^{-34}i$	0.1	$-0.0394524 + 1.49602 \times 10^{-19}i$	-0.0394784
2	$0.0992115 - 2.57189 \times 10^{-34}i$	0.0996052	$-0.0787492 + 2.98614 \times 10^{-19}i$	-0.0789568
3	$0.0982287 - 2.49433 \times 10^{-34}i$	0.0988156	$-0.117735 + 4.46447 \times 10^{-19}i$	-0.118279
4	$0.0968583 - 2.40692 \times 10^{-34}i$	0.0976329	$-0.156256 + 5.92518 \times 10^{-19}i$	-0.15729
5	$0.0951057 - 2.31002 \times 10^{-34}i$	0.09606	$-0.194161 + 7.36251 \times 10^{-19}i$	-0.195834
6	$0.0929776 - 2.20399 \times 10^{-34}i$	0.0941016	$-0.231299 + 8.77078 \times 10^{-19}i$	-0.233757
7	$0.0904827 - 2.08927 \times 10^{-34}i$	0.091764	$-0.267525 + 1.01444 \times 10^{-18}i$	-0.270907
8	$0.0876307 - 1.96631 \times 10^{-34}i$	0.089055	$-0.302695 + 1.14781 \times 10^{-18}i$	-0.307134
9	$0.0844328 - 1.83558 \times 10^{-34}i$	0.0859836	$-0.33667 + 1.27664 \times 10^{-18}i$	-0.342291
10	$0.0809017 - 1.69761 \times 10^{-34}i$	0.0825607	$-0.369316 + 1.40043 \times 10^{-18}i$	-0.376236

VI. TIME INFERENCE

This section details the time taken by Mathematica® to evaluate the formulated algorithms. An inference can be drawn that Euler’s algorithm solves the system of ODEs in less time than the RK4. By computation the average time for each algorithm is represented in Table VII. As seen in Table VII, the Euler’s algorithm computes faster although, it does not give an accurate approximation but the RK4 algorithm gives a better approximation despite having a lower computational speed compared to the Euler’s algorithm.

TABLE VII. THE COMPUTATIONAL TIME TAKEN IN EVALUATING THE EULER’S AND RUNGE-KUTTA ALGORITHM

	Example 1		Example 2	
	Euler’s	RK4	Euler’s	RK4
Absolute Timing	0.000553	0.014443	0.001603	0.003870
Average Timing	0.000732	0.002925	0.000684	0.002868

VII. CONCLUSION

In this paper, we considered an algorithm which was simulated using Mathematica® 12.2.0 installed on a 512 SSD, 64 bit laptop. The built-in function of Mathematica® with

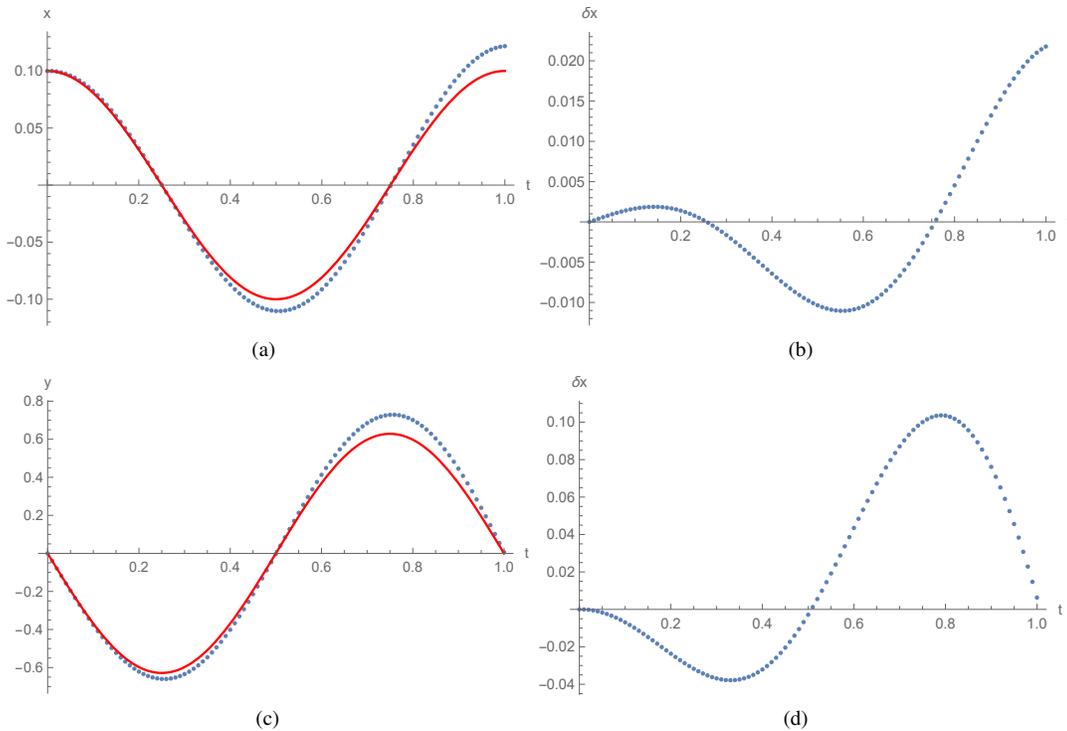


Fig. 9. (a) Graphical Representation: Exact and Numerical Solution of $x(t)$ (b) Error Plot of $x(t)$ (c) Graphical Representation: Exact and Numerical Solution of $y(t)$ (d) Error Plot of $y(t)$ using the Euler's Algorithm in Example 3.

absolute timing was used to generate a better approximation in the three examples considered in preceding sections. The accuracy of this algorithm depends on the system of ODEs under consideration. The algorithm was built to simulate system of ODEs by converting second-order ODE to a system of first-order IVP, with graphical representation of exact and numerical solution. The error plot with the same step size $h = 0.1$ was also graphically illustrated. Considering the parametric plot of example 1, the plot shows that the solution presented by the both algorithm, that RK4 algorithm gives a better approximation than the Euler's algorithm. This was also the case in example 2, the Runge-Kutta algorithm showed that the solution obtained from the exact and numerical solution by comparative analysis obtained accurate result and the same goes for example 3. It has been observed by the results shown in Figure 13, through the parametric plot that the algorithm works well. Table VII shows how the Euler's algorithm computes faster but does not give accurate approximations, like that of RK4 which computes slower but gives better approximation. In future work, we intend to harness the algorithm to solve higher orders of ODE that can be recasted to system of ODEs. We also aim to adopt the algorithms to solve other differential equations such as Lotka-Volterra model, and stiff equations, and investigate the algorithm time taken in execution, using Mathematica®. By investigation, it is shown in Figure 9 that the parametric plot of the system of ODEs using the (Euler's and RK4) algorithms concludes that RK4 algorithm works well and gives better approximation. In Equation (25) and Equation (26), we observed the appearance of complex conjugate values in the exact solution. This was due to the low value of damping ($\delta = 0.01$) used in the equation of motion, describing the particle attached to a stationary wall. This explanation

also covers for the complex values observed in Table V and Table VI. The numerical values in the aforementioned tables showed close approximations, despite the complex part. The values of the complex parts are also negligible as they are really small (having dimensions of 10^{-34} , 10^{-19} and 10^{-18}). Comparing Figure 9 (b) and (d) against Figure 10 (b) and (d), we can observe that the Runge-Kutta algorithm also gave a better approximation to the Euler algorithm. The errors for the Runge-Kutta algorithm was in order 5×10^{-2} while Euler algorithm was in order 1×10^{-1} .

ACKNOWLEDGMENT

The authors wish to thank Tshwane University of Technology for their support and the Department of Higher Education and Training, South Africa.

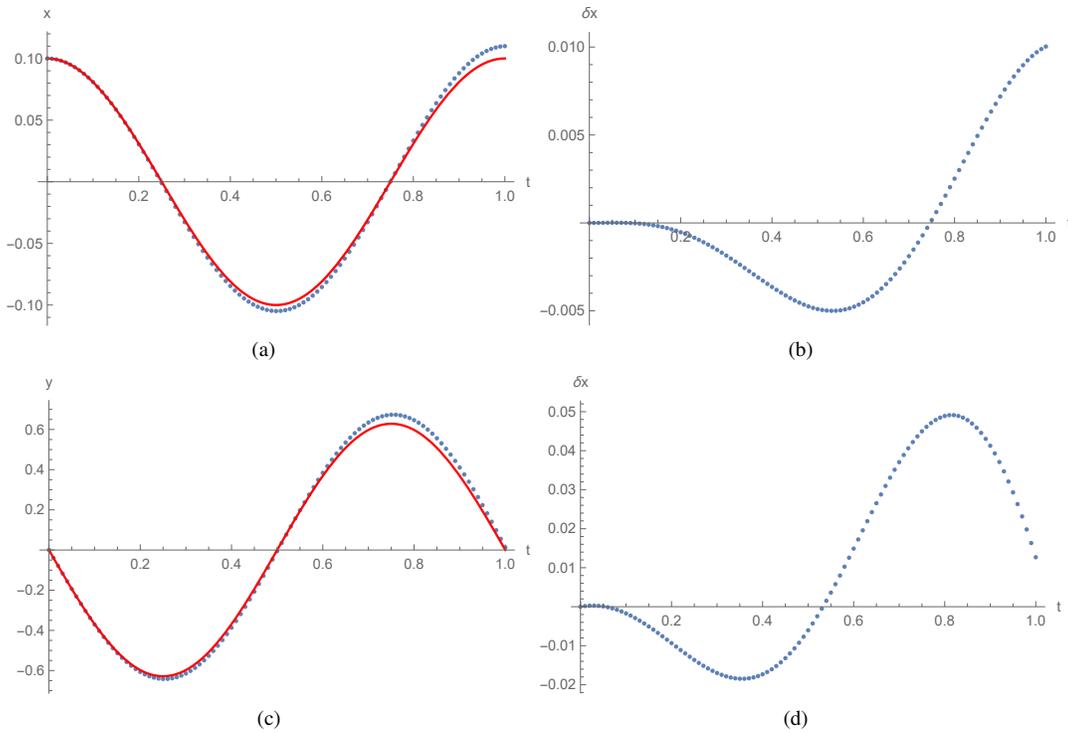


Fig. 10. (a) Graphical Representation: Exact and Numerical Solution of $x(t)$ (b) Error Plot of $x(t)$ (c) Graphical Representation: Exact and Numerical Solution of $y(t)$ (d) Error Plot of $y(t)$ using the Runge-Kutta Algorithm in Example 3.

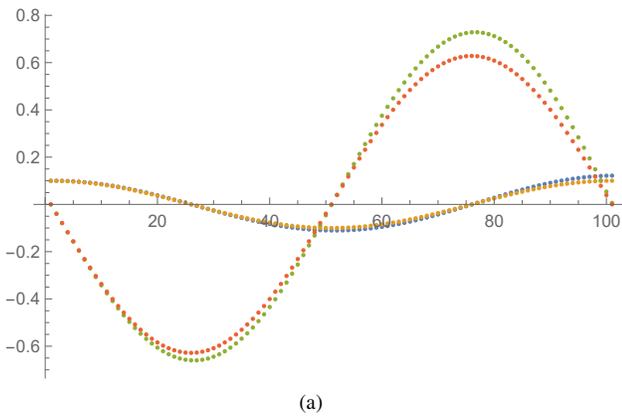


Fig. 11. (a) Graph of Exact and Numerical Solution, Euler's Algorithm in Example 3.

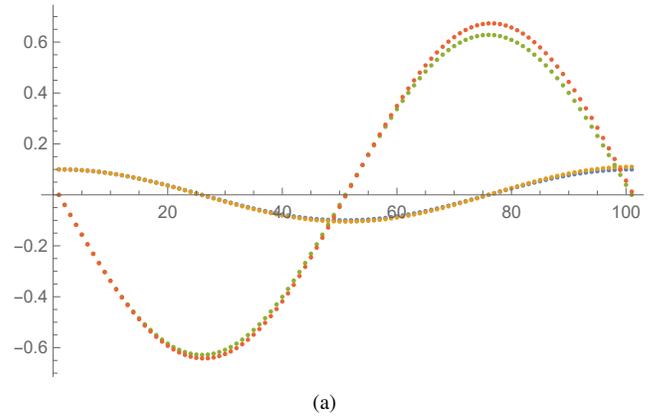


Fig. 12. (a) Graph of Exact and Numerical Solution, Runge-Kutta Algorithm in Example 3.

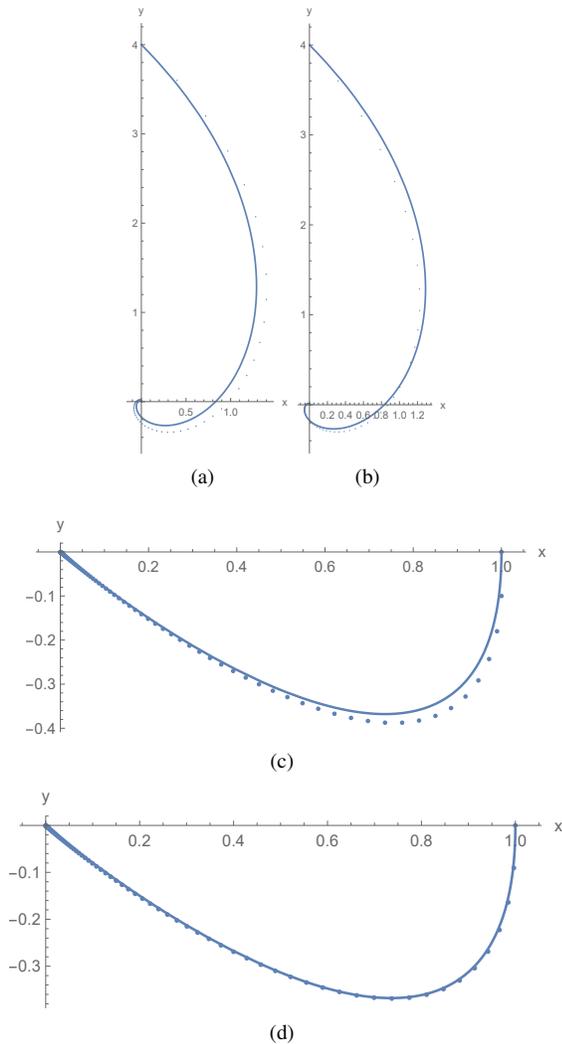


Fig. 13. [a,b] Parametric plot: Euler's and RK4 algorithm for Example 1
[c,d] Parametric plot: Euler's and RK4 algorithm for Example 2

REFERENCES

- [1] F. C. Hoppensteadt and C. S. Peskin., Mathematics in Medicine and the Life Sciences. *Springer-Verlag, New-York,* 1992.
- [2] AA. Adeniji, and MY. Shatalov, and I. Fedotov, and AC. Mkolesia, Introduction to numerical methods. *Discontinuity Nonlinearity and Complexity*, 10(3), 523-534, 2021
- [3] L. Koss. Visual arts, design, and differential equations. *Journal of Mathematics and the Arts*, 11(3): 129–158, 2017.
- [4] W. Brian. 2015-Koss-Differential Equations In Art and Film. *SIMIODE*.
- [5] J. RChasnov. Introduction to numerical methods. *The Hong Kong University of Science and Technology*, 60, 2012.
- [6] Ü. Göktaş, D. Kapadia. Methods in Mathematica for solving ordinary differential equations. *Mathematical and Computational Applications*, 16(4): 784–796, 2011.
- [7] Abell, LL. Martha and Braselton, P. James. Differential equations with Mathematica. *Academic Press*, 2016.
- [8] M. Sofroniou, and R. Knapp. Wolfram Mathematica Tutorial Collection-Advanced Numerical Differential Equation Solving in Mathematica. *Wolfram Research. Inc*, 2008.
- [9] DF. Griffiths, DJ. Higham. Euler's method, Numerical Methods for Ordinary Differential Equations. *Springer, London*, 19-31, 2010.
- [10] CR. Nathan, W. Dorathy Euler's Method for Systems of Differential Equations. *Applications of Calculus to Biology and Medicine*, 117-122, 2017.
- [11] B. Yong The comparison of fourth order Runge-Kutta and homotopy analysis method for solving three basic epidemic models. *Journal of Physics: Conference Series*, 1317(1), 12-20, 2019.
- [12] AT. Kolar Comparison of numerical methods for solving a system of ordinary differential equations: accuracy, stability and efficiency. *Mälardalen University, School of Education, Culture and Communication.*, 65, 2020.
- [13] Blanchard, Paul and Devaney, Robert L and Glen, R Hall, and Jong-Eao Lee. *Differential Equations: A Contemporary Approach*, Thomson Learning, 2007.
- [14] N. Shawagfeh, D. Kaya. Comparing numerical methods for the solutions of systems of ordinary differential equations. *Applied mathematics letters*, 17(3):323-8, 2004.
- [15] P, B Conrad. Ordinary Differential Equations: A Systems Approach. , 2010.

Maneuverable Autonomy of a Six-legged Walking Robot: Design and Implementation using Deep Neural Networks and Hexapod Locomotion

Hiep Xuan Huynh¹, Nghia Duong-Trung², Tran Nam Quoc Nguyen³, Bao Hoai Le⁴, Tam Hung Le⁵
Can Tho University, Can Tho city, Vietnam^{1,3,4,5}
Technische Universität Berlin, Berlin, Germany²

Abstract—Automatically real-time synthesizing behaviors for a six-legged walking robot pose several exciting challenges, which can be categorized into mechanics design, control software, and the combination of both. Due to the complexity of controlling and automation, numerous studies choose to gear their attention to a specific aspect of the whole challenge by either proposing valid and low-power assumption of mechanical parts or implementing software solutions upon sensorial capabilities and camera. Therefore, a complete solution associating both mechanical moving parts, hardware components, and software encouraging generalization should be adequately addressed. The architecture proposed in this article orchestrates (i) interlocutor face detection and recognition utilizing ensemble learning and convolutional neural networks, (ii) maneuverable automation of six-legged robot via hexapod locomotion, and (iii) deployment on a Raspberry Pi, that has not been previously reported in the literature. Not satisfying there, the authors even develop one step further by enabling real-time operation. We believe that our contributions ignite multi-research disciplines ranging from IoT, computer vision, machine learning, and robot autonomy.

Keywords—Six-legged walking robot; hexapod locomotion; Raspberry Pi; deep neural networks

I. INTRODUCTION AND MOTIVATION

Controlling a multi-legged robot receives a vast research endeavor in the literature. At first, researchers scrutinized the advantages of a legged robot over a wheeled one by investigating the degrees of freedom, mechanical structure, gait generation, and body stabilization on difficult terrain [1]. Then a scheme of the locomotion control approach by mimicking the biological movement of legs and basic motion pattern was investigated. The simulation and modeling were developed on either the eight-legged or six-legged robot. The feedback taken from the environment was collected from motor sensors [2]. By mimicking the physical structure of legged animals, engineers can implement a more stable and faster walking robot. A six-legged walking robot whose kinetic chain of legs based on the biomechanics of the cockroach was prototyped by [3]. The motion mechanisms were controlled by a PLC controller, and the two-degree-of-freedom robot only moved in a straight direction. Another work researched dynamic forward legged locomotion toward a static object using a single monolithic controller. Two virtual quadrupedal robots were used [4]. Another two papers have done intensive experiments and modeling of the robot stability and energy consumption analysis [5], [6]. Many other research articles have strengthened the aspect of sensor-based feedback, and mechanics can be found in [7], [8], [9].

Computer vision on Raspberry Pi is overwhelming on its own field [10], [11], [12]. The tasks can be narrowed to object detection and recognition, home automation, and surveillance, combining camera and software libraries [13], [14]. The idea of real-time face detection deployed on a telepresence robot was proposed in [15], where two Raspberry Pi boards were utilized to achieve real-time detection. They developed a technique to enable real-time detection on a Raspberry Pi and using the result to control the pan and tilt unit of a telepresence robot. The user can either robot's movement by turning or rotating the smartphone or set the robot to follow the face of someone who is at the robot place.

This article aims to bring the two separate research directions into one choir where robot automation, IoT, computer vision, and machine learning operate several complex tasks in real-time. Many contributions of owner work are follows. First, we propose a prototype and mechanical structure of a six-legged working robot. Second, we propose an architecture of convolutional neural networks that work very effectively regarding highly accurate face recognition and be able to deploy on a Raspberry Pi. Third, we implement a complete robot automation solution.

The article is organized as follows. Section (II) briefly presents the background on Haar-like features, boosting technique in machine learning, the cascade of weak classifiers to form a strong one, and convolutional neural networks. These materials help readers understand the underlying techniques that address the tasks of face detection and recognition. Section (III) summarizes several main hardware components used in the experimental robot. Many software libraries, e.g., especially related to convolutional neural networks, are introduced in Section (IV). Our main contributions of this research article are presented in Section (V), where the authors describe the system architecture and its three sub modules, e.g., face detection, owner recognition, and kinematic automation. Section (VI) present the experiments and experimental remarks. And Section (VII) concludes our work.

II. BACKGROUND AND STATE OF THE ART

A. Haar-like Features

Object detection utilizing Haar feature-based cascade classifiers is an effective object detection method [16]. It is a machine-learning-based model where a cascade function is trained from a lot of positive and negative observations. In

our experiments, images of and images without faces are the positive and negative observations, respectively.

Concretely, a general image makes up of pixel in which a person or a computer vision algorithm can recognize many distinctive shapes or patterns, partly depending on the level of contraction. During the creation of a Haar cascade, various parts of the image are cropped and scaled so that we consider only a few pixels that we call a window at a time. We will subtract some of the grayscale pixel values from others to measure the window's similarity to specific common shapes where a dark region meets a light area. If a window is very similar to one of these archetypes, it can be selected as a feature at similar positions and magnifications relative to each other. We expect to find similar features across all images of the same investigated subject.

The problem of determining the face of a person in an image is the problem of binary classification: face or not. To identify faces with Haar features, we apply Haar's features throughout the image, areas that are considered to be most similar to Haar features will be marked. Haar-like featurings is the most basic method of facial recognition. It will add a Haar feature to the whole image. The area that looks like it will identify it as the face. There will be many areas in the image where the algorithm recognizes a face. However, we do not just use a Haar feature but use a lot of features in different image positions and sizes to exclude these areas.

B. Adaptive Boosting

In machine learning [17], [18], we might take random guessing as a default baseline for evaluation reference to other learning models. A learning model is weak when it is slightly better than random guessing, while the strong one accurately classifies elements most of the time. In practice, it may not be competent to entirely rely upon the performance of just one machine learning model. Ensemble learning [19], [20] offers a systematic solution to combine the predictive power of multiple learners referring to learning a weighted combination of base models [21], [22]. The resultant model gives the aggregated output from several models. The aggregation models could be either from the same learning algorithm or different learning algorithms. The objective is to produce a classifier called C , where the output is the prediction confidence. We denote N as the number of observations. Hence, the total error E of classifier C is defined as follows:

$$E(C) = \sum_{i=1}^N \exp(-c_i C(\mathbf{x}_i)) , \quad (1)$$

where $c_i = \pm 1$ for the classification of observation \mathbf{x}_i . Note that we consider binary classification for a simplified discussion. In Equation (1), c_i and $C(\mathbf{x}_i)$ have the same sign if C highly confidently assign sample \mathbf{x}_i with large absolute value of $C(\mathbf{x}_i)$. Consequently, its exponential error $\exp(-c_i C(\mathbf{x}_i))$ contribute very little to the total error $E(C)$ and visa versa.

AdaBoost, short for Adaptive Boosting, is widely used as an ensemble learner. It focuses on classification problems and aims to convert a set of weak classifiers into a strong one. It uses in conjunction with many other types of

learning algorithms to improve performance. Conceptually, AdaBoost iteratively generates a series of classifiers C_m . In each iteration, the classifier is improved by concentrating on the misclassified observations. Therefore, each classifier is more accurate than its predecessor. It is generated as a linear combination of weak classifiers, and their coefficient gives the confidence of these weak classifiers.

Let initialize C_1 to the weak classifier $f_1(\mathbf{x})$ which misclassifies the least training samples. We denote β_1 as its coefficient, and the value is chosen concerning minimize the following error:

$$\begin{aligned} E(C_1) &= \sum_{i=1}^N \exp(-c_i \beta_1 f_1(\mathbf{x}_i)) \\ &= \sum_{c_i \neq f_1(\mathbf{x}_i)} \exp(\beta_1) + \sum_{c_i = f_1(\mathbf{x}_i)} \exp(-\beta_1) , \end{aligned} \quad (2)$$

where c_i and $f_1(\mathbf{x}_i)$ obtain the values of $+1$ or -1 . Then, the difference with respect to β_1 is calculated as follows:

$$\frac{dE(C_1)}{d\beta_1} = \sum_{c_i \neq f_1(\mathbf{x}_i)} \exp(\beta_1) - \sum_{c_i = f_1(\mathbf{x}_i)} \exp(-\beta_1) . \quad (3)$$

We denote N_A as the number of observations that have been accurately classified. Hence, we have

$$\frac{dE}{d\beta_1} = (N - N_A) \exp(\beta_1) - N_A \exp(-\beta_1) . \quad (4)$$

By setting Equation (4) to zero, we solve for β_1 to find the extreme as follows:

$$\begin{aligned} \exp(2\beta_1) &= \frac{N_A}{N - N_A} \\ \beta_1 &= \frac{1}{2} \log \frac{N_A}{N - N_A} . \end{aligned} \quad (5)$$

As previously mentioned, the weak learners should be better than random guessing. Consequently, N_A cannot be less than $N/2$, and the more considerable value of N_A is, the more confidence it indicates. The total error through m iterations is calculated as follows:

$$\begin{aligned} E(C_m) &= \sum_{i=1}^N \exp(-c_i C_{m-1}(\mathbf{x}_i)) - c_i \beta_m f_m(\mathbf{x}_i) \\ &= \sum_{i=1}^N \exp(-c_i C_{m-1}(\mathbf{x}_i)) \exp(-c_i \beta_m f_m(\mathbf{x}_i)) \\ &= \exp(-\beta_m) \sum_{c_i = f_m(\mathbf{x}_i)} v_{i,m} \\ &\quad + \exp(\beta_m) \sum_{c_i \neq f_m(\mathbf{x}_i)} v_{i,m} , \end{aligned} \quad (6)$$

where $v_{i,m}$ is the weights.

Weight is the exponential error that the current weak classifier produces when predicting observation \mathbf{x}_i . Hence, Equation (6) can be split as follows:

$$E(C_m) = \exp(-\beta_m) \sum_{i=1}^N v_{i,m} + (\exp(\beta_m) - \exp(-\beta_m)) \sum_{c_i \neq f_m(\mathbf{x}_i)} v_{i,m} . \quad (7)$$

As we can see in Equation (7), the second summation depends on the classifier f_m . Consequently, the classifier which produces the sum of weights over the accurately classified observations is the largest should be selected.

Assuming that k_m is selected, we differentiate the total error concerning β_m and set the resultant difference to zero, solving for β_m . Then β_m is as follows:

$$\beta_m = \frac{1}{2} \log \frac{\sum_{c_i=f_m(\mathbf{x}_i)} v_{i,m}}{\sum_{c_i \neq f_m(\mathbf{x}_i)} v_{i,m}} . \quad (8)$$

The more accurate observations f_m classifies, the larger its weights are. The higher the weights are, the more confident it gets. The more assured it achieves, the larger β_m is produced.

C. Cascade of Classifiers

We will have a sequence of classifiers, in which each classifier is built using the Adaboost algorithm. Now, let all the windows go through this sequence of classifiers. The first classifier eliminates the most negative sub-windows and passes through the positive sub-window. Here, the classifier is very simple, and therefore, the computational complexity is also very low. Of course, because it is so simple, among the windows that are recognized as faces, there will be a large number of windows that are misidentified (not faces). The windows passed by the first classifier will be viewed as considered by the classifier later: if the classifier thinks it's not the face is removed, if the classifier thinks it's a face, then we pass through and move to the rear class. Later the classifier is more complex, demanding more calculation. People call sub-windows (templates) that the classifier doesn't remove as patterns that are hard to identify. The deeper these patterns get into the sequence of classifiers, the harder it is to identify. Only windows that pass through all the classifiers will decide that face.

D. Convolutional Neural Networks

The high-level general CNN architecture is illustrated in Fig. 1. Input layers are places that raw images are loaded into. The raw images are specified by their width, height, and number of channels. Typically, RGB values of each pixel form three channels. Convolutional layers apply a patch of locally connecting neurons to transform the input data from the previous layer. Consequently, the layer calculates a dot product between the area of the neurons in the input layer and the weights to which they are locally connected in the output

layer. Typically, a convolution is defined as a mathematical operation describing a rule for how to merge two sets of information. Within this work, the authors briefly present the CNN background that leads to the proposed architecture in Section V-B. Further details on CNN and its next-generation investigating transfer learning [23], [24], [25], [26], [27] leaves to interesting readers.

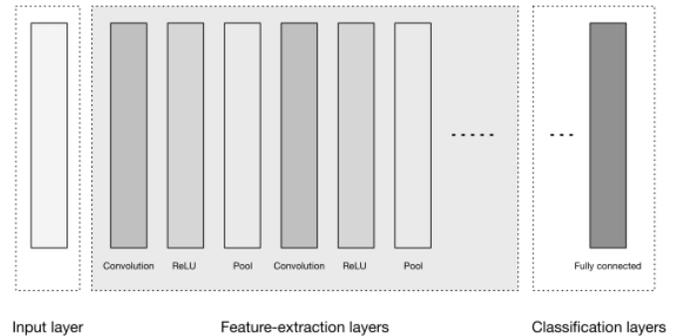


Fig. 1. High-level General CNN Architecture.

III. HARDWARE COMPONENTS

A. Raspberry Pi 3 Model B

The Raspberry Pi (RP) is a small, cheap, single-size, single-chip computer board that comes with CPU, GPU, USB ports, and I/O pins and is capable of performing several functions. Simple as a regular calculator [28], [29]. RP 3 Model B comes with a 64-bit quad-core processor on a circuit board with WiFi and Bluetooth and USB features. It has a processing speed of 700 MHz to 1.4 GHz in which the RAM ranges from 256 to 1GB. The device's CPU is considered to be the brain of the device responsible for executing instructions based on mathematical and logical operations. The board is equipped with Broadcom video core cable, which is mainly used to play video games through the device. RP 3 comes with GPIO (General Purpose Input Output) pins that are essential for maintaining a connection with other electronic devices. These input terminals receive commands and operate based on the device program. An Ethernet port is integrated on this device to establish a communication line with other devices. The board has four USB ports that can be used to connect a keyboard, mouse, or USB 3G connection to access the internet and an SD card added to store the operating system. The power connector is a basic part of the circuit board used to supply 5 V to the board. We can use any source to set the power for the circuit board. However, we prefer to connect the power cable via the laptop's USB port to provide 5 V. HDMI port is used to connect LCD or TV, can support cables versions 1.3 and 1.4. RCA Video port is used to connect the old TV screen using 3.5mm jack without supporting the HDMI port.

B. 32 Channel USB/UART Servo Motor Controller Driver Board

A 32 Servo control circuit can be used with the software on the computer, wireless controller PS2, Android app, Arduino (or other microcontrollers) through UART communication [30]. UART stands for Universal Asynchronous Receiver/-Transmitter. It is a physical circuit in a microcontroller, or

a stand-alone IC. A UART's main purpose is to transmit and receive serial data. The 32 Servo control circuit is capable of simultaneously controlling 32 servos smoothly, from which users can coordinate applications to control systems for robotic arms, industrial production lines, and machines. Servo Controller software and instruction can be found at RTROBOT¹.

Several important technical specifications of the board are summarized as follows. Physical size 63mm x 45mm. Operation voltage 5V. Servo Motor Input Voltage is 4.2V to 7.2V (According to a specific servo). CPU 32bit. Baud Rate (USB) 115200. Baud Rate (Bluetooth UART) 4800, 9600, 19200, 38400, 57600, and 115200. Flash Capacity 16M. Servo Motor Controller the Number Of Simultaneous 32. Max Action Groups 256. Control precision 1us. LEDs are dedicated to CPU power, servo motor power, and wireless remote control.

C. Camera Pi v1.3

Raspberry Pi camera has a built-in 5 Megapixel camera with high light sensitivity, able to shoot well in many different lighting conditions, both indoors and outdoors. A special feature that the camera brings is high-definition photography during movie recording. No additional USB ports are needed for the camera as the camera is securely attached to the CSI socket. This helps limit bandwidth bottlenecks for USB processing chips on Raspberry circuits. The camera cable length has been carefully calculated when it reached the required length while ensuring the speed of image transmission from the module to the RP. The important technical specifications of the Pi camera follow. Resolution 2592 x 1944 (5 megapixels). Lens $f = 3.57$ mm, $f / 2.8$. 65-degree viewing angle. Focus range 0.69m to infinity (at 1.38m). Support 1080p @ 30 fps with H.264 (AVC) codec, 720p @ 60fps and 640x480p @ 60/90 fps. CSI interface. Three dimensions are 25 mm x 25 mm x 10 mm. Weight about 2.8g.

D. Digital RC Servo LD-1501MG Motor

Digital RC Servo LD-1501MG motor is used in complex structured robots such as the humanoid robot, biped Robot, and spider robot due to the outstanding characteristics² compared to the traditional Analog RC Servo MG995 and MG996. RC servos or RC actuators convert electrical commands from the receiver or control system, back into physical movement. A servo plugs into a specific receiver, gyro, or FBL controller channel to move that specific part of the RC model. This movement is proportional, meaning that the servo will only move as much as the transmitter stick on your radio is moved, or as much as the gyro/FBL system instructs it to move. Digital RC Servo LD-1501MG motor has a metal gearbox, fast response speed with strong traction according to manufacturer torque figures up to 17Kg.cm. The important technical specifications of the motor are the following. The operation voltage is 5 to 7.4VDC. No-load current of 100mA, with load \dot{c} 500mA. Speed 0.16sec / 60° at 7VDC. Traction torque of 13Kg.cm at 6VDC; 15KG.cm at 6.5VDC; 17KG.cm at 7VDC. Weight 60g. Dimensions 40 x 20 x 40.5mm. Cable length 30cm. We use 18 motors in the implementation to achieve 4 degrees of freedom of the robot's body, which encourages us to implement complex movement and dancing.

¹<https://rtrobot.org/software/>

²<https://www.rchelicopterfun.com/rc-servos.html>

E. Power supply and Low voltage circuit board

The AC power cord (AC) 220V is connected to L and N pins at the input on the power supply (12V, 10A). The 12V output of power adapter, V+ pin, and V- pin, are connected to positive pin (+) and negative pin (-) on the low voltage circuit board. The low voltage circuit board consists of 2 outputs; one output can be adjusted voltage from 1.25 - 32V used to power the servo motor control module, the other output is a 5V USB port used to power the RP. In this case, the appropriate voltage to supply the servo motors is 6 - 6.5V. We can adjust by rheostat on the low voltage circuit board.

F. Mechanic legs

This robot model is inspired by real spiders. However, we can only partially reproduce the movements of spiders on this robot because of the limits of joints on the robot and the complexity of the control. The simple robot model will have fewer joints, so the number of servos (actuators) will be less beneficial, which can include benefits such as reducing the weight of the robot, reducing power consumption and increasing working time. The minimum condition for a robot to move is that each leg must be made up of two or more moving joints. The fact that our robot has three motion joints makes the robot move smoothly and motion control easier. The spider robot model has six legs, each of which consists of three servos (actuators) that are caught in the coal. Each side of the robot consists of three legs that allow the robot to move. The 3D sketch is done by Glovius CAD³.

IV. SOFTWARE LIBRARIES

The Tensorflow library [31] is an open-source library developed by Google and made available to the developer community in 2015. Google Brain has developed the platform for image search, voice recognition, traces data, and many Google services. Keras [32], [33] is an open-source library for neural networks written in the Python language. Keras is a high-level API that can be used in conjunction with deep learning libraries such as Tensorflow, Theano, and CNTK. In this article, we use the Keras library in combination with the Tensorflow library [34]. OpenCV library [35] is a leading open-source library for computer vision, image processing, machine learning, and GPU acceleration features in real-time operation. OpenCV is designed for efficient computing and with a heavy focus on real-time applications. Raspberry Pi uses multiple operating systems: Raspbian, Ubuntu, Windows 10 IoT, and Chromium OS. In this article, the system uses the Raspian operating system⁴ [36], [37] Raspian is the basic and most common operating system and is provided by the Raspberry Pi Foundation.

V. SYSTEM DESIGN

The system architecture can be disintegrated into three sub modules, e.g., face detection, owner recognition, and kinematic automation. The principal task in the face detection module is to detect faces from sequential pictures taken by the Raspberry camera. This task is discussed in Section V-A in more details. Then, only the cropped faces go through the process of owner

³<https://www.glovius.com/>

⁴<https://www.raspbian.org/>

recognition module where (i) the robot owners are recognized, and (ii) robot behaviors and kinematic locomotion are executed. We are going to describe these progress in Section V-B. The last module is more on the kinematics aspect, referred to Section V-C, where the authors discuss how to control the legs in case of owner detection and no owner detection. We have associated several techniques to achieve the harmonic maneuverability of detection, recognition and automation in real-time. The proposed system architecture is illustrated in Figure 2. Our hardware design and implementation of a six-legged walking robot are presented in Figure 3.

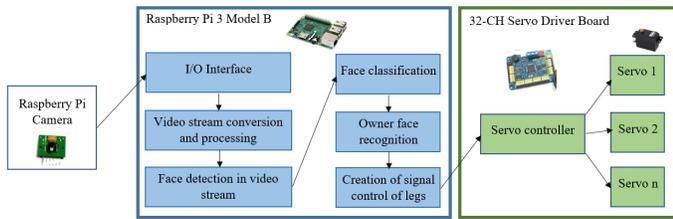


Fig. 2. The Overall Software and Hardware Architecture.

A. Module 1: Face Detection

The facial detection system is a computer application that automatically identifies a face and then identifies someone from a digital photo or a video resource. One way to do this is to compare pre-selected facial features from the image and a face database to remove non-face components. Building an automatic control program for spider robots using control signals as a result of the face recognition system with input data is frames cut from the video directly from the camera through the method of using Use the Haar Cascade technique. This RP analyzes the image and returns framed and stored faces for data training later. In general, the method is built on a combination of four components, e.g., Haar features, an integral image, Adaboost, and a cascade of classifiers. The progress of module 1 face detection is illustrated in Fig. 4.

After the faces are detected, they will be manually reclassified to verify faces, to eliminate false positives in the photo set. After the test is stored with the same number of three image files of three team members, these three files will be three data files used to train for the program in the next face recognition.

B. Module 2: Owner Recognition

In this study, the machine learning model used to classify faces is a CNN model, also known as a convolution neural network. The model has two parts: feature learning and classification. We briefly describe them in Section V-B1 and Section V-B2.

1) *Feature learning*: This section includes layers: Conv2D, MaxPooling2D, and Dropout. They extract the image features, learn the image features (during training), reduce the image size, and blur the image. The function of Conv2D is to extract image specifications, learn these specifications (during training). The Conv2D layers use filters to get image features (curves and lines).

The element retained during the training of extracting these properties is not the image channel. If all channels are stored, the data of the model after training will be very large. Besides, adding a new image and predicting will be no different from finding value in the database. So the values stored in this section are mostly kernels. The initial values of the kernels are randomly generated, or we can assign the initial values to them. The kernel matrix values will be updated during the Backpropagation process.

2) *Classification*: This part is two fully connected layers, which consists of 1 dense layer containing 512 nodes and one dense layer consisting of 3 nodes, calculating values and make a classification. Note that the output contains three nodes because we set the number of robot's owners is three. The values stored after training in this part are the weights that are updated during the backpropagation process. The sigmoid activation function is used in the last layer to predict the percentage of an object with an input sample. The class with the highest score will be the final prediction.

We proposed the CNN architecture presented in Fig. 5. We implemented the experimental model in Keras and TensorFlow. The model consists of four convolution Conv2D layers. The first, second, and fourth convolution layers accompany a Max-Pooling2D a Dropout layers. The third convolution layer is just attached to a MaxPooling2D layer. The last two layers are dense. Default Adam [38] is the optimization method. Categorical cross entropy [39], [40] is the loss measurement.

C. Module 3: Maneuverable Automation

A multi-legged robot [41], [42] maintains a tremendous potential for maneuverability over different terrain, particularly in comparison to wheeled robot. It possesses less complexity than a human-like walking robot regarding stability and development effort. A six-legged robot inspired by biology is one of the most common design [43], [44]. It is important to develop a good control mimicking the kinematic behavior of the complex six-legged robotic mechanism. The overall design of our proposed automation is presented in Fig. 7. We use the OpenCV library to process data received from the RP camera. The system selects one frame out of five received frames. The reason is that the robot takes a certain amount of time to complete the current motion, then we perform the identification and send the next control signal. During the experiment, we found that this ratio gives the most natural motion of the robot. When there is a frame, the system will detect the face cut, if any. These cropped faces are then processed by the machine learning model developed on the Keras and TensorFlow library. Next, the system controls the movements of the robot according to the predicted result. The system checks the current rotation angle of each servo motor and finds the appropriate pulse value for the action to be performed. The control signal will be transmitted to each motor, and the robot forms a motion.

We program the robot to perform some basic movements based on the results of the face recognition of the owner. The specific movements of the robot are as follows. If the owner is Hoai Bao, the robot performs the stand-up move. If the owner is Hung Tam, the robot performs the sit-down operation. If the owner is Nam Tran, the robot will move towards the owner. If

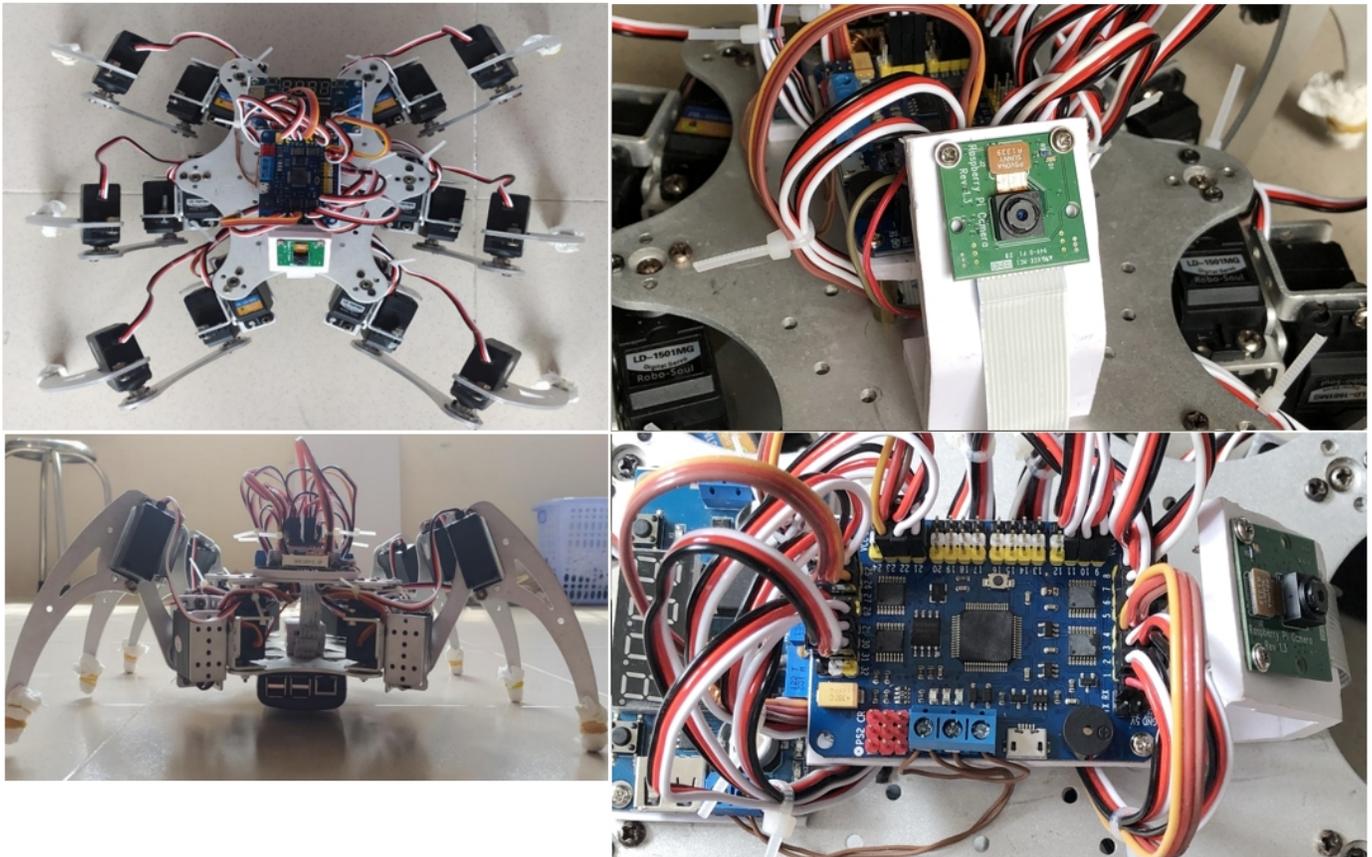


Fig. 3. Several Views of the Six-legged Walking Robot.

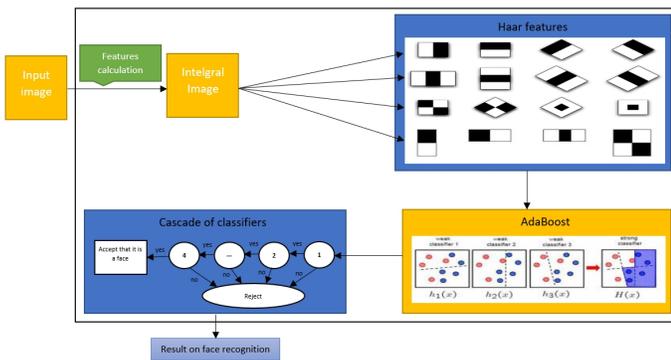


Fig. 4. The Architecture of Module 1: Face Detection.

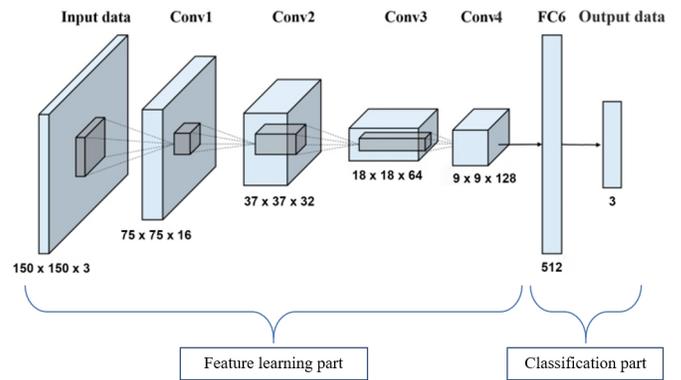


Fig. 5. The Architecture of Module 2: Face Recognition.

one of the three owners is not recognized, keep the previous movement. To control the rotation of the servo motor, we make a change in the pulse value. For LD-1510MG servo motors used in this study, the pulse range is from 500 to 2400 us. Fig. 6 depicts the pulse and sweep angle of the motor.

VI. EXPERIMENTS

A. Hardware and Software Setup

Our six-legged walking robot consists of several hardware components such as one RP 3 model B, one 32 channel US-B/UART servo motor controller driver board, one RP camera

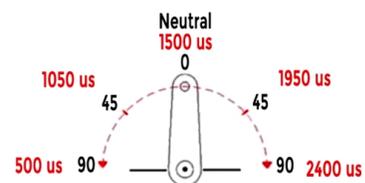


Fig. 6. Motion Angle of a Motor.

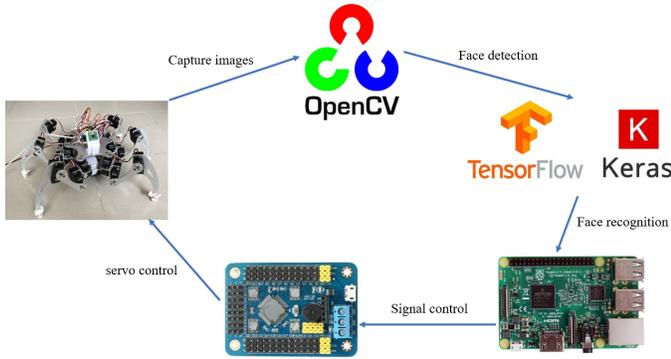


Fig. 7. The Architecture of Module 3: Maneuverable Automation.

v1.3, 18 digital RC servo LD-1501MG motors, one power supply 12V 10A, one low voltage circuit board, mechanic legs, and other components. The hardware connection is presented in Table I. The RP 3 model B is considered the robot brain where most of the controls are calculated from here.

TABLE I. CONNECTIONS OF MAIN HARDWARE DEVICES

Device	Connector Type	Device
Raspberry Pi	CSI - CSI	Pi Camera v.1.3
	RX - TX	32 Channel USB/UART Servo Motor Controller Driver Board
	TX - RX	
	GND - GND	
32 Channel USB/UART Servo Motor Controller Driver Board	GND - GND	Servos
	VCC - VCC	
	heart beat signals	

B. Legs Controller

Each servo motor is then connected to the control circuit. For ease of programming and maintenance, we specify the wiring order as follows: 6 front two-servo motors will be connected to pin 1 to 6 of the motor control module. The six servo motors of the middle two legs will be connected to pin 7 to 12. And the last six servo motors of the rear legs are connected to pin 13 to 18. To check whether the servo motors are correctly installed, we connect the servo motor control module to the computer via the USB interface and use the software⁵ to check the operation of each servo.

In the program, after having information on face recognition, this signal is the input for the control signal for the spider robot. The signals are transmitted except the Raspberry Pi to the servo control circuit by UART communication. Now we need to determine the pulse level for each leg of the robot is the servo. The balance joints smoothly work together, similar to the movements of spiders, in reality, to shape the movements of the spider robots. The control signals are transmitted in the form of pulses with an execution time of ms. Example #1P2500T100 where 1 is the pin position 1, P2500 means 2500 pulses, and T100 is the execution time of 100ms. The simulation of robot moving forward, backward and turning-left is presented in Fig. 8, 9, and 10.

Control angles are calculated manually using RTROBOT control simulation software. We check and select the appropriate pulse width to create motion for 18 robot leg joints.

The settings are saved into a configuration file containing foot control information and pulse value for each motor. Activities are grouped into blocks. The time taken for a moving group is fixed the same by 1 second to avoid control signal congestion. After detecting a face, the RP calls control signals that are pre-configured according to each face we have previously classified. The control signal is executed in series, but the execution speed of the foot joints is fast so that it will see simultaneous movements.

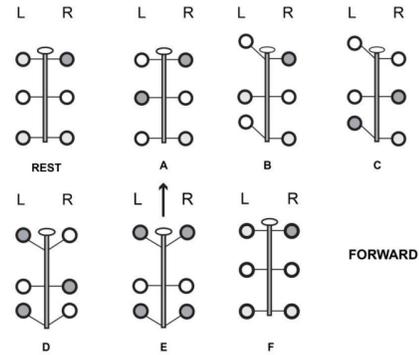


Fig. 8. Simulation of Robot Moving Forward.

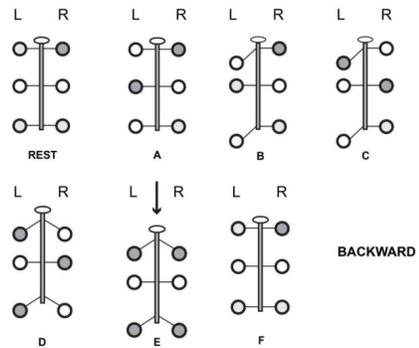


Fig. 9. Simulation of Robot Moving Backward.

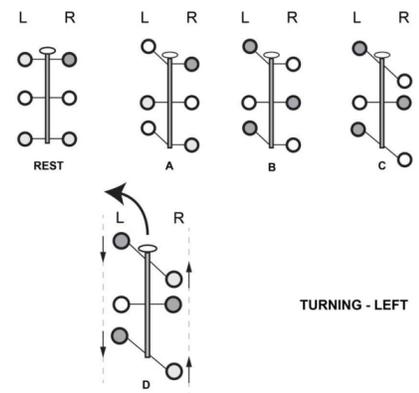


Fig. 10. Simulation of Robot Turning Left.

C. Face Detection and Image Collection

To apply the Haar recognition method with Adaboost, we use the *detectMultiScale* function to search for faces in the

⁵<https://rtrobot.org/software/>

image. Each image will be scanned in turn, and the program will confirm whether it is a human face or not and determine the number of faces in that frame by sliding the Haar symbols on the image, helping us identify All faces in the photo. The *detectMultiScale* function is a face search, which is part of the *CascadeClassifier* class. The codes are presented in Listing (1).

```
1 cascade = cv.CascadeClassifier(  
2   haarcascade_frontalface_default.xml)  
3 gray = convert_to_gray(imagePath)  
4 faces = cascade.detectMultiScale( gray ,  
5   scaleFactor = 1.1 ,  
6   scaleNeighbor =5, minSize(5,5) , maxSize  
7   (100,100))
```

Listing 1: Detect faces with detectMultiScale function.

The .xml file in Listing (1) contains the classifiers built on the standard database sets that is part of the OpenCV library.

The face detection on video is similar to the image because the video is a series of frames. Then we will capture each frame and save it. The program will read the captured frames again. Then we let the face detection program run on all image data and save all identifiable faces into another folder with the purpose of building data for the training set to be used in the recognition face. The codes in Listing (2) present how images are captured from Pi camera.

```
1 cap=cv.VideoCapture(0)  
2 i=0  
3 while (i!=15):  
4   frame=cap.read()  
5   frame=cv.flip(frame,-1)  
6   cv.imshow("Frame",frame)  
7   name= "%s" %i  
8   name= name + ".png"  
9   i=i+1  
10  cv.imwrite(name,frame)  
11  cap.release()  
12  cv.destroyAllWindows()
```

Listing 2: Images are captured using Pi camera.

Faces after being detected, the image will be manually reclassified to check again to identify faces correctly, removing false identities included in the image file. After the test is stored with the same number of three image files of three team members, these three files will be three data files used to train for the program in the next face recognition. Through experiments with model *haarcascade_frontalface_default.xml* with images containing human faces, the program recognized at a good level.

To increase the data for training, we use the solution augmentation technique to provide more data. *ImageDataGenerator* function is used in the implementation. The codes is presented in Listing (3).

```
1 image_gen = ImageDataGenerator(rescale  
2   =1./255, horizontal_flip = True,  
3   zom_range=0.25, rotation_range=45)  
4 train_data_gen = image_gen.  
5   flow_from_directory(batch_size=1000,  
6   directory="C:\\Test\\Image\\output\\",  
7   shuffle=True,
```

```
target_size=(IMG_HEIGHT, IMG_WIDTH))
```

Listing 3: ImageDataGenerator function for augmentation.

D. Face Recognition

1) *Datasets*: The experimental dataset is a set of images of the face of the robot's owners. This volume is divided into three parts, one part is used to train the machine learning program (training set), the other part is used for validation (validation set), and one part is used to check the accuracy. The original image's resolution taken by Pi Camera is 2048 x 1536. Then face regions are detected and cropped to 150 x 150.

The training set consists of 2250 photos divided into three folders. Each folder contains pictures of a specific person (Hoai Bao, Nam Tran, Hung Tam). These samples were taken from Pi cameras in environments with different light intensity and different angles. Besides, we also use augmentation methods such as: rotating the image with an angle of x degrees (with x being the number of degrees), enlarging the image, zooming the image, flipping the image, or combining many of them to get more photo patterns and to avoid overfitting. The validation set of 750 photos is also divided into three folders. Each folder contains pictures of a specific person (Hoai Bao, Nam Tran, Hung Tam). These images are also taken from the Pi camera. However, they are received at random. These samples will be reviewed after each epoch (number of model training) as one of the model's goodness test metrics. Similar to the validation set, the test set of 741 images is also divided into three folders corresponding to three people who are considered the owners of the robot.

2) *Evaluation Metrics*: Suppose we are solving a binary classification task with a labeled dataset $\mathcal{D} = \{\mathbf{x}_i, y_i\}$. Given a threshold parameter ϕ that guilds our decision rule $g(\mathbf{x})$ and count the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). We also define m_+ the total of condition positives, m_- the total of condition negatives, \hat{m}_+ the total predicted condition positives, \hat{m}_- the total predicted condition negatives, and m the total population.

We can compute the sensitivity, also known as true positive rate (TPR) or recall by using:

$$\text{TPR} = \frac{\text{TP}}{m_+} \approx P(\hat{y} = 1|y = 1) . \quad (9)$$

Similarly, we can compute the fall-out, also known as false positive rate or probability of false alarm by using:

$$\text{FPR} = \frac{\text{FP}}{m_-} \approx P(\hat{y} = 1|y = 0) . \quad (10)$$

The true negative rate (TNR) or specificity is defined as follows:

$$\text{TNR} = \frac{\text{TN}}{m_-} \approx P(\hat{y} = 0|y = 0) . \quad (11)$$

The false negative rate (FNR) or miss rate is calculated as follows:

$$\text{FNR} = \frac{\text{FN}}{m_+} \approx P(\hat{y} = 0 | y = 1) . \quad (12)$$

If we work with a dataset for binary classification when the number of negatives is very large or a dataset for multi-class text prediction when a class imbalance exists, considering TPR, FPR, TNR, and FNR themselves is not very informative. Before going further, we define positive predictive value (PPV) or precision as follows:

$$\text{PPV} = \frac{\text{TP}}{\hat{m}_+} \approx P(y = 1 | \hat{y} = 1) . \quad (13)$$

By combining Equation (9 and 13), we can compute F1-score as follows:

$$\text{F1-score} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} . \quad (14)$$

3) *Evaluation Results:* Model trained on a 2-core 4-thread laptop, 12GB RAM, Nvidia GTX 950M GPU. The size of each batch is 32 images. Model converges after 10 epochs. The model is then tested on test data to evaluate its final accuracy. The TP, FN, and FP over the test set are presented in Table II.

TABLE II. THE TP, FN, AND FP OVER THE TEST SET

No.	Owner face	TP	FN	FP
1	Nam Tran	247	16	8
2	Hung Tam	247	12	5
3	Hoai Bao	247	15	3
Total		741	43	16

From the evaluation result presented in Table II, we can easily calculate Precision \approx **0.98**, Recall \approx **0.95**, and F1-score \approx **0.96**.

4) *Experimental Remarks:* Model training can be done on Raspberry or regular computers. Because RP is an embedded computer, the hardware configuration and processing speed are relatively limited. Training on Raspberry takes a lot of time. In this case, the training time for one epoch is more than 2 hours (the model used in this topic is ten epoch training) for a training set of 2250 images, and each image size is 150x150. Besides, we have to turn off all the applications so that the training program has enough Ram to execute. However, we still achieved experimental results similar to those done on laptops. During experiments, we also try the VGG-16 model [45], which is very lightweight to compare the prediction accuracy with our proposed CNN model. However, the VGG-16 model cannot fit into RP's memory when loading images. Note that the image's resolution is only 150 x 150.

VII. CONCLUSIONS

Throughout the article, the authors present Haar-like features, ensemble learning, and CNN as the methodology background. We discuss hardware components and mechanical legs, which build a six-legged working robot. The critical contribution is the proposed system design, where three sub modules, e.g., face detection, face recognition, and kinematic

automation, work harmonically in real-time. Therefore, a complete solution associating both mechanical moving parts of a walking robot, IoT hardware components, and software has been adequately demonstrated. The proposed design and implementation has orchestrated several complex tasks. First, we address the task of interlocutor face detection and recognition, utilizing ensemble learning and convolutional neural networks. Second, we develop maneuverable automation of a six-legged robot via hexapod locomotion. And last but not least, we deploy the proposed system on a Raspberry Pi that has not been previously reported in the literature. The experiments are well described to encourage further reproducibility and improvement.

REFERENCES

- [1] E. Celaya and J. M. Porta, "A control structure for the locomotion of a legged robot on difficult terrain," *IEEE Robotics & Automation Magazine*, vol. 5, no. 2, pp. 43–51, 1998.
- [2] B. Klaassen, R. Linnemann, D. Spenneberg, and F. Kirchner, "Biomimetic walking robot scorpion: Control and modeling," *Robotics and autonomous systems*, vol. 41, no. 2-3, pp. 69–76, 2002.
- [3] C. Loughlin, S. Soyguder, and H. Alli, "Design and prototype of a six-legged walking insect robot," *Industrial Robot: An International Journal*, 2007.
- [4] J. Auerbach and J. C. Bongard, "How robot morphology and training order affect the learning of multiple behaviors," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 39–46.
- [5] S. S. Roy and D. K. Pratihari, "Dynamic modeling, stability and energy consumption analysis of a realistic six-legged walking robot," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 2, pp. 400–416, 2013.
- [6] —, "Kinematics, dynamics and power consumption analyses for turning motion of a six-legged robot," *Journal of Intelligent & Robotic Systems*, vol. 74, no. 3-4, pp. 663–688, 2014.
- [7] G. Zhong, H. Deng, G. Xin, and H. Wang, "Dynamic hybrid control of a hexapod walking robot: Experimental verification," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 5001–5011, 2016.
- [8] T. Maiti, Y. Ochi, D. Navarro, M. Miura-Mattausch, and H. Mattausch, "Walking robot movement on non-smooth surface controlled by pressure sensor," *Adv. Mater. Lett*, vol. 9, no. 2, pp. 123–127, 2018.
- [9] F. Tedeschi and G. Carbone, "Design of a novel leg-wheel hexapod walking robot," *Robotics*, vol. 6, no. 4, p. 40, 2017.
- [10] V. Patchava, H. B. Kandala, and P. R. Babu, "A smart home automation technique with raspberry pi using iot," in *2015 International conference on smart sensors and systems (IC-SSS)*. IEEE, 2015, pp. 1–4.
- [11] A. Pajankar, *Raspberry Pi computer vision programming*. Packt Publishing Ltd, 2015.
- [12] I. Iszaidy, R. Ngadiran, R. Ahmad, M. Jais, and D. Shuhaizar, "Implementation of raspberry pi for vehicle tracking and travel time information system: A survey," in *2016 International Conference on Robotics, Automation and Sciences (ICORAS)*. IEEE, 2016, pp. 1–4.
- [13] C. Kaymak and U. Aysegul, "Implementation of object detection and recognition algorithms on a robotic arm platform using raspberry pi," in *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*. IEEE, 2018, pp. 1–8.
- [14] J. Cicolani, *Beginning Robotics with Raspberry Pi and Arduino: Using Python and OpenCV*. Apress, 2018.
- [15] K. Janard and W. Marungrsith, "Accelerating real-time face detection on a raspberry pi telepresence robot," in *Fifth International Conference on the Innovative Computing Technology (INTECH 2015)*. IEEE, 2015, pp. 136–141.
- [16] A. Sharifara, M. S. M. Rahim, and Y. Anisi, "A general review of human face detection including a study of neural networks and haar feature-based cascade classifier in face detection," in *2014 International Symposium on Biometrics and Security Technologies (ISBAST)*. IEEE, 2014, pp. 73–78.

- [17] A. C. Faul, *A Concise Introduction to Machine Learning*. CRC Press, 2019.
- [18] N. Duong-Trung, *Social Media Learning: Novel Text Analytics for Geolocation and Topic Modeling*. Cuuillier Verlag, 2017.
- [19] L. Nanni, S. Brahnam, and A. Lumini, "Face detection ensemble with methods using depth information to filter false positives," *Sensors*, vol. 19, no. 23, p. 5242, 2019.
- [20] Z. Wu, W. Lin, and Y. Ji, "An integrated ensemble learning model for imbalanced fault diagnostics and prognostics," *IEEE Access*, vol. 6, pp. 8394–8402, 2018.
- [21] A. Saeed, A. Al-Hamadi, and A. Ghoneim, "Head pose estimation on top of haar-like face detection: A study using the kinect sensor," *Sensors*, vol. 15, no. 9, pp. 20945–20966, 2015.
- [22] S. W. Cho, N. R. Baek, M. C. Kim, J. H. Koo, J. H. Kim, and K. R. Park, "Face detection in nighttime images using visible-light camera sensors with two-step faster region-based convolutional neural network," *Sensors*, vol. 18, no. 9, p. 2995, 2018.
- [23] N. Duong-Trung, L.-D. Quach, and C.-N. Nguyen, "Towards Classification of Shrimp Diseases Using Transferred Convolutional Neural Networks," *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 4, pp. 724–732, 2020.
- [24] —, "Learning deep transferability for several agricultural classification problems," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 1, 2019.
- [25] N. Duong-Trung, L.-D. Quach, M.-H. Nguyen, and C.-N. Nguyen, "A combination of transfer learning and deep learning for medicinal plant classification," in *Proceedings of the 2019 4th International Conference on Intelligent Information Technology*, 2019, pp. 83–90.
- [26] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions," *IEEE Access*, vol. 5, pp. 14347–14357, 2017.
- [27] N. Duong-Trung, L.-D. Quach, M.-H. Nguyen, and C.-N. Nguyen, "Classification of grain discoloration via transfer learning and convolutional neural networks," in *Proceedings of the 3rd International Conference on Machine Learning and Soft Computing*, 2019, pp. 27–32.
- [28] S. Monk, *Raspberry Pi cookbook: Software and hardware problems and solutions*. " O'Reilly Media, Inc.", 2016.
- [29] D. Molloy, *Exploring raspberry PI*. Wiley Online Library, 2016.
- [30] "Basics of uart communication," Apr 2017. [Online]. Available: <https://www.circuitbasics.com/basics-uart-communication/>
- [31] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp. 265–283.
- [32] A. Gulli and S. Pal, *Deep learning with Keras*. Packt Publishing Ltd, 2017.
- [33] F. Chollet *et al.*, "Keras," <https://github.com/fchollet/keras>, 2015.
- [34] J. Brownlee, *Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras*. Machine Learning Mastery, 2016.
- [35] A. F. Villán, *Mastering OpenCV 4 with Python: a practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7*. Packt Publishing Ltd, 2019.
- [36] W. Harrington, *Learning raspbian*. Packt Publishing Ltd, 2015.
- [37] P. Singh, P. Nigam, P. Dewan, and A. Singh, "Design and implementation of a raspberry pi surveillance robot with pan tilt raspbian camera," *International Journal of Nanotechnology and Applications*, vol. 11, no. 1, pp. 69–73, 2017.
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [39] Z. Zhang and M. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," in *Advances in neural information processing systems*, 2018, pp. 8778–8788.
- [40] R. Gómez, "Understanding categorical cross-entropy loss, binary cross-entropy loss, softmax loss, logistic loss, focal loss and all those confusing names," *URL: https://gomburu.github.io/2018/05/23/cross_entropy_loss/*, 2018.
- [41] S. Peng, X. Ding, F. Yang, and K. Xu, "Motion planning and implementation for the self-recovery of an overturned multi-legged robot," *Robotica*, vol. 35, no. 5, pp. 1107–1120, 2017.
- [42] A. Mahapatra, S. S. Roy, and D. K. Pratihari, "Multi-legged robots—a review," in *Multi-body Dynamic Modeling of Multi-legged Robots*. Springer, 2020, pp. 11–32.
- [43] B. He, S. Xu, Y. Zhou, and Z. Wang, "Mobility properties analyses of a wall climbing hexapod robot," *Journal of Mechanical Science and Technology*, vol. 32, no. 3, pp. 1333–1344, 2018.
- [44] G. Zhong, L. Chen, Z. Jiao, J. Li, and H. Deng, "Locomotion control and gait planning of a novel hexapod robot using biomimetic neurons," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 2, pp. 624–636, 2017.
- [45] Y. T. Li and J. I. Guo, "A vgg-16 based faster rcnn model for pcb error inspection in industrial aoi applications," in *2018 IEEE international conference on consumer electronics-Taiwan (ICCE-TW)*. IEEE, 2018, pp. 1–2.

Comprehensive Analysis of Augmented Reality Technology in Modern Healthcare System

Jinat Ara¹
Jahangirnagar University
Bangladesh

Faria Benta Karim²
University of Asia Pacific
Bangladesh

Mohammed Saud A Alsubaie³
Taif University
Saudi Arabia

Yeasin Arafat Bhuiyan⁴
Dhaka University
Bangladesh

Muhammad Ismail Bhuiyan⁵
Sylhet MAG Osmani Medical College
Bangladesh

Salma Begum Bhyan⁶
University of the Sunshine coast
Australia

Hanif Bhuiyan^{7*}
Data61, CSIRO and CARRS-Q,
Queensland University of Technology, Australia

Abstract—The recent advances of Augmented Reality (AR) in healthcare have shown that technology is a significant part of the current healthcare system. In recent days, augmented reality has proposed numerous intelligent applications in the healthcare domain including, wearable access, telemedicine, remote surgery, diagnosis of medical reports, emergency medicine, etc. These developed augmented healthcare applications aim to improve patient care, increase efficiency, and decrease costs. Therefore, to identify the advances of AR-based healthcare applications, this article puts on an effort to perform an analysis of 45 peer-reviewed journal and conference articles from scholarly databases between 2011 and 2020. It also addresses concurrent concerns and their relevant future challenges including, user satisfaction, convenient prototypes, service availability, maintenance cost, etc. Despite the development of several AR healthcare applications, there are some untapped potentials regarding secure data transmission, which is an important factor for advancing this cutting-edge technology. Therefore, this paper also analyzes distinct AR security and privacy including, security requirements (i.e., scalability, confidentiality, integrity, resiliency, etc.) and attack terminologies (i.e. sniffing, fabrication, modification, interception, etc.). Based on the security issues, in this paper, we propose an artificial intelligence-based dynamic solution to build an intelligent security model to minimize data security risks. This intelligent model can identify seen and unseen threats in the threat detection layer and thus can protect data during data transmission. In addition, it prevents external attacks in the threat elimination layer using threat reduction mechanisms.

Keywords—Augmented Reality (AR); healthcare applications; healthcare challenges; AR-based healthcare security issues; dynamic security solution

I. INTRODUCTION

Augmented Reality (AR) stands for the comprehension of modern technology that performs through detection and identification of critical intuition and represents the computer-generated perceptual information (Wikipedia). In recent days, AR is considered one of the prominent advanced technologies for its augmentation prototypes. It enhanced its effectiveness and influenced almost every aspect of life, such as emergency service [2], healthcare management [5][3], industrial implementation, education and training [25][33], design and planning [1][4], and many more.

Therefore, several pieces of research have been undertaken to generate more effective AR prototypes [4] and shown

outstanding achievements [5]. According to Urakov et al. (2019) [6], AR enhances the visualization of the 3-dimensional holographic image through sophisticated glasses. Again, El-hariri et al. (2018) suggested that augmented reality is the visualization process of 3-D imagining data [7]. In some other studies conducted by Vavra et al. (2017) [8] and Kim et al. (2017) [9], augmented reality was described as a projection and interaction-based technology to interact and project computer-generated images in a real environment. Current AR-related research has resulted in several advanced prototypes that promote AR as a prominent technology and call for future research to explore new prototypes and strategies [10][11].

These days, the advancement of AR technology in healthcare has become highly significant with its vast arrays of applications. AR-based healthcare systems are being incorporated with computer vision [6], object detection and identification [18], image processing [10][16], image segmentation [12] and cloud computing technology [14][15]. AR-based healthcare solution has achieved several prominent advancements, such as more secure connectivity and privacy across individual patients [17], more speed while providing quality diagnosis and treatment with more reliability [35]. Besides, wireless services, early diagnosis, real-time monitoring, online consultation, tumor detection, diagnosis specialization, m-health service, and personal assessment are the emerging prototypes of AR-based healthcare solutions, as shown in Fig. 1. Similarly, Medication adherence and rehabilitation treatment have also been enhanced through its utilization. Recently, modern healthcare society introduced an e-health policy to implement prominent augmented devices: AR headset and smart glasses [18]. Therefore, day by day, the popularity of augmented technology in the healthcare system has grown dramatically, and more and more works proposed towards its development.

However, from our empirical observation, it is found that previous review studies have some drawbacks, such as i) a significant number of previous review studies did not present a brief discussion of augmented healthcare applications and their services, ii) few of studies put an effort to provide a comprehensive analysis of the previous works and didn't focus on current limitation of this emerging technology, iii) a majority of them considered only m-health application (medication plan, healthy food guidelines, glucose level monitoring, heartbeat

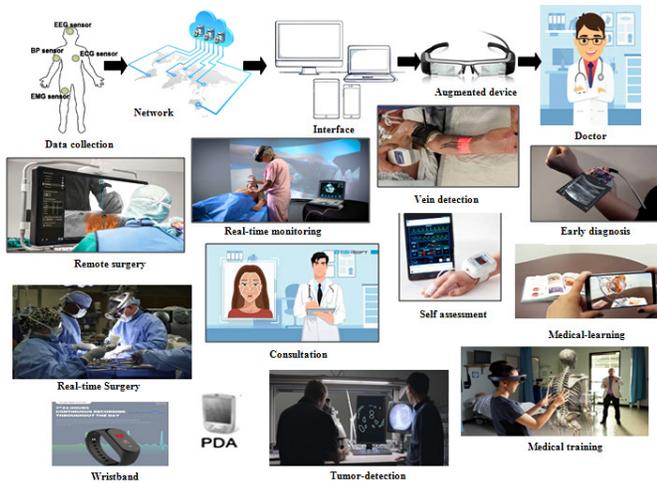


Fig. 1. Current Trends of Augmented Reality in Healthcare Platform.

calculation, etc.) for review/analysis rather than focusing on clinical applications (examines the abnormal joint function, tumor location detection, vein detection, diagnose breast cancer, etc.), and iv) some of the papers addressed recent challenges of augmented healthcare application in short and did not address the possible attacks with proper solutions to mitigate these challenges. Therefore, based on our scientific observation and the author's best knowledge, it can be stated that a comprehensive analysis would be beneficial to identify the limitation and challenges of AR-based healthcare applications. Moreover, this analysis may help to find the solution to these challenges as well as for the further improvement of current applications. This paper aims to investigate the impacts of augmented reality in the healthcare system by addressing above mentioned issues. A brief discussion on augmented healthcare applications (m-health and clinical applications) with their services in the healthcare system is present in this study. Also, a wide range of challenges of the augmented platform in the healthcare system has been demonstrated here with a possible solution against future attacks. Moreover, this paper can be influential in having insights into existing augmented-based healthcare applications. For a further study of AR healthcare solutions, interested readers referred to [23][24][25].

This article has organized as follows: Paper collection and selection for review described in Section 2. Section 3 demonstrates through healthcare applications development prototypes, several augmented mobiles and clinical healthcare applications with services, and current challenges. Section 4 summed up our contribution by addressing several security issues with our proposed intelligent security architecture. Finally, the article concluded with a conclusion and a recommended direction for future research.

II. RESOURCES OF RELATED PAPERS

A. Paper Selection Process

To find the published research articles, we have considered five popular databases in the computer science and health informatics domain for recent development and updated approaches (shown in Fig. 2).

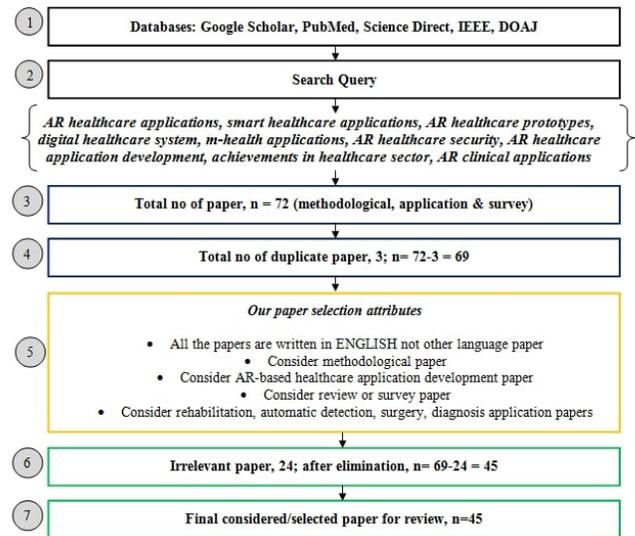


Fig. 2. Related Article Collection and Selection Through Search Query and Predefined Criteria using PRISMA Flow Diagram.

We have considered search queries for searching the relevant papers from the selected database according to the relevancy of our work. For this work, we investigated both peer-reviewed journal and conference articles from 2011 to 2020. Primarily, 72 related articles were found relevant, and after further evaluation, 27 articles have excluded, and 45 articles have finally selected according to our aim. The paper selection process performs following the PRISMA approach. Fig. 2 shows the article extraction and selection process using the PRISMA flow diagram. In Fig. 2, at stage-1, to extract recent publications of augmented and healthcare technology, we considered IEEE, Google Scholar, PubMed, Science Direct, and Directory of open access journal (DOAJ) databases. At stage-2, article searching performs through several keywords such as Augmented Reality (AR) healthcare application, Digital healthcare system, AR-based smart-healthcare applications, AR healthcare application developments, AR-based m-health, recent AR achievements in the health sector, AR healthcare Security, AR clinical application, etc. We added the advanced search option to reveal the most recent articles of our intended domain. Regarding search keywords, we have extracted 72 papers that belong to the methodological approaches, real-life AR healthcare applications development, comparative analysis, survey, and review papers (at stage-3). We performed a manual checking to identify duplicates papers. Three papers have been identified as duplicate and eliminated (stage-4). At stage-5, we set our selection attributes to evaluate the collected papers according to their relevancy with our study aim. Those articles that did not match our selection criteria (Fig. 1, stage-5) did not consider for review. Mostly, those papers that are not written in English and not directly related to the healthcare domain have been marked irrelevant and excluded. After evaluating, 24 articles were removed from the total number (stage-6), and 45 papers were selected for final review (stage-7).

B. Papers Statistics Information

Here, we have categorized the previous articles according to their year of publication, publication database, proposed

TABLE I. CATEGORICAL OVERVIEW OF SELECTED PAPERS FOR REVIEW

No.	Authors	Country	Publication year	Published database	Category	Journal	Conference
1	White et al. [1]	Ireland	2018	Google Scholar	Methodological	—	Yes
2	Jeong et al. [2]	South Korea	2017	DOAJ	Review	Yes	—
3	Agrawal et al. [3]	India	2018	IEEE	Application Development	—	Yes
4	Riva et al. [4]	Italy	2016	DOAJ	Comparative Analysis	Yes	—
5	Sutherland et al. [5]	Canada	2019	Google Scholar	Review	Yes	—
6	Urakov et al. [6]	USA	2019	Science Direct	Application Development	Yes	—
7	El-Hariri et al. [7]	Canada	2018	DOAJ	Methodological	Yes	—
8	Vávra et al. [8]	Czech Republic	2017	PubMed	Review	Yes	—
9	Kim et al. [9]	South Korea	2017	Google Scholar	Survey	Yes	—
10	Frajhof et al. [10]	Brazil	2018	DOAJ	Methodological	Yes	—
11	Mahmood et al. [11]	UK	2018	Science Direct	Comparative Analysis	Yes	—
12	Wilhelm et al. [12]	Germany	2018	PubMed	Review	Yes	—
13	Liu et al. [13]	France	2018	IEEE	Comparative Analysis	—	Yes
14	Jones et al. [14]	USA	2018	DOAJ	Comparative Analysis	Yes	—
15	Chen et al. [15]	China	2011	IEEE	Methodological	—	Yes
16	Ahad et al. [16]	Bangladesh	2018	PubMed	Comparative Analysis	Yes	—
17	Pereira et al. [17]	Chile	2019	Science Direct	Methodological	Yes	—
18	Khor et al. [18]	UK	2016	PubMed	Review	Yes	—
19	Pereira-Azevedo et al. [19]	Netherlands	2018	PubMed	Comparative Analysis	Yes	—
20	Vigliarolo et al. [20]	Italy	2019	DOAJ	Review	Yes	—
21	Garrett et al. [21]	Canada	2018	DOAJ	Comparative Analysis	Yes	—
22	Wong et al. [22]	USA	2018	PubMed	Review	Yes	—
23	Yoon et al. [23]	USA	2018	Google Scholar	Review	Yes	—
24	Belmustakov et al. [24]	USA	2018	Science Direct	Review	Yes	—
25	Uppot et al. [25]	USA	2019	PubMed	Review	Yes	—
26	Rabbi et al. [26]	Pakistan	2012	Google Scholar	Review	Yes	—
27	Pflugl et al. [27]	Switzerland	2017	Google Scholar	Application Development	—	Yes
28	García-Cruz et al. [28]	Spain	2018	Science Direct	Comparative Analysis	Yes	—
29	Nakamoto et al. [29]	USA	2012	DOAJ	Review	Yes	—
30	Daher et al. [30]	USA	2017	IEEE	Comparative Analysis	—	Yes
31	Hemant et al. [31]	Brazil	2020	Google Scholar	Comparative Analysis	Yes	—
32	Moro et al. [32]	Australia	2017	Google Scholar	Comparative Analysis	Yes	—
33	Lahanas et al. [33]	Greece	2016	DOAJ	Survey	Yes	—
34	Maier-Hein et al. [34]	Germany	2013	Science Direct	Methodological	Yes	—
35	Chen et al. [35]	USA	2018	Google Scholar	Application Development	—	Yes
36	Tabrizi et al. [36]	Germany	2015	Google Scholar	Application Development	Yes	—
37	Cavalcanti et al. [37]	Brazil	2018	PubMed	Review	Yes	—
38	Hussain et al. [38]	France	2018	Google Scholar	Application Development	—	Yes
39	Monge et al. [39]	Portugal	2018	IEEE	Methodological	—	Yes
40	Hsiao et al. [40]	Taiwan	2015	Google Scholar	Methodological	Yes	—
41	Shinomiya et al. [41]	Japan	2018	Science Direct	Review	Yes	—
42	Gallos et al. [42]	Greece	2018	PubMed	Survey	Yes	—
43	Ingeson et al. [43]	Sweden	2018	Google Scholar	Methodological	Yes	—
44	Chamberlain et al. [44]	USA	2016	PubMed	Methodological	—	Yes
45	Hayhurst et al. [45]	England	2018	Google Scholar	Review	Yes	—

category, journal paper, and conference paper. The total number of 45 articles have summarized in Table 1. These articles have been selected based on some predetermined criteria (Fig. 2). From Table 1, the following observation may conclude. First, from 2018 to the present, AR healthcare application has become a serious concern for most the computer scientist and healthcare professionals. Second, the number of comparative analysis and review of AR healthcare applications have dramatically increased in recent days. In contrast, the number of application development articles has decreased due to challenges for ensuring appropriate services and security issues. The detailed statistics of selected articles have shown in the following sections.

1) *Study year and reference database:* Here, we conclude the statistics of the considered articles according to their publication year (2011-2020) and referencing databases (PubMed, Science Direct, DOAJ, Google Scholar, and IEEE), which shows in Fig. 3. Among 45 articles, a significant number of works have published in 2018, 2017, and 2016.

Among five databases, PubMed published ten articles, Science Direct published seven articles, DOAJ published nine articles,

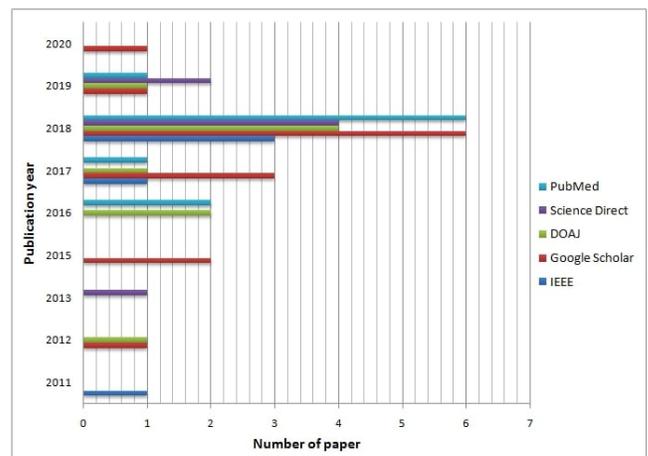


Fig. 3. Number of Articles According to their Publication Year and Referencing Databases.

Google Scholar published thirteen articles, and IEEE published five articles. Moreover, a maximum number of the article we have found for the year 2018. Besides, Google Scholar & PubMed is the more popular database or resource for augmented healthcare domains.

2) *Category of study articles:* To understand the effectiveness of previous works, we analyzed our selected forty-five articles into five categories: methodological, application development, comparative analysis, review, and survey paper. At the same time, we classified these studies into journals and conference publications. Fig. 4 shows that among a total of 45 articles, review, comparative analysis, and methodological papers are maximum in consideration. Among the explained studies, the proposal for methodological studies is ten, six for application developments, eleven for comparative analysis, fifteen for review paper, and three for the survey. The majority of these journal and conference proceedings come from healthcare engineering, computer science, biomedical engineering, and bioinformatics.

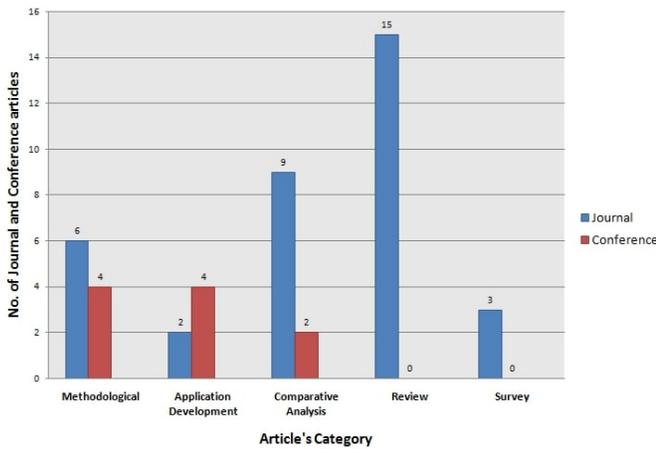


Fig. 4. Number of Articles in Each Category with Paper Status.

3) *Country of study articles:* In this study, we considered articles from both developing and developed countries. Figure 5 shows the number of articles according to their country where majority from the USA, Canada, Germany and Brazil and few from UK, South Korea, Greece, Italy, and France. Compared to these countries Australia, India, Bangladesh, Pakistan, and others show less interest in AR-based healthcare research.

III. COMPREHENSIVE ANALYSIS OF SELECTED ARTICLES

After article selection (explain in Section 2), our primary objective is to perform a comprehensive analysis of these 45 selected articles. To do our analysis, here we have considered our research questions regarding how augmented technology facilitated in healthcare domain; identify the recent augmented healthcare applications with services, and current challenges of existing augmented healthcare applications. After exploring these aspects, based on our observation, we proposed a model to mitigate the recent augmented healthcare application's limitation (Section IV-C).

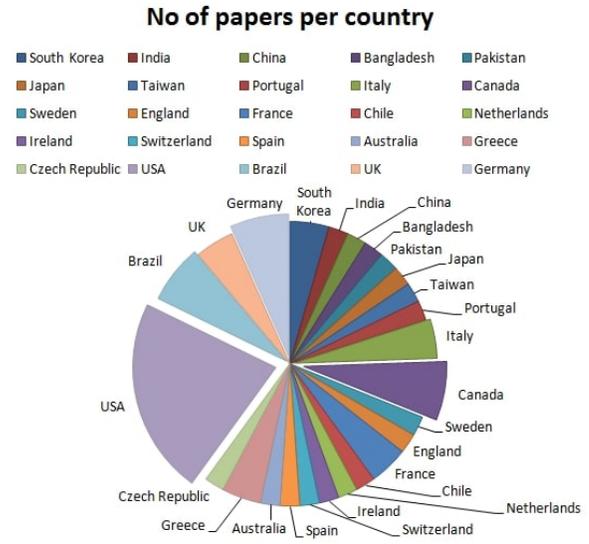


Fig. 5. Number of Articles Per Country.

A. Augmented Reality in Healthcare System

Nowadays, AR technology has attracted significant attention to developing applications that healthcare professionals and patients directly use. Augmented healthcare solutions may apply to a diverse array of fields, including diagnosis, surgery, rehabilitation, monitoring, guidance, education, etc. Generally, AR technology facilitated the healthcare system through its numerous services and platform [29]. The AR service platform refers to a framework that focuses on information acquisition, proper data organization; ensuring data support and providing consistent services, as shown in Fig. 6.

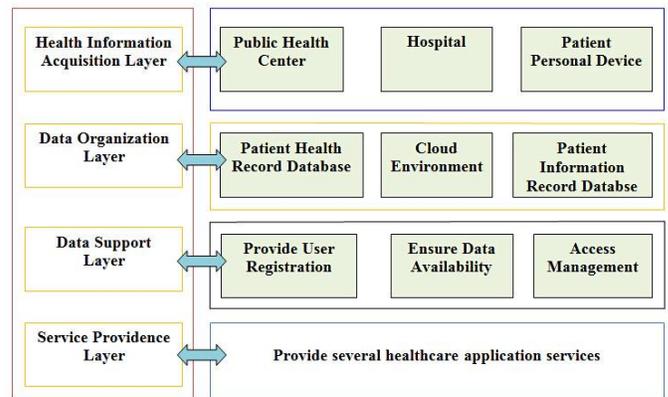


Fig. 6. The Process of Augmented Healthcare Application Services.

1) *Health information acquisition layer:* It refers to collecting healthcare information or patient health reports shared by patients or public health centers or hospitals to make available patient health records to medical specialists or doctors. Generally, healthcare or hospital authority collects these health data and stores them on the hospital database.

2) *Data organization layer:* The data organization layer is the primary layer of the AR-based healthcare application

development process. It benefits transferring real-world information into a virtual environment. This layer extracts the health information from the healthcare database and stores it into a cloud environment to support data preprocessing and model creation. Cloud environment or cloud server refers to data (health data) in the server [27]. Cloud databases provide services for healthcare applications during data examination or visualization. The data preprocessing performed to reduce the noise in MRI and X-ray data. Finally, model creation helps create a visual object for the effective organization of 3-D data [26]. This process allows the generation of a 3-D model from 2-D or 3-D health images or data.

3) *Data support layer*: The data support layer provides services through user registration, data availability, and accessing opportunities. To reduce unknown threats during communication it prerequisites the registration panel. Based on authenticity confirmation, it ensures data availability and accordingly grants access. The authentication process of numerous prototypes ensures through the network interface. Network interface provides the network layer to allow secure communication of the entire prototype of AR healthcare application. Network Layer ensures security through network platforms and network protocols. A network platform is a medium that generates frequencies for communication. Third-generation technology (3G) and fourth-generation technology (4G) are the most popular network platform or network technologies for AR-based healthcare applications. Managing data and ensuring secure connectivity are the prior responsibilities of network protocols. Two widely used protocols are User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). UDP and TCP are responsible for data transmission where UDP is connectionless, reliable, and TCP, errorless.

4) *Service providence layer*: The service providence layer refers to the user interface that allows the extraction of virtual environment data according to availability and accessibility. This layer is also called the process of data collaboration. It helps healthcare specialists overlay the virtually created imagining model on real-time patients' bodies using AR devices. During surgery or diagnosis, on-demand health data accessing is the prime objective of this interface. The majority of healthcare solutions are using smart displays or glass. To detect an object, image, or text projection, Head Mounted Display and Holographic Display device are significant. In recent days, HMD has implemented various healthcare solutions like surgery, diagnosis, assessment, etc. Some prominent functionality of it includes recognizing voice comments, tracking eye movements and gestures movements. At present, powerful and prototype-based HMD devices are OHMDs and OST-HMD. Similarly, a Holographic display is a lighting technology that visualizes a three-dimensional image of objects with better accuracy. Among several holographic displays, HoloLens and Oculus rift is the most used one in healthcare applications. In contrast, smart glass is a wearable, hands-free control system that collects patient's physical information and provides accurate solutions during surgery and diagnosis. It helps to take pictures, record videos, recognize voice instruction, and lessen the requirement of looking on a different screen while examining the medical test report [28]. Smart glasses are being categorized as optical glass and video glass for this study. It is worth mentioning that the majority of AR-based healthcare application uses optical glasses for its augmented functionality.

The most prominent optical glasses are VuFine and Google Glass. Video glasses are called personal media viewers as they provide hands-free convenience. These are categorized into Vuzix, Epson Moverio and Atheer air.

B. AR Healthcare Applications and Services

According to our observation regarding the context of 45 healthcare articles, no standard definition of AR healthcare services has been found. As a result, it gets hard to categorize the healthcare application and their services. However, here, we classified augmented healthcare applications into six categories: including Diagnosis, Surgery, Monitoring, Guidance, Rehabilitation, and Medical Education. We have summarized the available AR healthcare applications and their services as following.

1) *Diagnosis Applications*: Eye Diagnosis, Cancer Diagnosis, Cardiac Diagnosis, Tumor Detection, Vein Detection, and Lung Diagnosis [34] [19] [35] [31] [3] [36].

a) *Services*:

- Identifies the patient's exact eye diseases: Glaucoma, dry eye condition, etc.
- Using a digital contact lens measures blood sugar levels through a multi-sensor for retinal implanted people.
- Used to diagnose breast cancer in sentinel lymph nodes and prostate cancer in prostatectomy specimens.
- To understand cardiac inner and deeper structure.
- Analyzes the heart condition according to normal, murmur and extra-systolic sound.
- Cardiac data examination, prediction and cardiac arrhythmia treatment.
- Identifying skin incision, skull craniotomy and tumor location.
- Provides feedback on lung conditions and is used to diagnose respiratory disease.
- M-application use for vein detection.

2) *Surgery Applications*: Orthopedic Surgery, Cholelithotomy Surgery, Dental Surgery, craniofacial Surgery, HIP Surgery, Urology Surgery, Laparoscopy Surgery [22] [29] [38] [41].

a) *Services*:

- Helps orthopedic surgeon to examine abnormal joint function.
- Identifies the bone structure through 3D computed tomography (CT) data for orthopedic surgery.
- Identify skin incision, skull craniotomy and tumor location for Cholelithotomy surgery.
- Prevents damaging tissues, blood vessels and dental nerves during dental surgery.
- Helps in Endoscopic endonasal transsphenoidal surgery.

- Used for craniofacial surgery applied in treating orbital hypertelorism, hemifacial microsomia, mandibular angle split osteotomy related abnormalities.
- Use for safely HIP placement.
- Enhances urology surgical procedures.
- Provides automatic relevant information to help dental surgeons.
- Used in stone extraction and biopsy procedure.
- Used for treating orbital hypertelorism, hemifacial microsomia, mandibular angle split osteotomy related abnormalities.
- Used for laparoscopy surgery like endophytic tumor surgery.
- Identifies the bone structure through 3D computed tomography (CT) data.
- Reduces blood loss and risk of chronic renal insufficiency.

3) *Monitoring Applications:* Glucose Level Monitoring [42], Insulin Dosages Monitoring [43]

a) *Services:*

- Continuously monitors glucose level, insulin dosages and suggests appropriate foods.
- Helpful for real-time health monitoring and recommending appropriate specialist.
- Automatically connects to the emergency number to seek help or to find defibrillators.
- Measures blood sugar levels through multi-sensor for retinal implanted people.

4) *Guidance Applications:* Grocery Shopping, Medication Plan, Smoking Consequence, Exercise Plan, Medicine Scheduling, Hearing Impaired Guidance [44] [45].

a) *Services:*

- Provides exact guidelines about healthy food.
- Provides information about allergic food, low-fat diets, and general caloric intake.
- Support for medication plan and medication restrictions.
- Helps hearing impaired people for museum visits.
- Shows the consequence of smoking and stimulates the side effects.
- Shows medicine functionality and using procedure.
- Spreads awareness about exercise and make people fit and healthy.
- Recognizes speech commands and provides information on medical needs.
- Provides exact information about drug names and dosages.

5) *Rehabilitation Applications:* Smart Physical Rehabilitation [20] [37][39] [40]

a) *Services:*

- Provides inpatients and outpatient rehabilitation facilities for musculoskeletal, neurological, r-hematological and cardiovascular systems.

6) *Medical Education Applications:* Training, 4D-Visualization [21] [32] [33].

a) *Services:*

- Represents the human organs into 4D anatomies without dissection.
- Helps dental students with both theoretical and practical training.

C. Current Challenges of AR Healthcare Applications

One primal focus of this research is to identify the current challenges of existing AR healthcare applications. Also, we noticed several possible future challenges related to augmented healthcare applications. According to the perspective of our considered articles, here we have conceded eight challenges. Several other challenges might exist in the current AR healthcare application that is beyond our scope of this review.

1) *Challenges with Data security:* Ensure healthcare data security is the most significant aspect of AR healthcare applications. Surprisingly, several exiting applications have no standard security protocols and recently proposed approaches also have a poor security concern. To ensure data security, the implementation of network connection encryption is essential. Providing a complete encrypted or protected data transferring environment for healthcare applications is difficult but adequately needed.

2) *Challenges with Specialized platform:* The majority of augmented healthcare applications have no standard platform and lack appropriate guidelines. The incompetency to provide authentic and accurate services makes it more insecure and complex. Launching a standard and specialized platform that refers to applications package interface (APIs), framework, and appropriate libraries are challenging for modern healthcare platforms. Failure to ensuring a specialized platform may hamper in manages of code, classes, and documentation also.

3) *Challenges with Limited services:* From 2011 to the present, several applications have developed for a limited number of healthcare aspects, for example, diagnosis, treatment, assessment, etc. Yet, more healthcare branches should disclose to provide an effective solution, for example, cardiac surgery, Hematopoietic stem-cell transplantation, etc.

4) *Challenges with Healthcare information accessing interoperability:* The Healthcare Leadership Council (HLC) published several guidelines related to healthcare IT interoperability. For accessing and sharing information, ensuring HLC guidelines are a big concern.

5) *Challenges with Scope of adaptability:* Complex functionality and lack of user guidance make a barrier to healthcare application adaptability. Generally, augmented healthcare application is the reflection of integrated digital components. These integrated digital components might introduce some technical barriers, especially for the healthcare specialist and the end-user. Overcoming the technical obstacle is one of the utmost challenges.

6) *Challenges with Lack of availability:* Several AR healthcare applications launch in the last few years. But a majority of them have been used in the developed country only. High configuration prototypes, training requirements, and cost issues are responsible for these applications limitation in underprivileged countries. Ensuring these applications available for all the country is a challenging issue.

7) *Challenges with Implementation cost:* Furthermore, the development and implementation cost of the AR healthcare device is another vital issue. Due to the high development cost, several healthcare providers disagree to reconstruct their healthcare platform. Therefore, an increased range of implementation costs raises the barrier to introducing modern and digital technology into healthcare solutions.

8) *Challenges with Healthcare device acceptability:* The current augmented platform is committed to providing a user-friendly environment with fluent functionality. But sometimes, issues associated with improving consistency and reducing errors can be more viable. Overcoming this challenge may help to increase the device's acceptability. Developing an application with effective functionalities is a relatively focusable issue for AR technology.

9) *Challenges with Inappropriate prototypes:* Due to inappropriate prototypes, sometimes, unexpected abnormalities such as false recognition, lack of semantic knowledge reduce the effectiveness of augmented healthcare applications. Ensure appropriate prototypes are one of the distinct challenges of healthcare application.

10) *Challenges with Service availability:* Instead of numerous AR applications, providing constant services or long-duration services such as surgery visualization, projection, and interaction for as long as 9-12 hours in a continuous manner is relatively challenging.

11) *Challenges with Real-time geometry awareness:* Real-time geometry awareness is another major issue for researchers and developers. It associates with the process of annotation, tracking location, object detection, and appropriate measurement. Surprisingly, most abnormalities arise during geometry transformation in the augmented application. Therefore, awareness regarding real-time geometry can be a challenging issue for AR platforms.

12) *Challenges with Limited research:* According to the research statistics, it has shown that AR research on the healthcare domain is insufficient due to limited scope, less popularity for healthcare professionals, ambiguous knowledge, and limited resources. Therefore, most healthcare researchers have not shown their interest to research in the healthcare domain, specifically on augmented platforms.

IV. OUR CONTRIBUTION

Section 3 shows that several challenges have existed in the recent augmented healthcare platform. Among several current AR healthcare application issues, ensuring data security is one of the major concerns. Generally, augmented healthcare application provides services through a physical layer with overlaid computer-generated information [30]. Therefore, it is not free from external threats, and cybersecurity risk might introduce on the physical layer. AR healthcare applications should address these external threats and common security vulnerabilities during the development phase. Therefore in this section, we addressed several security requirements and attack terminologies to ensure data security and added a security model that might help prevent threats or network attacks during data transmission.

A. Security Requirements

Security requirements of augmented healthcare applications refer to the property that protects the application platform from unexpected attacks or threats. Figure 7 shows some required security requirements that should focus during application development to ensure application effectiveness and modality.



Fig. 7. Security Requirements of AR Platform.

For example, during the collaboration of virtual objects and real-time data, an intruder may attack. The augmented platform should provide security during virtual and real-time object collaboration. At present, the number of augmented healthcare devices has increased and started to connect with the network. Therefore, the augmented platform should ensure confidentiality to prevent unregistered access to medical information to provide medical data security. However, the assorted number of things can modify data while accessing or transmitting. Healthcare solutions should introduce integrity to ensure

medical information or health data security. Another crucial requirement is resiliency to provide a protective environment if health devices are compromised.

Another major security issue of current healthcare applications is availability. It ensures that data are stored and available in a secure environment and used by the registered authority under secure networks. To increase efficiency and protect patient data from untraceable attacks, ensuring the tracking functionality of all the activities with an authorized detector is essential. Most of the time, lack of awareness about DoS attacks can create opportunities for malicious attackers to hamper data safety while transmitting. Therefore, security awareness should consider sincerely. As we previously mentioned, augmented healthcare devices are wearable technologies to connect to the network anywhere. Hence, deriving a security model for dynamic network topology is highly required.

Designing a dynamic security protocol could increase the security of augmented applications. Surprisingly, augmented Reality Markup Language (ARML) has a lack of security control policy. Lack of security issues and Limited standardization makes it vulnerable, which in turn can introduce security vulnerabilities. In recent days, augmented reality uses computer-aided objects provided by third-party vendors that pave the way for threats such as spoofing, sniffing, data manipulation, and man-in-middle attacks. The AR platform should be aware of these threats and be sure of the authenticity of the generated content during the transmission process to make the application reliable. As augmented healthcare application often uses various unprotected web browsers, it can lead to security breach too.

B. External Attack Terminology

The augmented platform is not free from external attacks. Attackers may introduce several security threats and affect existing and future augmented healthcare applications, devices, and networks. Sometimes these threats are predictable, while some are hard to detect or predict. In this section, we are going to mention several external threats that reduce application and network effectiveness. Fig. 8 shows some attack terminologies that can make systems vulnerable and reduce effectiveness.

For example, by data-stealing or spreading malicious codes during data transmission or communication, intruders may replace original data or make it unreliable. Through unauthorized access, attackers may modify the data and create confusion and reduce network performance. Also, false information may get injected into the network and minimize data reliability, affect the hardware platform by altering program code, information, or reprogramming. Sometimes an attack can be introduced through software platforms such as operating systems or software applications and take advantage of software vulnerabilities that cause buffering, resource destruction or loss.

C. Proposed Data Security Model

Augmented reality-based developed healthcare applications and services are not yet robust but continuing to develop. Therefore, summing up the entire possible vulnerabilities, threats, and attacks associated with augmented medical solutions is difficult. With the expansion of healthcare devices, networks, and applications, several unknown or unseen threats

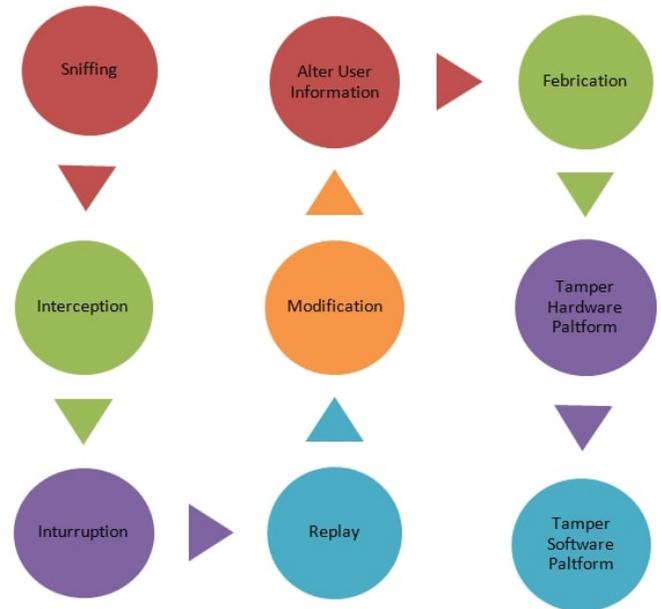


Fig. 8. Possible Attacks for Security Threats.

may initiate during communication or data transmission, or data storage in the cloud.

A dynamic security protocol is essential to mitigate unseen data threats and provide data transmission security. Therefore, security services should design with dynamic properties. An artificial intelligence-based dynamic mechanism might be capable of identifying these unseen attacks. Here we proposed a security model for AR-based healthcare solutions. The proposed model operates through knowledge-based services. Figure 9 presents the proposed scheme that follows three security layers: threat detection, threat reduction, and data protection.

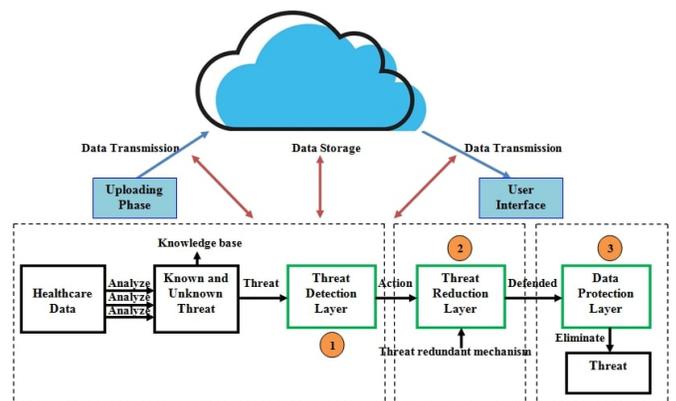


Fig. 9. Proposed Intelligent Security Model.

The threat detection layer is designed to receive healthcare data from healthcare devices and networks and analyzes captured information to identify security threats. Then, to reduce the attack during data transmission, a threat reduction layer is designed. The data protection layer helps protect data by

defending all identified attacks. This security layer performs through dynamic algorithms and provides a shield against attacks. Upon detection, the threat detection layer issues action commands and shares detection experience with the threat reduction layer. The threat reduction layer defends threats using a redundant threat mechanism. Finally, the data protection layer eliminates detected threats and protects valuable data.

V. CONCLUSION

Researchers are vigorously working to invent better technological solutions to enhance the modern health care system throughout the world. The motive of these inventions is to bring a dramatic change in the healthcare sector and reduce the existing complexities. In this paper, we have provided a brief analysis of various AR-based healthcare services and applications. A broad view of the current development strategies regarding present healthcare application architecture and health data processing and accessing procedures have been layout. The article, to some extent, tried to facilitate further development by pointing out several uncovered issues regarding concurrent security requirements and future challenges. The discussion performed in this article on standardization, data availability, service quality, and data protection may help in several ways for future research on AR-based healthcare applications and services. Moreover, this paper also illustrated the importance of augmented reality-based healthcare applications, backed up with present-day market data, which may increase the contribution of several stakeholders for further development.

REFERENCES

- [1] White. G, Cabrera. C, Palade. A, and Clarke. S, Augmented reality in iot, In International Conference on Service-Oriented Computing, pp. 149-160, Springer, Cham, https://doi.org/10.1007/9783-030-17642-6_13, 2018.
- [2] Jeong. B, and Yoon. J, Competitive intelligence analysis of augmented reality technology using patent information, Sustainability, 9(4), pp. 497, <https://doi.org/10.3390/su9040497>, 2017.
- [3] Agrawal. D, Mane. S. B, Pacharne. A, and Tiwari. S, IoT Based Augmented Reality System of Human Heart: An Android Application, In 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 899-902, IEEE, <https://doi.org/10.1109/ICOEI.2018.8553807>, 2018.
- [4] Riva. G, Baños. R. M, Botella. C, Mantovani. F, and Gaggioli. A, Transforming experience: the potential of augmented reality and virtual reality for enhancing personal and clinical change, Frontiers in psychiatry, 7, 164, <https://doi.org/10.3389/fpsy.2016.00164>, 2016.
- [5] Sutherland. J, Belec. J, Sheikh. A, Chepelev. L, Althobaity. W, Chow. B. J, and La Russa. D. J, Applying modern virtual and augmented reality technologies to medical images and models, Journal of digital imaging, 32(1), pp. 38-53, <https://doi.org/10.1007/s10278-018-0122-7>, 2019.
- [6] Urakov. T. M, Wang. M. Y, and Levi. A. D, Workflow caveats in augmented reality-assisted pedicle instrumentation: cadaver lab, World neurosurgery, 126, e1449-e1455, <http://doi.org/10.1016/j.wneu.2019.03.118>, 2019.
- [7] El-Hariri. H, Pandey. P, Hodgson. A. J, and Garbi. R, Augmented reality visualization for orthopaedic surgical guidance with pre- and intra-operative multimodal image data fusion, Healthcare Technology Letters, 5(5), pp. 189-193, <http://doi.org/10.1049/htl.2018.5061>, 2018.
- [8] Vávra. P, Roman. J, Zonča. P, Ihnát. P, Němec. M, Kumar. J, and El-Gendi. A, Recent development of augmented reality in surgery: a review, Journal of healthcare engineering, 2017.
- [9] Kim. S. K, Kang. S. J, Choi. Y. J, Choi. M. H, and Hong. M, Augmented-reality survey: from concept to application, KSII Transactions on Internet and Information Systems (TIIS), 11(2), pp. 982-1004, <https://doi.org/10.3837/tiis.2017.02.019>, 2017.
- [10] Frajhof. L, Borges. J, Hoffmann. E, Lopes. J, and Haddad. R, Virtual reality, mixed reality and augmented reality in surgical planning for video or robotically assisted thoracoscopic anatomic resections for treatment of lung cancer, Journal of Visualized Surgery, 4, pp. 143, <https://doi.org/10.21037/jovs.2018.06.02>, 2018.
- [11] Mahmood. F, Mahmood. E, Dorfman. R. G, Mitchell. J, Mahmood. F. U, Jones. S. B, and Matyal. R, Augmented reality and ultrasound education: initial experience, Journal of cardiothoracic and vascular anesthesia, 32(3), pp. 1363-1367, <https://doi.org/10.1053/j.jvca.2017.12.006>, 2018.
- [12] Wilhelm. D, Vogel. T, Ostler. D, Marahrens. N, Kohn. N, Koller. S, and Kranzfelder. M, Enhanced visualization: from intraoperative tissue differentiation to augmented reality, Visceral medicine, 34(1), pp.52-59, <https://doi.org/10.1159/000485940>, 2018.
- [13] Liu. Y, and Tang. P, The prospect for the application of the surgical navigation system based on artificial intelligence and augmented reality, In 2018 IEEE international conference on artificial intelligence and virtual reality (AIVR), pp. 244-246, IEEE, <https://doi.org/10.1109/AIVR.2018.00056>, 2018.
- [14] Jones. L. D, Golan. D, Hanna. S. A, and Ramachandran. M, Artificial intelligence, machine learning and the evolution of healthcare: A bright future or cause for concern?, Bone & joint research, 7(3), pp. 223- 225, <https://doi.org/10.1302/2046-3758.73.BJR-2017-0147.R1>, 2018.
- [15] Chen. M, Ling. C, and Zhang. W, Analysis of augmented reality application based on cloud computing, In 2011 4th International Congress on Image and Signal Processing, Vol. 2, pp. 569-572, IEEE, <https://doi.org/10.1109/CISP.2011.6100311>, 2011.
- [16] Ahad. M. A. R, Kobashi. S, and Tavares. J. M. R, Advancements of image processing and vision in healthcare, Journal of Healthcare Engineering, pp.3-6, <https://doi.org/10.1155/2018/8458024>, 2018.
- [17] Pereira. N, Kufeke. M, Parada. L, Troncoso. E, Bahamondes. J, Sanchez. L, and Roa. R, Augmented reality microsurgical planning with a smartphone (ARM-PS): a dissection route map in your pocket, Journal of Plastic, Reconstructive & Aesthetic Surgery, 72(5), pp. 759-762, <https://doi.org/10.1016/j.bjps.2018.12.023>, 2019.
- [18] Khor. W. S, Baker. B, Amin. K, Chan. A, Patel. K, and Wong. J, Augmented and virtual reality in surgery—the digital surgical environment: applications, limitations and legal pitfalls, Annals of translational medicine, 4(23), <https://doi.org/10.21037/atm.2016.12.23>, 2016.
- [19] Pereira-Azevedo. N. M, and Venderbos. L. D, eHealth and mHealth in prostate cancer detection and active surveillance, Translational andrology and urology, 7(1), pp. 170, <https://doi.org/10.21037/tau.2017.12.22>, 2018.
- [20] Viglialoro. R. M, Condino. S, Turini. G, Carbone. M, Ferrari. V, and Gesi. M, Review of the augmented reality systems for shoulder rehabilitation, Information, 10(5), pp. 154, <https://doi.org/10.3390/info10050154>, 2019.
- [21] Garrett. B. M, Anthony. J, and Jackson. C, Using mobile augmented reality to enhance health professional practice education, Current Issues in Emerging eLearning, 4(1), pp. 10, <https://scholarworks.umb.edu/ciee/vol4/iss1/10>, 2018.
- [22] Wong. K, Yee. H. M, Xavier. B. A, and Grillone. G. A, Applications of augmented reality in otolaryngology: a systematic review, Otolaryngology–Head and Neck Surgery, 159(6), pp. 956-967, <https://doi.org/10.1177/0194599818796476>, 2018.
- [23] Yoon. J. W, Chen. R. E, Kim. E. J, Akinduro. O. O, Kerezoudis. P, Han. P. K, and Quinones-Hinojosa. A, Augmented reality for the surgeon: systematic review, The International Journal of Medical Robotics and Computer Assisted Surgery, 14(4), e1914, <https://doi.org/10.1002/rcs.1914>, 2018.
- [24] Belmustakov. S, Bailey. C, and Weiss. C. R, Augmented and virtual reality navigation for interventions in the musculoskeletal system, Current Radiology Reports, 6(9), pp. 1-10, <https://doi.org/10.1007/s40134-018-0293-5>, 2018.
- [25] Uppot. R. N, Laguna. B, McCarthy. C. J, De Novi. G, Phelps. A, Siegel. E, and Courtier. J, Implementing virtual and augmented reality tools for radiology education and training, communication, and clinical care, Radiology, 291(3), pp. 570-580, <https://doi.org/10.1148/radiol.2019182210>, 2019.
- [26] Rabbi. I, Ullah. S, and Khan. S. U, Augmented Reality Tracking Techniques: A Systematic Literature Review Protocol, IOSR Journal of Computer Engineering (IOSRJCE) ISSN, pp. 2278-0661, <https://doi.org/10.9790/0661-0222329>, 2012

- [27] Pflugl. S, Lerch. T, Vasireddy. R, Boemke. N, Tannast. M, Ecker. T. M, and Zheng. G, Augmented Marker Tracking for Periacetabular Osteotomy Surgery: A Cadaver Study, *CAOS*, 1, pp. 54-57, <https://doi.org/10.29007/9mbb>, 2017.
- [28] García-Cruz. E, Bretonnet. A, and Alcaraz. A, Testing smart glasses in urology: clinical and surgical potential applications, *Actas Urológicas Españolas*, 42(3), pp. 207-211, <https://doi.org/10.1016/j.acuroe.2018.02.004>, 2018.
- [29] Nakamoto. M, Ukimura. O, Faber. K, and Gill. I. S, Current progress on augmented reality visualization in endoscopic surgery, *Current opinion in urology*, 22(2), pp. 121-126, <https://doi.org/10.1097/MOU.0b013e3283501774>, 2012.
- [30] Daher. S, Optical see-through vs. spatial augmented reality simulators for medical applications, In 2017 IEEE Virtual Reality (VR), pp. 417-418, IEEE. <https://doi.org/10.1109/VR.2017.7892354>, 2017.
- [31] Hemanth. J. D, Kose. U, Deperlioglu. O, and de Albuquerque. V. H. C, An augmented reality- supported mobile application for diagnosis of heart diseases, *The Journal of Supercomputing*, 76(2), pp. 1242-1267, <https://doi.org/10.1007/s11227-018-2483-6>, 2020.
- [32] Moro. C, Štromberga. Z, Raikos. A, and Stirling. A, The effectiveness of virtual and augmented reality in health sciences and medical anatomy, *Anatomical sciences education*, 10(6), pp. 549-559, <https://doi.org/10.1002/ase.1696>, 2017.
- [33] Lahanas. V, Georgiou. E, and Loukas. C, Surgical simulation training systems: box trainers, virtual reality and augmented reality simulators, *International Journal of Advanced Robotics and Automation*, 1(2), pp. 1-9, <https://doi.org/10.15226/2473-3032/1/2/00109>, 2016.
- [34] Maier-Hein. L, Mountney. P, Bartoli. A, Elhawary. H, Elson. D, Groch. A, and Stoyanov. D, Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery, *Medical image analysis*, 17(8), pp. 974-996, <https://doi.org/10.1016/j.media.2013.04.003>, 2013.
- [35] Chen. P. H, Gadepalli. K, MacDonald. R, Liu. Y, Nagpal. K, Kohlberger. T, and Stumpe. M. C, An augmented reality microscope for real-time automated detection of cancer, In Proc. Annu. Meeting American Association Cancer Research, 2018.
- [36] Tabrizi. L. B, and Mahvash. M, Augmented reality-guided neurosurgery: accuracy and intraoperative application of an image projection technique, *Journal of neurosurgery*, 123(1), pp. 206-211, 2015.
- [37] Cavalcanti. V. C, Santana. M. I. D, Da Gama. A. E, and Correia. W. F, Usability assessments for augmented reality motor rehabilitation solutions: a systematic review, *International Journal of Computer Games Technology*, 2018, <https://doi.org/10.1155/2018/5387896>, 2018.
- [38] Hussain. R, Lalande. A, Marroquin. R, Girum. K. B, Guigou. C, and Grayeli. A. B, Real-time augmented reality for ear surgery, In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 324-331, Springer, Cham, https://doi.org/10.1007/978-3-030-00937-3_38, 2018.
- [39] Monge. J, and Postolache. O, Augmented reality and smart sensors for physical rehabilitation, In 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), pp. 1010-1014, IEEE. <https://doi.org/10.1109/ICEPE.2018.8559935>, 2018.
- [40] Hsiao. K. F, and Rashvand. H. F, Data modeling mobile augmented reality: integrated mind and body rehabilitation, *Multimedia Tools and Applications*, 74(10), pp. 3543-3560, <https://doi.org/10.1007/s11042-013-1649-8>, 2015.
- [41] Shinomiya. A, Shindo. A, Kawanishi. M, Miyake. K, Nakamura. T, Matsubara. S, and Tamiya. T, Usefulness of the 3D virtual visualization surgical planning simulation and 3D model for endoscopic endonasal transsphenoidal surgery of pituitary adenoma: technical report and review of literature, *Interdisciplinary Neurosurgery*, 13, pp. 13-19, <https://doi.org/10.1016/j.inat.2018.02.002>, 2018.
- [42] Gallos. P, Georgiadis. C, Liaskos. J, and Mantas. J, Augmented Reality Glasses and Head-Mounted Display Devices in Healthcare, *Studies in health technology and informatics*, 251, pp. 82-85, 2018.
- [43] Ingeson. M, Blusi. M, and Nieves. J. C, Smart augmented reality mHealth for medication adherence, In *AIH@ IJCAI*, pp. 157-168, 2018.
- [44] Chamberlain. D., Jimenez-Galindo. A, Fletcher. R. R., and Kodgule. R, Applying augmented reality to enable automated and low-cost data capture from medical devices, In Proceedings of the Eighth International Conference on Information and Communication Technologies and Development, pp. 1-4, <https://doi.org/10.1145/2909609.2909626>, 2016.
- [45] Hayhurst. J, How augmented reality and virtual reality is being used to support people living with dementia—Design challenges and future directions, *Augmented reality and virtual reality*, pp. 295-305, https://doi.org/10.1007/978-3-319-64027-3_20, 2018.

Identification of Abusive Behavior Towards Religious Beliefs and Practices on Social Media Platforms

Tanvir Ahammad¹

Department of Computer Science
and Engineering
Jagannath University, Bangladesh

Md. Khabir Uddin²

Department of Computer Science
and Engineering
Jagannath University, Bangladesh

Tamanna Yesmin³

Department of Computer Science
and Engineering
Uttara University, Bangladesh

Abdul Karim⁴

Department of Computer Science
and Engineering
Jagannath University, Bangladesh

Sajal Halder⁵

Department of Computer Science
RMIT University
Melbourne, Australia

Md. Mahmudul Hasan⁶

Department of Computer Science
and Engineering
Dhaka International University, Bangladesh

Abstract—The ubiquitous use of social media has enabled many people, including religious scholars and priests, to share their religious views. Unfortunately, exploiting people’s religious beliefs and practices, some extremist groups intentionally or unintentionally spread religious hatred among different communities and thus hamper social stability. This paper aims to propose an abusive behavior detection approach to identify hatred, violence, harassment, and extremist expressions against people of any religious belief on social media. For this, first religious posts from social media users’ activities are captured and then the abusive behaviors are identified through a number of sequential processing steps. In the experiment, Twitter has been chosen as an example of social media for collecting dataset of six major religions in English Twittersphere. In order to show the performance of the proposed approach, five classic classifiers on n-gram TF-IDF model have been used. Besides, Long Short-term Memory (LSTM) and Gated Recurrent Unit (GRU) classifiers on trained embedding and pre-trained GloVe word embedding models have been used. The experimental result showed 85% accuracy in terms of precision. However, to the best of our knowledge, this is the first work that will be able to distinguish between hateful and non-hateful contents in other application domains on social media in addition to religious context.

Keywords—Social media; religious abuse detection; religious keywords; religious hatred; feature extraction; classifier

I. INTRODUCTION

In the modern age of Information and Communication Technology (ICT), social media platforms seem to be an indispensable part of human lifestyle. Everyday, millions of individuals share their opinions, ideas, thoughts, feelings, experiences through social media services such as Twitter, Facebook, YouTube, Instagram and so on [1], [2]. As the people are not only sharing opinions but also connecting with new people, creating groups and making friendships, so social media has become an important source of information with variety of communities [3]. Moreover, different types of people, even many people who would be speechless previously, now use social media for different purposes in response to the fulfillment of the desires or goals of the mind [4]. Despite

of having such a positive effect on freedom of expression on social media, we can not ignore its negative impacts. Some people either intentionally or unintentionally involved in abusive activities such as spreading false news and videos, trolling, cyberbullying and many more [5].

The people concerned with religion find the social media as useful platforms for sharing religious beliefs, rules and rituals to large audiences. They want to increase affiliations and trust on their religion through different types of activities on social media [6], [7]. During COVID-19 pandemic time, this scenario has increased significantly [8]. However, some people use hate speech against other religions to justify their religion. Furthermore, exploiting people’s religious beliefs and practices, some extremist groups spread provocative and hateful content on social media. Consequently, the falsified and hateful information creates instability among the religious communities since social media news spread so fast by liking and sharing [9], [10]. Moreover, all these religious narrowness increase the likelihood of militant propaganda. As the people are very sensitive regarding religion, so it is essential to restrict the abusers who spread religiously offensive or hate speech on social media.

As religion refers to the ideological identity of a person [11], so expressing hatred by hurting a religion is like hurting a person’s identity. In the literature, there is limited research on detecting religious abuse on social media. In our previous work [12], we addressed the concept of identification of religious abusers on Twitter. We only considered verifying whether abusive activities originated from spamming sources. Besides, there have been many works on social media about hate speech, such as identifying hate speech on Arabic social media [13], [14], which do not mention religious abuses. There is also another similar research in the literature demonstrating hate speech identification on vulnerable communities [15]. Detecting spamming activities in social media [16] is another type of research that is somewhat related to religious abuse detection. However, it should be admitted that not all abusers may use spamming techniques in spreading offensive speech

regarding religion. There has been some more researches on identification of spreading propaganda on social media regarding different issues, like jihadist propaganda [17], [18], [19] and propaganda of COVID-19 deadly virus [20], [21], [22], [23] and then showing how the propaganda is analyzed [24]. Most of these are focused on Arabic social media context and targeted to single community. Therefore, there is a need for identifying abusive/offensive expressions from the posts and comments of every user who expresses religious views on social media.

Each social media has its own policy regarding different abusive activities, but it is still difficult to find out those activities in every user's posts. More specifically, identifying religiously abusive contents on social media are hard due to the noisy data structure. In other words, social media data consists of misspelled user-generated data (as users express freely anything they want without following any rules in writing), jargon, heterogeneity in data with the mixture of texts, urls, videos, emojis and so on [25].

In this paper, we focus on identifying abusive attitudes towards religion on social media. We refer to those activities religiously abusive that represents hatred, violence, harassment and extremist expressions against people of any religion or community, as depicted in Fig. 1 for an example. However, to find abusive contents, first we retrieved social media posts using a predefined set of religious keywords for six major religious beliefs and filtered them in order to remove unwanted information. Then, we labeled the filtered data with lexicon and rule based approaches. After that, we extracted features using traditional and advanced deep learning strategies from the texts after preprocessing. Finally, we fed the dataset into classifiers to identify abusive speech/content. In our experiments, we used Twitter as a social media platform for collecting the dataset. We compared the classical classification models, including Naïve Bayes, Support Vector Machine (SVM), Random Forest, Logistic Regression, and MLP (Multilayer Perceptron) on TF-IDF feature extraction method. We also compared trained word embedding with pre-trained GloVe embedding in terms of LSTM and GRU models. The experimental result shows that we can obtain 85% overall accuracy in state-of-the-art performance in terms of precision.



Fig. 1. An Example of Abusive Speech on Social Media Against Religious Communities.

To the best of our knowledge, this is first work that represents an approach to identify the abusive attitudes towards religion on social media platforms. That is, we can detect

religiously abusive activities on any social media platform. In sum, our contributions are as follows:

- Our proposed approach examines the application of identifying religious abusive expressions on social media platforms.
- We identify abusive behaviour from social media users' posts for major religious beliefs in the world.
- We can use the proposed approach to detect hate/offensive speech on other application domains in addition to religious context.

The rest of the article is organized as follows: Section II represents recent studies in the literature related to this paper. In Section III, we have demonstrated and explained every module of the proposed approach. Then, Section IV represents the experimental procedures including data collection, feature extraction, building classification models and so on. Next, we have shown and discussed experimental results elaborately in Section V. Finally, we have concluded the work with future directions in Section VI.

II. RELATED WORK

People use social media to express their feelings and find emotional gratifications for various reasons [26]. In religious point of view, the use of social media is making a significant contribution to the development of religious values. That is, the linkage between religion and social media has attracted many people over the years, including religious scholars and priests [27]. Despite of these advantages, abusive activities on religion can not be ignored. So, we need to pay attention to the negative impact of using social media in the context of religion. However, research on how to automatically detect offensive remarks from every user's posts and comments is limited in literature. Our previous work [12] highlighted the concept of detecting religious abusers on Twitter. It only showed that the spamming sources were identified as religiously offensive.

Detecting hate speech on Arabic twittersphere is very promising research tread in literature nowadays. Albadi et al. [28] published first publicly available annotated dataset and three lexicons with hate scores in order to detect religious hate speech from Arabic tweets. They classified extracted tweets as hate and not hate speech in terms of lexicon based, n-gram, GRU plus word embedding based, and GRU word embedding with handcrafted features including temporal, user and content features. The experimented result showed that feature based GRU model gave the best accuracy in terms of recall. However, their approach makes a pathway to find hate speech in Arabic tweets but it can not be applicable in many religious contexts such as Hinduism, Buddhism and so on. Besides, the annotated lexicons are unavailable for English contents.

Spreading offensive expressions in many contexts, such as Islamophobia, anti-Africa, and anti-Arab, on social media against people of different groups is not a good sign. Z. Mossie and J. H. Wang [15] were the first to find hatred against minor communities in Ethiopia from Amharic texts on Facebook. They annotated the dataset by the handcrafted method and then clustered the hate words using Word2Vec method. Although their approach successfully identified hate speech against vulnerable groups, but handcrafted features and

annotation tend to the possibility of biasing on any cultural group. There is another research on communal hatred detection that was published by B. Vidgen and T. Yasseri [29]. They identified hate speech in their work on Muslims in terms of non-Islamophobic, weak and strong Islamophobic speech on Twitter. They considered the names of Muslims and Mosques in their analysis. It is not the case that only name of Muslim or Mosque spread Islamophobia, rather the hateful behavior, threats and online harassment against Muslims incite Islamophobia. So, these are also essential to consider.

G. Jain et al. [16] proposed a novel deep learning based approach (combination of CNN and LSTM) to identify spam detection in social media. They considered spamming behavior of user and short text messages in their proposed model. Similar type of research was conducted by N. Sun et al. [30] to find near real-time spam on Twitter. They identified spams in terms of number of tweets issued by a user, number of retweeted, and fake accounts using traditional machine learning models. In articles [13] and [14] we found how deep learning models were used to identify hate speech in Arabic tweets that inspired us to do the work. Although these researches (on spam and hate speech detection) are somewhat related to our work, but it is acknowledged that all abusers may not use spamming tactics or use Arabic language in spreading offensive speech regarding religion. Therefore, we need to establish a model to analyse (lexically, syntactically, and/or semantically) the abusive/offensive behavior of each user when sharing religious views on social media platforms.

III. METHODOLOGY

This section represents our proposed approach that processes religious posts from social media users' activity and then identifies the abusers who spread hatred, violence, harassment, and extremist expressions among different religious communities. It consists of several segments with different functionalities, namely, Information Retrieval, Filtering, Labelling Class Attribute, Text Normalization, Text Vectorization, and Building and Evaluating Classifiers, represented in Fig. 2.

Information Retrieval (IR) is the method of capturing and extracting relevant users' activity from social media platforms. It follows three sequential processing steps. First of all, the social media from which we want to retrieve data needs to provide the required credentials (e.g., access token, consumer key, page id or post id) in compliance with all the terms and conditions. Then, it is necessary to identify the type of data (content language preferences, e.g., English, Japanese, or Arabic) we want to collect through religious search patterns or keywords. Finally, the collected data is transformed into tabular form by selecting different attributes.

Since not all users' activities or social media posts are related to religiously abusive, so only interested of these are considered for further processing. In other words, users' posts that are likely to have an impact on the religious community are filtered out. However, the filtering process follows three phase filtering schemes. In phase one, the information of the best possible impacted posts (or status) are considered, including keeping all unique posts, and posts that have length greater than a threshold. For example, the best impacted post length is between 70-100 in Twitter [31]. The second phase

filters information based on the reputation of the user account, calculated as

$$R = \frac{\text{Number of followers}}{\text{Number of friends} + \text{Number of followers}}$$

If the R (stands for reputation) is small and close to zero, then it is probable to be an abusive account, as the abusers generally tend to get more followers [32]. Phase third of the filtering represents users' activity including the duration of account, date of posts, number of sharing (or retweeting) posts, number of posts issued, and number of likes or dislikes. Thus, the information of the higher activity is kept for future processing steps.

Labelling class attribute defines introducing the class attribute in the filtered dataset. As dataset consists of different types of attributes, so labelling class attribute is created on text data (i.e., religious posts or status of users). At first, text data is cleaned in order to remove unwanted texts or symbols, such as URLs, hashtags, special symbols, and punctuation. Then, generate religious lexicon containing words and phrases with their own polarity scores to be abusive, non-abusive, or neutral. After that, the cleaned texts are attributed as abusive or non-abusive class label using rule-based and generated lexicon-based approaches.

Before feeding the corpus into machine learning models, it is essential to transform the texts into a standard form or normalized form. At the beginning of the normalization, the texts are fragmented into smaller units called tokens, e.g., 'this is religiously abusive' → 'this', 'is', 'religiously', 'abusive'. Then, remove all unwanted words such as 'about', 'above', 'across' from texts with predefined stop word list that is particularly applicable for social media text analytics. At the end of normalization, identical or near identical words (or tokens) are mapped into its base form, called stemming and lemmatization. That is, stemming and lemmatization help to convert the root form of inflected words. For example, the word "followers" and "followings" are transformed to "follow", its root form. Another example is representing of near similar words such as "keywords", "key-words" and "key words" to just "keywords".

After normalization, the textual data in the corpus is mapped into real valued vector form, called text vectorization or feature extraction from texts. That is, it is the process of making textual document into numerical vector form. In Fig. 2, text vectorization is illustrated by feature matrix which is defined as

- **w1, w2, w3,...,wn** represents features (n-grams or words);
- **Doc** represents textual documents (D1, D2, D3,...,Dn) in the corpus where each document indicates a social media post;
- **Class** represents class label attribute; and
- **Each row** represents a text data of a religious post containing feature values in it.

Building and evaluating classifiers is the final stage of the proposed approach where a classification model is built by feeding the feature matrix into it and then evaluate the model

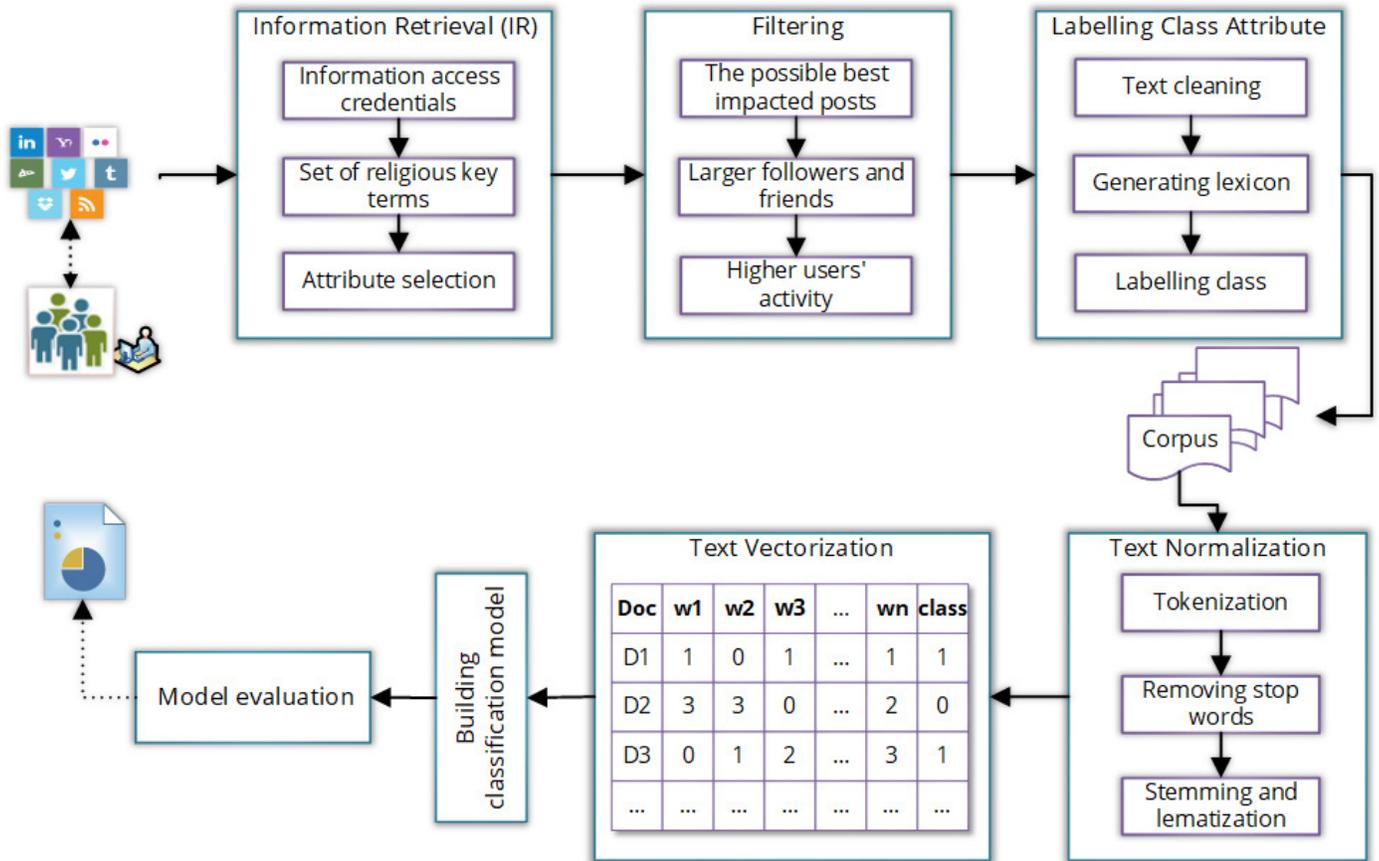


Fig. 2. Overall Architectural Diagram of the Proposed Approach.

with different performance metrics for showing how effectively the machine learning models classify religiously abusive or non-abusive contents.

IV. EXPERIMENTAL SETUP

In this paper, we focused on six major religious beliefs in worldwide¹, including Christianity (31.2%), Islam (24.1%), Secular/Nonreligious/Agnostic/Atheist (16%), Hinduism (15.1%), Buddhism (6.9%), and Judaism (0.2%). We selected Twitter as a social media platform for collecting data in our experimental purposes. We used TF-IDF, trained Word Embedding, and pre-trained GloVe model for feature extraction. For classification, we used different classifiers in the state-of-the-art. However, the overall experimental design flow is demonstrated in Table I.

A. Data Collection

We retrieved total 9,787 publicly available users' Tweet related to English language and letters using Twitter's search API² method with Python Tweepy³ library based on a predefined list of keywords as illustrated in Table II. As the Twitter API search approach returns object with a mix of root-level

attributes, so we highlighted the most fundamental attributes for our experiment. We then removed duplicates and filtered the tweets based on attributes, such as friends, followers, like count, retweets, and total tweet issued. These are demonstrated in the Algorithm 1 and Algorithm 2, respectively. However, after filtering process, we obtained 4,903 tweets for subsequent processing steps.

B. Labelling Class to Tweet Dataset

Before labelling class in filtered tweets, we cleaned the tweets to eliminate hash tags, mentions, and urls because these don't make any significant impact in detecting religious abusive tweets. We then applied ruled based and lexicon based analysis on tweet texts. In order to this, we considered various techniques including position of words, surrounding of words, contexts, parts of speech, phrases, religious slangs (e.g., bible thumper), punctuations (e.g., good!!!), and degree of modifier (e.g., very, kind of). We used VADER (Valence Aware Dictionary and sEntiment Reasoner) Lexicon Tool⁴ which is very effective in finding motifs from texts on social media platforms, especially for microblogging contexts. Using this tool, we calculated different polarity scores of words of a tweet text and then aggregated the scores. We then decided how much close the score for being abusive with a predefined threshold value. However, the Algorithm 3 demonstrates how

¹ <https://www.pewforum.org/2015/04/02/religious-projections-2010-2050/>

² <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets>

³ <https://github.com/tweepy/tweepy>

⁴ <https://github.com/cjhutto/vaderSentiment>

TABLE I. SUMMARY OF THE EXPERIMENTAL PROCEDURES

Experiment steps	Description
Dataset	Contains tweets of six Religion communities/trusts/beliefs
Total tweets retrieved	9,787
Unique tweets	5,235
Filtered tweets	4,903
Total class labels	3 labels- <ul style="list-style-type: none"> 0: non-abusive 1: abusive 2: neutral
Removed neutral class label samples	762
Total data samples for classification	4,141 samples for two class labels, i.e., <ul style="list-style-type: none"> non-abusive (0): 2074 abusive(1): 2067
Train and test split for classification	70% (2,898 samples) for training and 30% (1,243 samples) for testing
Feature extraction with classification models	<ul style="list-style-type: none"> TF-IDF: Naive Bayes, SVM, Random Forest, Logistic Regression, MLP Trained Word Embedding model: LSTM, GRU Pre-trained GloVe model: LSTM, GRU
Performance evaluation metrics for classification models	<ul style="list-style-type: none"> Accuracy F1-score Precision score Recall score Jaccard Similarity score ROC-AUC score Matthews Correlation Coefficient (MCC) Zero-one Loss

TABLE II. RELIGIOUS BELIEFS WITH RELATED KEYWORDS USED TO COLLECT TWEETS

Religious Beliefs	Keywords
Christianity	Christian, Roman Catholic, Christianity, Apocalypticism, Catholic Church, Baptism, Bible, Bishop
Islam	Quran, Islam, Muslims, Islamic State, Kurdish, Shia, Sunni, Jihad, Wahhabi, Islamphobia
Atheist	Atheist, Atheists, Atheism
Hinduism	Hindu, Bhagavad-Gita, Brahman, Mahabharata, Ma Kali, Ramayana, Durga, Saraswati, Jai Hanuman
Buddhism	Buddhism, Gautama Buddha, Bodhisattva, Buddha, Dalai Lama, Mahayana, Nirvana
Judaism	Judaism, Jews, Jew, Berit

Algorithm 1: Tweet Filtering ($Tweets_{keywords}$)

Data: $Tweets_{keywords}$: All tweets retrieved from religious keywords search approach
Result: Filtered tweets after satisfying different conditions

```

/* Initially set filtered tweet list to null */
1 filtered $Tweets \leftarrow \phi$ 
2  $unique_{tweets} \leftarrow UniqueTweets(Tweets_{keywords})$ 
3 for  $\forall tweet \in unique_{tweets}$  do
4   if  $tweet.followers \geq 100$  then
5     |  $filtered_{Tweets} = tweet$ 
6   else if  $tweet.friends \geq 100$  then
7     |  $filtered_{Tweets} = tweet$ 
8   else if  $tweet.retweet \geq 50$  then
9     |  $filtered_{Tweets} = tweet$ 
10  else if  $tweet.tweet_{like\_count} \geq 1$  then
11    |  $filtered_{Tweets} = tweet$ 
12  else if  $tweet.total_{tweet\_issued} \geq 50$  then
13    |  $filtered_{Tweets} = tweet$ 
14  else
15    |  $filtered_{Tweets} \notin unique_{tweets}$ 
16 return  $filtered_{Tweets}$ 

```

Algorithm 2: UniqueTweets ($Tweets$)

Data: $Tweets$: Set of all tweets retrieved from Tweet API search

Result: Unique tweets after removing duplicates

```

/* Initializing first tweet from Tweets */
1  $uniqueTweets \leftarrow tweet_1 \in Tweets$ 
2 for  $\forall tweet \in Tweets$  do
3   for  $\forall temp \in uniqueTweets$  do
4     if  $tweet \neq temp$  then
5       |  $uniqueTweets = tweet$ 
6 return  $uniqueTweets$ 

```

we can achieve labelling class attributes in the dataset. After including class attributes in filtered dataset, we obtained our corpus for classification.

C. Feature Extraction

In feature extraction, we transformed the texts (from corpus) into feature vector. Before doing that, we preprocessed or normalized the texts in a few consecutive ways. That is, we first removed special characters, numbers and punctuations; then expanded the contractions, lowering case, tokenization and removing stop words with our predefined stop word list; and finally mapped the words to their root form with lemmatization. For feature extraction, we used three different models, i.e., n-gram TF-IDF, trained deep learning based word embedding model, and pre-trained GloVe model. The intention of choosing these three methods is to verify whether the accuracy of the classification improves with feature extraction. However, we obtained different feature matrices for each model and then fed them in the classifiers.

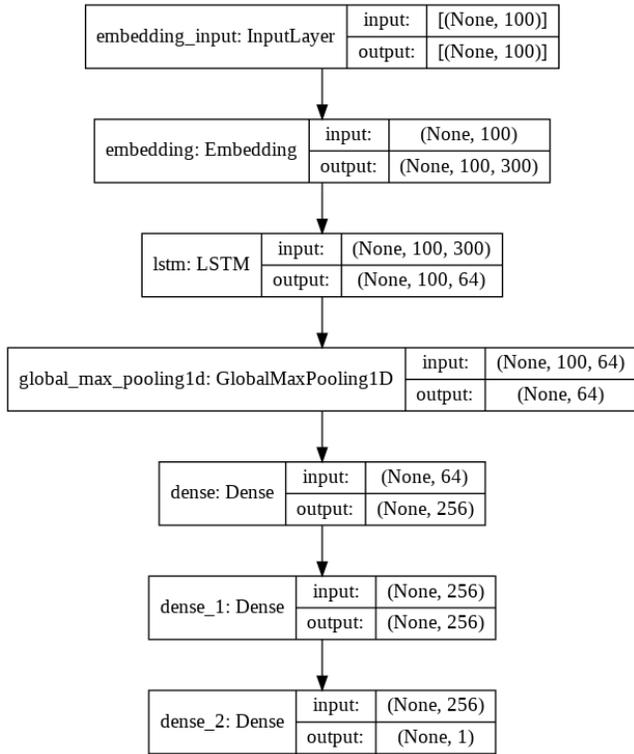


Fig. 3. Visualization of LSTM Model Configuration.

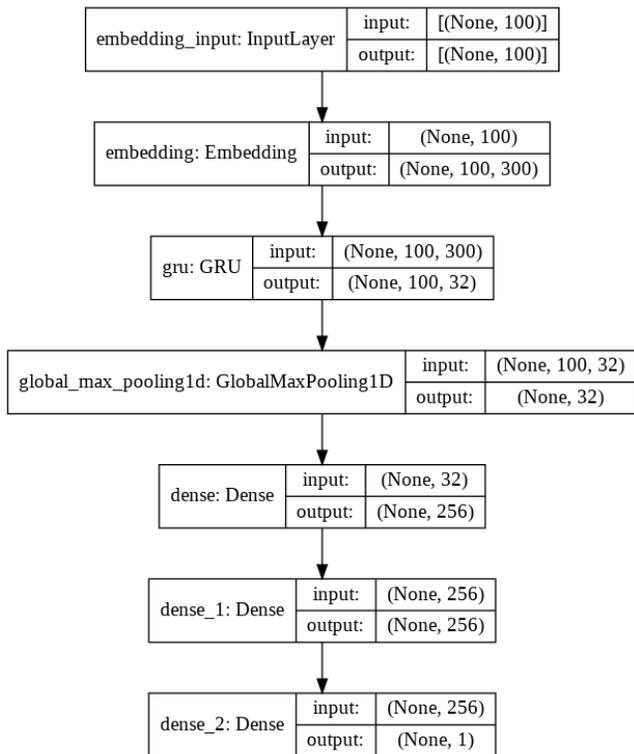


Fig. 4. Visualization of GRU Model Configuration.

Algorithm 3: Labelling Classes($Tweets_{Filtered}$)

Data: $Tweets_{Filtered}$: Set of all tweet texts after applying tweet filtering process
Result: List of class labels
 (0 : non – abusive, 1 : abusive, 2 : neutral)
 associated with each tweet

```

  /* Initialize class label list to null
  */
  1  $label_{class} \leftarrow \phi$ 
  2 for  $\forall texts \in Tweets_{Filtered}$  do
  3    $scores_{aggregated} = VADER\_Lexicon(texts)$ ;
  4   if  $scores_{aggregated} \geq 0.05$  then
  5      $label_{class} = 0$ ;
  6   else if  $scores_{aggregated} \leq 0.05$  then
  7      $label_{class} = 1$ ;
  8   else
  9      $label_{class} = 2$ ;
  10 return  $label_{class}$ 
  
```

D. Building Classifiers

As we focused on detecting religiously abusive contents from texts, so we skipped neutral class labels and considered only abusive and non-abusive classes only. That is, we considered 4,141 (out of 4,903) data samples containing abusive and non-abusive attributes for classification purposes. However, we split the dataset into train and test subsets, where 70% (2,898 samples) were for training and 30% (1,243 samples) for testing. For classification, we used Naïve Bayes, SVM, Random Forest, Logistic Regression, and MLP classifiers on TF-IDF feature matrix. On the other hand, LSTM, and GRU classification models were used on both trained and pre-trained GloVe embedding models because these classifiers are more efficient than the traditional machine learning based classifiers. We have trained all the classifiers more than 100 times with different parameters and then selected the best configuration for the final classification model building. MLP, LSTM, and GRU are neural network based classifiers that may not perform well with predefined parameters for all datasets, rather, depending upon the use case and problem statement, so we emphasized on choosing training parameters that show best performance on test data. The parameters of MLP include: input layer size=total features (i.e., more than 10,000), 3 hidden layers with neurons (125, 125, 125), optimization='adam', learning rate=0.0001, maximum iteration is between 2000-5000, and 1 output layer for deciding whether abusive or non-abusive; on the other hand, the best configuration of LSTM, and GRU networks are shown in Fig. 3 and Fig. 4, respectively. However, we then evaluated the classifiers with various performance metrics (see in Table I).

V. RESULT ANALYSIS AND DISCUSSIONS

As we focused on detecting abusive activities among major religious beliefs of worldwide in social media, so we haven't mentioned how much abusive activities are identified of a particular religion in our experiments. In this section, we have discussed the obtained results in different perspectives.

A. Exploratory Data Analysis

In section IV-A, we demonstrated how we collected tweets with a mix of fundamental attributes using different religious keywords. Among them, the hashtag attribute gives an overview of how religiously relevant datasets we have been able to collect. While retrieving the dataset, we found more than 100 unique hashtags, the most commonly used hashtags are shown in the Fig. 5. It shows that the most occurring hashtags are somehow religiously related, so it can be said that the more occurring religious hashtags, the more likely users have involved in religious tweets.

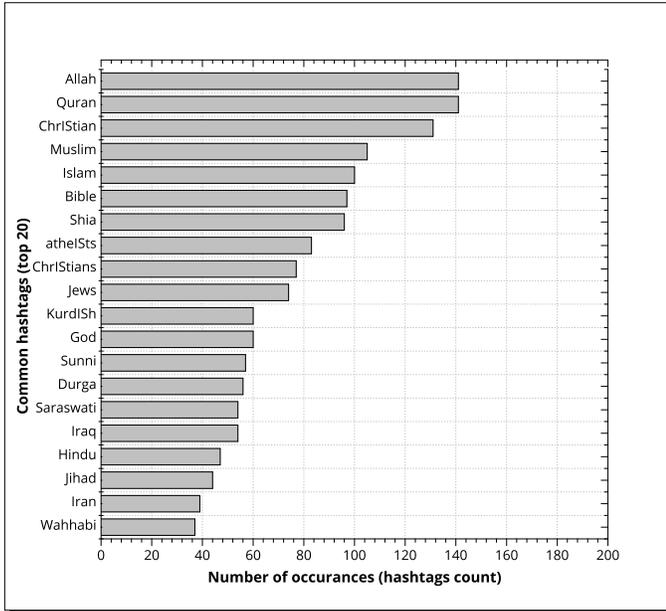


Fig. 5. Most Frequently Occurring Hashtags in the Collected Tweets.

Users in Twitter share their opinions, ideas and thoughts on diverse religious issues with a large number of audiences. So, taking these opportunities, some users may involve in the harassment of someone, sharing offensive posts or contents, spreading hate speech, and then encourage others to do so. In this paper, we collected many users' tweet (4,903 after filtering) containing both abusive and normal or neutral attitudes. To illustrate the behaviour of the users' tweet, we have shown them using the most occurring unigram, bigram and trigram frequencies Fig. 6, Fig. 7 and Fig. 8, respectively. Fig. 6 shows twenty most frequently occurring unigrams (single words) with their respective frequency counts out of 66,730 unique unigrams. Whereas Fig. 7 shows ten most occurring bigrams (2-adjacent words from a sequence of tokens) with the number of times, they appear sequentially in users' tweet out of 33,350 unique bigrams. On the other hand, Fig. 8 depicts ten most commonly occurring trigrams (3 consecutive words from a sequence of tokens) out of 22,243 unique trigrams on different religions.

We also analysed the number of religious abusers based on the location of tweet users. We found more than 1,500 specified locations that the user provided in their accounts profile. However, Fig. 9 shows ten locations where the highest percentage of abusive tweets were found (as there were 2,067 abusers detected out of 4,141 tweets). The locations indicated

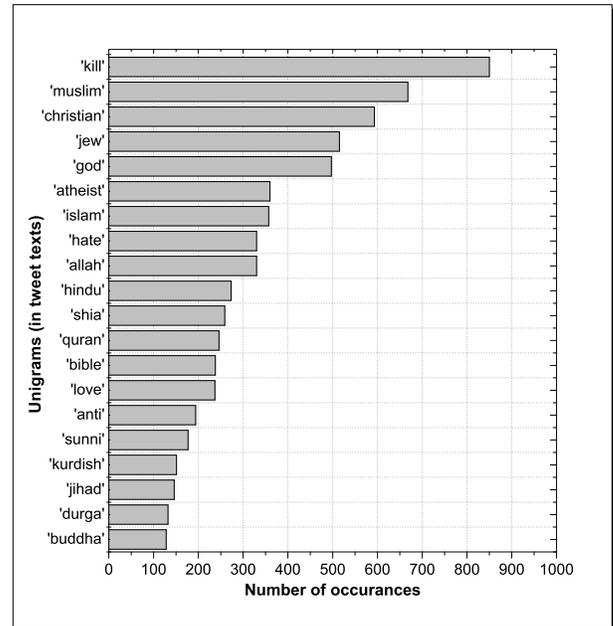


Fig. 6. Most Commonly Occurring Unigrams in users' Tweet on Different Religious Beliefs.

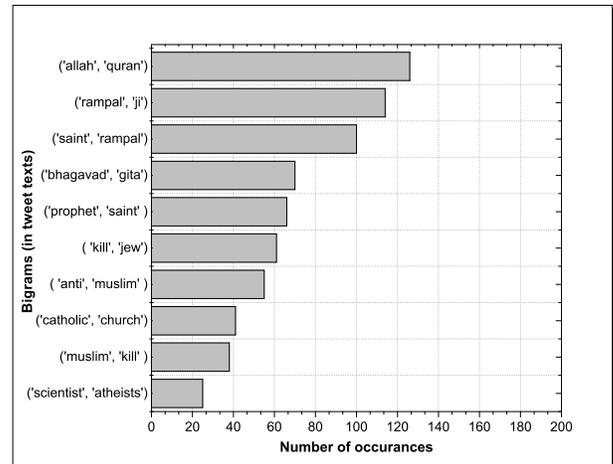


Fig. 7. Ten Most Commonly Occurring Bigrams in users' Tweet on Different Religious Beliefs.

in the figure are categorized by country to visualize the percentage of abusers.

As the text contents on Twitter is limited up to 280 characters, so many users try to attract more audiences to read or engage through posting of different lengths of tweets on it. In our analysis, we considered three different tweet lengths (in total of 4,141 tweets), i.e., '70-140' character limit, '140-280', and less than 70 character limit. These are shown in Fig. 10, where 72.52% of tweets are found within '140-280' character limit, 22% in between '70-140', and then 5.48% in less than 70 character limit. Moreover, the percentage of abusers and normal users in '140-280' length are also higher than that of other character limits. Fig. 10 also shows that the abusive tweets in '140-280' character length are higher than that of other limits, whereas normal or non-abusive tweets in '70-

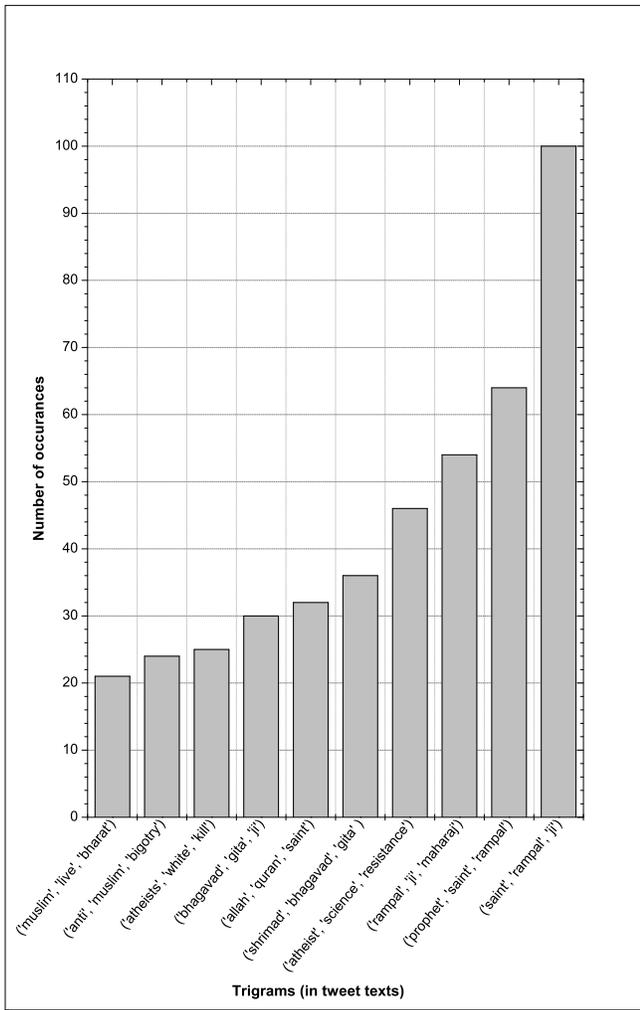


Fig. 8. Ten Most Commonly Occurring Trigrams in users' Tweet on Different Religious Beliefs.

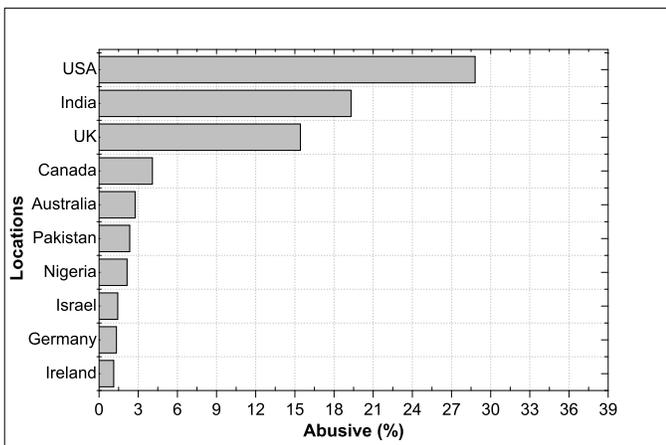


Fig. 9. The Highest Percentage of Abusers Detected in Different Countries.

140' and '<70' character limit are higher than that of abusive tweets. So, in sum, the most common length of tweets in our dataset are between 140 and 280 characters, which indicates that the users posted long tweets to express their thoughts on

religion.

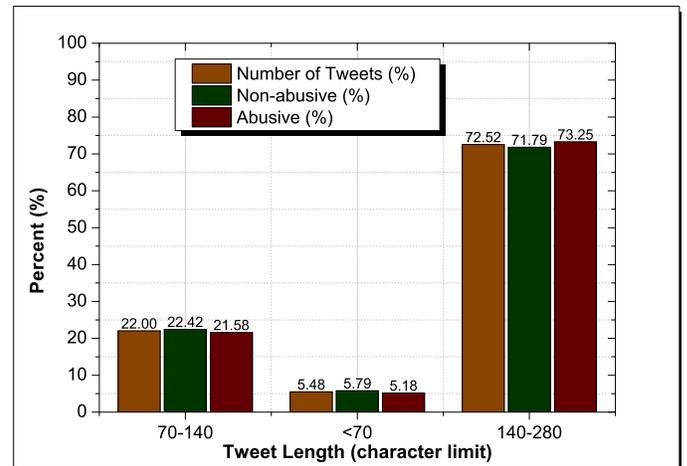


Fig. 10. Percentage users' Activities in Different Character Limits of Tweets.

B. Evaluating Classifiers

We evaluated the classification models using new unseen testing dataset. As we trained the models more than 100 times with different parameters and selected the best configuration for the final model building, so we evaluated the classifiers with the best performance on testing data (i.e., 1,253 samples with abusive 618, and non-abusive 625). Table III shows a comparative summary of the performance of the classifiers on TF-IDF feature matrix. We can see that SVM shows the best performance compared to other classification models in terms of different classification metrics (except in Jaccard Similarity score). So, the overall accuracy and loss of SVM are 83% and 17%, respectively, which is indeed promising. This means that we have been able to classify 83% correctly of abusive and non-abusive tweets. Table III also indicates that the MLP model shows quite similar performance to SVM except for (0.1-0.4)% marginal difference. However, to illustrate the performance of the models visually, the confusion matrix of each model is presented in Fig. 11. We can see that the number of false positives is higher than that of false negatives in Naïve Bayes and MLP classifiers. On the other hand, false negatives are higher than the false positives in SVM, Random Forest, and Logistic Regression.

As deep learning based feature extraction models are best suited for large features, including millions of parameters [33], so we focused on improving the classification accuracy in terms of trained and pre-trained deep learning models for more than 70,000 n-gram features in our dataset. In Table IV, the performance of LSTM and GRU classifiers are shown on trained word embedding feature extraction model. We can see that the LSTM performs better than that of GRU classifier in all classification metrics. So, we have achieved 84% overall accuracy and 16% loss on trained deep learning word embedding using LSTM. The confusion matrix of the two classifiers is depicted in Fig. 12. It shows that the number of false positives is higher than the false negatives in both LSTM and GRU models.

GloVe embedding is learned in one task and used to solve another identical task. We used this pre-trained embedding

TABLE III. CLASSIFICATION MODEL EVALUATION BASED ON TF-IDF FEATURE EXTRACTION MODEL

Classifiers	Accuracy	F1-score	Precision	Recall	Jaccard Similarity score	ROC-AUC score	MCC score	Zero-one loss
Naïve Bayes	0.798	0.798	0.802	0.799	0.678	0.799	0.601	0.201
SVM	0.832	0.832	0.832	0.832	0.708	0.832	0.665	0.167
Random Forest	0.811	0.811	0.813	0.811	0.671	0.811	0.625	0.188
Logistic Regression	0.820	0.820	0.821	0.820	0.689	0.820	0.641	0.179
MLP	0.830	0.830	0.830	0.830	0.709	0.830	0.661	0.170

TABLE IV. CLASSIFICATION MODEL EVALUATION BASED ON TRAINED DEEP LEARNING WORD EMBEDDING MODEL

Classifiers	Accuracy	F1-score	Precision	Recall	Jaccard Similarity score	ROC-AUC score	MCC score	Zero-one loss
LSTM	0.844	0.844	0.845	0.844	0.735	0.844	0.689	0.156
GRU	0.837	0.837	0.838	0.837	0.726	0.837	0.675	0.163

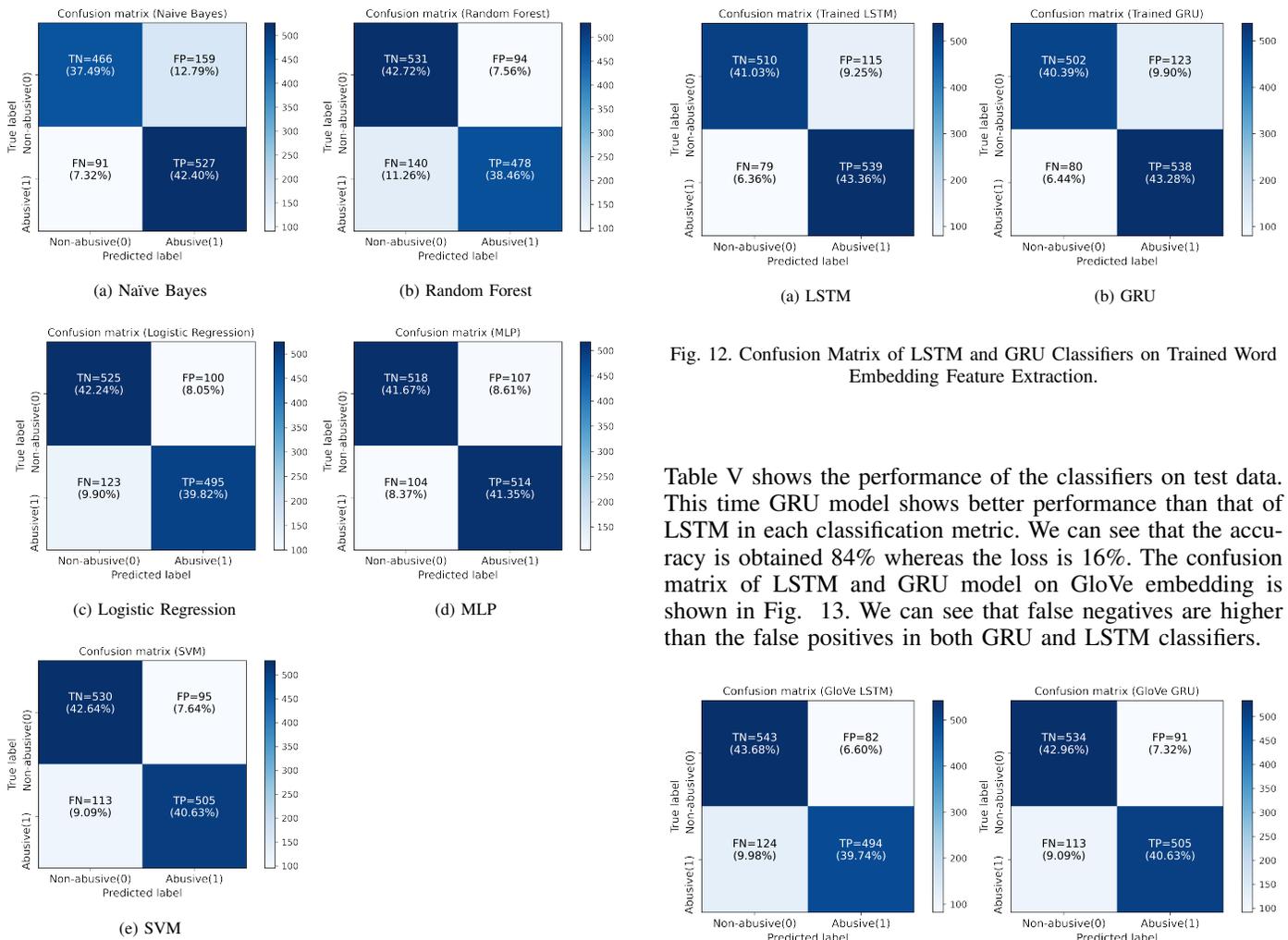


Fig. 11. Confusion Matrix of Traditional Classifiers on n-gram TF-IDF Feature Method.

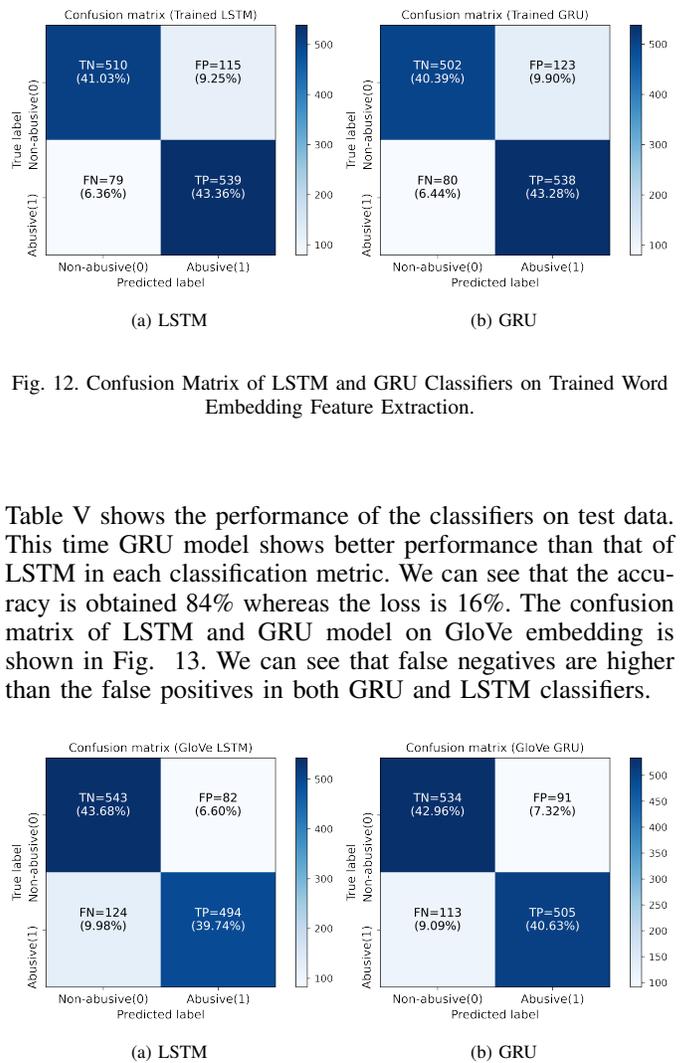


Fig. 12. Confusion Matrix of LSTM and GRU Classifiers on Trained Word Embedding Feature Extraction.

Table V shows the performance of the classifiers on test data. This time GRU model shows better performance than that of LSTM in each classification metric. We can see that the accuracy is obtained 84% whereas the loss is 16%. The confusion matrix of LSTM and GRU model on GloVe embedding is shown in Fig. 13. We can see that false negatives are higher than the false positives in both GRU and LSTM classifiers.

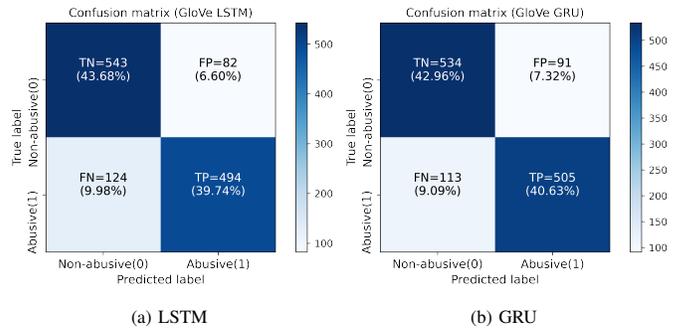


Fig. 13. Confusion Matrix of LSTM and GRU Classifiers on Pre-trained GloVe Embeddings Model.

model to create the embedding matrix on training dataset and then fed it into LSTM and GRU models for classification.

However, we have seen that SVM shows the best accuracy on TF-IDF, LSTM on trained embedding and GRU on GloVe

TABLE V. CLASSIFICATION MODEL EVALUATION BASED ON PRE-TRAINED GLOVE MODEL. THE BOLD NUMBERS REPRESENT THE BEST RESULTS

Classifiers	Accuracy	F1-score	Precision	Recall	Jaccard Similarity score	ROC-AUC score	MCC score	Zero-one loss
LSTM	0.834	0.834	0.836	0.834	0.706	0.834	0.670	0.166
GRU	0.836	0.836	0.836	0.836	0.712	0.836	0.672	0.164

TABLE VI. CLASSIFIERS PERFORMANCE COMPARISON BASED ON TF-IDF, TRAINED WORD EMBEDDING AND PRE-TRAINED GLOVE MODEL. THE BOLD NUMBERS REPRESENT THE BEST RESULTS.

Feature model extraction	Accuracy	F1-score	Precision	Recall	Jaccard Similarity score	ROC-AUC score	MCC score	Zero-one loss
TF-IDF	0.832	0.832	0.832	0.832	0.709	0.832	0.665	0.167
Trained word embedding	0.844	0.844	0.845	0.844	0.735	0.844	0.689	0.156
Pre-trained GloVe embedding	0.836	0.836	0.836	0.836	0.712	0.836	0.672	0.164

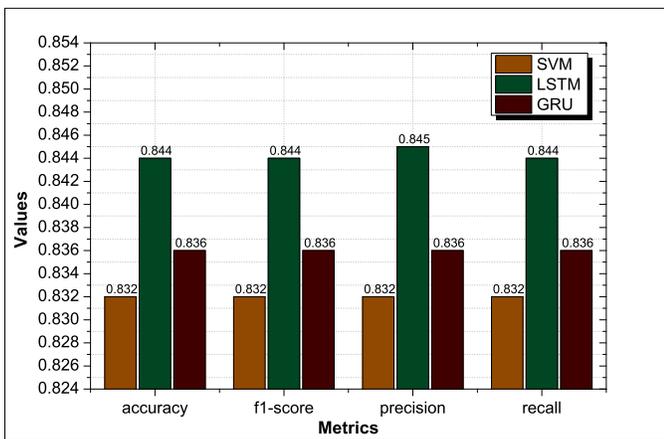


Fig. 14. Classification Performance Evaluation of SVM, LSTM, and GRU Models.

embedding. Now we compare each feature extraction method in terms of classification metrics, as presented in Table VI. The classifier on trained word embedding performs better than the other classification models. That is, LSTM model shows the best classification accuracy (84%) compared to others. The overall accuracy of the classifiers on three feature extraction methods is depicted in Fig. 14. So, if we consider precision, then 85% overall accuracy will be achieved.

VI. CONCLUSIONS

This paper focuses on identifying religiously abusive users on social media. It is the first research devoted to analyse and classify user activities regarding diverse religious beliefs and practices on any social media platform. The main contribution of this research is to establish an approach for detecting religiously abusive activities from users' social media posts. For conducting the experiment, Twitter has been selected as a social media data source. There were many users' activities (approx. 10,000) collected using a set of predefined religious keywords in English Twittersphere. Then the tweets that were redundant in nature and contained less user involvement or attraction were filtered. We then labelled the dataset using

rule based and lexicon based approaches. The labelled dataset (tweet texts) was fed into classifiers after extracting features using three different methods. The performance of the classifiers were evaluated with various classification metrics on test data. The obtained results indicate that SVM model showed 0.832 (83%) accuracy compared to others on TF-IDF, whereas LSTM gave 0.844 (84%) accuracy on trained embedding, and GRU showed 0.836 ($\approx 84\%$) accuracy on GloVe model. However, the LSTM model on trained word embedding showed 85% precision in state-of-the-art performance. Finally, we will be able to use the proposed approach for identification of any hatred/offensive speeches on any social media platform besides religious context.

Although this paper shows the identification of abusive behaviors on different religious beliefs with good accuracy, there are few constraints that were needed to be addressed. Firstly, more train and test data were desirable to take into consideration. Secondly, imbalance class distributions should be used to examine the performance of classifiers (since we used almost balanced class distributions) in experiment. However, in future, we will consider these limitations and explore on user activities of different languages. We will also identify whether the abusive contents are spread by human or social robots. In addition, we want to find the semantic similarity of hate/ non-hate speech on any religion with defined abusive key-terms in users' posts.

ACKNOWLEDGMENT

This work was partially supported by Research, Jagannath University, Dhaka, Bangladesh.

REFERENCES

- [1] M. A. Tocoglu, O. Ozturkmenoglu, and A. Alpkocak, "Emotion analysis from turkish tweets using deep neural networks," *IEEE Access*, vol. 7, pp. 183 061–183 069, 2019, doi: 10.1109/ACCESS.2019.2960113.
- [2] Y. Qi, F. Aleksandr, and F. Andrey, "I know where you are coming from: On the impact of social media sources on ai model performance (student abstract)," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 10, 2020, pp. 13 971–13 972, doi: 10.1609/aaai.v34i10.7258.

- [3] K. S. Sweet, J. K. LeBlanc, L. M. Stough, and N. W. Sweany, "Community building and knowledge sharing by individuals with disabilities using social media," *Journal of computer assisted learning*, vol. 36, no. 1, pp. 1–11, 2020, doi: 10.1111/jcal.12377.
- [4] D. L. Hoffman and T. Novak, "Why do people use social media? empirical findings and a new theoretical framework for social media goal pursuit," *Empirical Findings and a New Theoretical Framework for Social Media Goal Pursuit (January 17, 2012)*, 2012, doi: 10.2139/ssrn.1989586.
- [5] C. V. Baccarella, T. F. Wagner, J. H. Kietzmann, and I. P. McCarthy, "Social media? it's serious! understanding the dark side of social media," *European Management Journal*, vol. 36, no. 4, pp. 431–438, 2018, doi: 10.1016/j.emj.2018.07.002.
- [6] P. J. Brubaker and M. M. Haigh, "The religious facebook experience: Uses and gratifications of faith-based content," *Social Media+ Society*, vol. 3, no. 2, p. 2056305117703723, 2017, doi: 10.1177/2056305117703723.
- [7] A. Baraybar-Fernández, S. Arrufat-Martín, and R. Rubira-García, "Religion and social media: Communication strategies by the spanish episcopal conference," *Religions*, vol. 11, no. 5, p. 239, 2020, doi: 10.3390/rel11050239.
- [8] F. Lendriyono, "Public's perception on social media towards new normal during covid-19 pandemic in indonesia: Content analysis on religious social media accounts," in *IOP Conference Series: Earth and Environmental Science*, vol. 717, no. 1. IOP Publishing, 2021, p. 012039, doi:10.1088/1755-1315/717/1/012039.
- [9] T. Buchanan, "Why do people spread false information online? the effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation," *Plos one*, vol. 15, no. 10, p. e0239666, 2020, doi: 10.1371/journal.pone.0239666.
- [10] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018, doi: 10.1126/science.aap9559.
- [11] L. Peek, "Becoming muslim: The development of a religious identity," *Sociology of religion*, vol. 66, no. 3, pp. 215–242, 2005, doi: 10.2307/4153097.
- [12] T. Ahammad, M. K. Uddin, A. Karim, and S. Halder, "A framework for detecting and tracking religious abuse in social media," in *2019 International Conference on Computing, Power and Communication Technologies (GUCON)*. IEEE, 2019, pp. 206–211.
- [13] S. Alsafari, S. Sadaoui, and M. Mouhoub, "Hate and offensive speech detection on arabic social media," *Online Social Networks and Media*, vol. 19, p. 100096, 2020, doi: 10.1016/j.osnem.2020.100096.
- [14] R. Alshalan and H. Al-Khalifa, "A deep learning approach for automatic hate speech detection in the saudi twittersphere," *Applied Sciences*, vol. 10, no. 23, p. 8614, 2020, doi: 10.3390/app10238614.
- [15] Z. Mossie and J.-H. Wang, "Vulnerable community identification using hate speech detection on social media," *Information Processing & Management*, vol. 57, no. 3, p. 102087, 2020, doi:10.1016/j.ipm.2019.102087.
- [16] G. Jain, M. Sharma, and B. Agarwal, "Spam detection in social media using convolutional and long short term memory neural network," *Annals of Mathematics and Artificial Intelligence*, vol. 85, no. 1, pp. 21–44, 2019, doi:10.1007/s10472-018-9612-z.
- [17] A. Badawy and E. Ferrara, "The rise of jihadist propaganda on social networks," *Journal of Computational Social Science*, vol. 1, no. 2, pp. 453–470, 2018, doi: 10.1007/s42001-018-0015-z.
- [18] A. Al-Rawi and J. Groshek, "Jihadist propaganda on social media: An examination of isis related content on twitter," *International Journal of Cyber Warfare and Terrorism (IJCWT)*, vol. 8, no. 4, pp. 1–15, 2018, doi: 10.4018/IJCWT.2018100101.
- [19] L. Wakeford and L. Smith, "Islamic state's propaganda and social media: Dissemination, support, and resilience," in *ISIS propaganda: A full-spectrum extremist message*. Oxford University Press, 2020, pp. 155–187, doi:10.1093/oso/9780190932459.003.0006.
- [20] A. M. U. D. Khanday, Q. R. Khan, and S. T. Rabani, "Identifying propaganda from online social networks during covid-19 using machine learning techniques," *International Journal of Information Technology*, pp. 1–8, 2020, doi: 10.1007/s41870-020-00550-5.
- [21] D. A. Broniatowski, D. Kerchner, F. Farooq, X. Huang, A. M. Jamison, M. Dredze, and S. C. Quinn, "The covid-19 social media infodemic reflects uncertainty and state-sponsored propaganda," *arXiv preprint arXiv:2007.09682*, 2020.
- [22] O. D. Apuke and B. Omar, "Fake news and covid-19: modelling the predictors of fake news sharing among social media users," *Telematics and Informatics*, vol. 56, p. 101475, 2021, doi: 10.1016/j.tele.2020.101475.
- [23] H. Nguyen and A. Nguyen, "Covid-19 misinformation and the social (media) amplification of risk: A vietnamese perspective," *Media and Communication*, vol. 8, no. 2, pp. 444–447, 2020, doi: 10.17645/mac.v8i2.3227.
- [24] D. D. Chaudhari and A. V. Pawar, "Propaganda analysis in social media: a bibliometric review," *Information Discovery and Delivery*, 2021, doi: 10.1108/IDD-06-2020-0065.
- [25] Y. K. Dwivedi, G. Kelly, M. Janssen, N. P. Rana, E. L. Slade, and M. Clement, "Social media: The good, the bad, and the ugly," *Information Systems Frontiers*, vol. 20, no. 3, pp. 419–423, 2018, doi: 10.1007/s10796-018-9848-5.
- [26] A. Whiting and D. Williams, "Why people use social media: a uses and gratifications approach," *Qualitative Market Research: An International Journal*, 2013, doi: 10.1108/QMR-06-2013-0041.
- [27] M. S. Kgate, "Social media and religion: Missiological perspective on the link between facebook and the emergence of prophetic churches in southern africa," *Verbum et Ecclesia*, vol. 39, no. 1, pp. 1–6, 2018, doi: 10.4102/ve.v39i1.1848.
- [28] N. Albadi, M. Kurdi, and S. Mishra, "Investigating the effect of combining gru neural networks with handcrafted features for religious hatred detection on arabic twitter space," *Social Network Analysis and Mining*, vol. 9, no. 1, pp. 1–19, 2019, doi: 10.1007/s13278-019-0587-5.
- [29] B. Vidgen and T. Yasseri, "Detecting weak and strong islamophobic hate speech on social media," *Journal of Information Technology & Politics*, vol. 17, no. 1, pp. 66–78, 2020, doi: 10.1080/19331681.2019.1702607.
- [30] N. Sun, G. Lin, J. Qiu, and P. Rimba, "Near real-time twitter spam detection with machine learning techniques," *International Journal of Computers and Applications*, pp. 1–11, 2020, doi: 10.1080/1206212X.2020.1751387.
- [31] A. B. Boot, E. T. K. Sang, K. Dijkstra, and R. A. Zwaan, "How character limit affects language usage in tweets," *Palgrave Communications*, vol. 5, no. 1, pp. 1–13, 2019, doi: 10.1057/s41599-019-0280-3.
- [32] W. Herzallah, H. Faris, and O. Adwan, "Feature engineering for detecting spammers on twitter: Modelling and analysis," *Journal of Information Science*, vol. 44, no. 2, pp. 230–247, 2018, doi: 10.1177/0165551516684296.
- [33] H. Liang, X. Sun, Y. Sun, and Y. Gao, "Text feature extraction based on deep learning: a review," *EURASIP journal on wireless communications and networking*, vol. 2017, no. 1, pp. 1–12, 2017, doi: 10.1186/s13638-017-0993-1.

Fish Disease Detection System: A Case Study of Freshwater Fishes of Bangladesh

Juel Sikder¹, Kamrul Islam Sarek², Utpol Kanti Das³
Dept. of Computer Science and Engineering
Rangamati Science and Technology University
Rangamati, Bangladesh

Abstract—The proposed system is designed for automatic detection and classification of fish diseases in freshwater especially Rangamati Kaptai Lake and Sunamganj Hoar area of Bangladesh. Our experimental result is indicating that the proposed approach is significantly an accurate and automatic detection and recognition of fish diseases. This study presents fish disease detection based on the K-means and C-means fuzzy logic clustering method to segment the filtering image. Gabor's Filters and Gray Level Co-occurrence Matrix (GLCM) are used to extract the features from the segmented regions. Finally Multi-Support Vector Machine (M-SVMs) is used for classification of the test image. The proposed system demonstrated a comparison between K-means clustering and C-means fuzzy logic. The proposed methodology gave 96.48% accuracy using K-means and 97.90% using C-means fuzzy logic which is the highest accuracy rate to compare other existing methods. The proposed system has been experimented in the MATLAB environment on infected fish images of Rangamati Kaptai Lake and Sumangan Hoar area. It is a challenging task of fisheries farming in Hoar areas and Lake areas to detect fish diseases initially. The proposed methodology can detect and classify different fish diseases in early stages and also contributes to improved results for fish disease detection.

Keywords—K-means; c-means fuzzy logic; multi-SVM; detection

I. INTRODUCTION

Fisheries play a vital role in social and economic development and deserve the potential for future development in Bangladesh's agricultural economy. This sector represents one of the most prolific and energetic sectors in Bangladesh. It subsidizes 3.65% of the national GDP and approximately one-fourth (23.81%) of the GDP. Bangladesh is considered as one of the most suitable countries in the world for freshwater rural fisheries because of its resources and agro-climatic condition [1]. Fisheries and aquaculture paid 25% of agricultural GDP and 4.4% of national GDP in 2012 [1]. Fish biodiversity of the Rangamati Kaptai Lake and the Sunamganj Hoar area is decreasing day by day, due to territory deprivation and different man-made causes. It also worsens the normal health condition of fish and causes death of the fishes which results in great economic loss. Early detection of infected fish is a necessary step in preventing the spread of the disease. This paper recognizes and identifies the different fish diseases of Rangamati Kaptai Lake and Sunamganj Hoar which impacts the control of disease and maintenance of a strong relationship between living creatures, and the environment of Rangamati and Sunamganj. The prosperity of a fish culture depends on its precise regulation, fish life time, and cultural environment. However, unexpected use of feed and fertilizer are the main

reasons of failures in fish production. Traditional and semi-intensive culture systems are also responsible for the fish disease. Therefore, the search for a more efficient, less expensive, and more accurate method of diagnosing fish disease cases is very important [2]. In this research, we conducted the identification of local freshwater fish using a software based approach. This research also approaches a method for different fish disease detection. Firstly, the pre-processing stage applied on the test image. In the pre-processing stage, the test image is resized [3] then RGB images converted into HIS. The noise filtering technique is also used to remove noise. In the second stage, K-means and C-means fuzzy logic clustering method applied to segment the filtering image. In the third stage, the system extracts the features from the segmented and infected regions using Gabor's Filters and GLCM. In the fourth stage, classification of the test image is done by Multi-SVMs algorithm. The proposed system is tested using different types of fish diseases, especially Tail and Fin Rot, EUS, Red spot, Bacterial gill rot, Parasitic Disease (Argulus), Broken Antennae, and Rostrum. Finally, fish diseases are detected and recognized that can be prescribed in its early stages [4].

This research is designed in six sections. Section II represents some related works and Section III describes the data collection process. Section IV gives a detailed description of the proposed methodology. Section V briefing the experimental results and discussing about the mentioned method and outcomes of the system. Section VI presents conclusions and the opportunity of probable coming work in this area.

II. RELATED WORKS

Some related works associated with the proposed method described here. Waleed, Ahmed, et al. [5] completed a survey on automatic recognition of different fish diseases. They proposed three several categories of fish diseases named Columnaries, Ichthyophthirius (Ich) and Epizootic ulcerative syndrome (EUS) to recognize and identify automatically. They show the effect of several color spaces on CNN ultimate execution using Alexnet. Chacon, Mario I., Luis Aguilar, and Abdi Delgado [6] proposed a method related to an optional defuzzification function, fuzzy operators and an image fuzzification function. The Contextuality of the system is described in three phases, geometric measurements, edge detection and image binarization. On binarization the fuzzy method describes the way to solve indeterminacy on the classification of pixels. The fuzzy edge detection showed how a particular perception could be demonstrated by a fuzzy system. Lastly, fuzzy geometry is an illustration of how to obtain a variable state

below of lighting to gather more consistent measurements. Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar [7] proposed a method that applies GLCM to extract second order statistical texture features. The Four features namely, Inverse Difference Moment, Correlation, Angular Second Moment and Entropy are computed using Xilinx FPGA. The image compression time can be diminished significantly by extracting the features of an image. These features are important in the motion approximation of videos and in real time pattern identification applications. Li, Wei, and Qian Du [8] proposed a system that showed the importance of extracted features using a simple Gabor filter for the NRS classifier. The accuracy of the experimented classifier results compared to conventional pixel-wise classifiers. Specifically, to extract spatial features in the PCA projected domain or a subset of original bands with band selection, a simple two-dimensional Gabor filter was implemented. Traditional classifiers, such as SVM, RLDA and SRC compared with applied classification techniques, i.e., PC-Gabor-NRS and BS-Gabor-NRS. Authors claimed that Experimental results can outperform other classification techniques. Lyubchenko, Valentin, et al. [9] proposed a color based image segmentation by determining the infected areas so that it can detect and diagnose fish diseases. To detect infected and healthy areas of the fish body, they used digital image processing and colored markers. Malik, Shaveta, Tapas Kumar, and A. K. Sahoo [4] proposed a method to segment and classify EUS and Non-EUS infected fish diseases. To enhance the image, they applied segmentation. Again, they collect useful information using different edge detection methods and also applied morphological operation. They used various feature descriptors to extract features to classify the fish disease using PCA. Sankar, M. Muni, et al. [10] proposed an approach to investigate diseased fish using image processing techniques. They applied capturing image sensing techniques for image acquisition and k-means clustering method. They also used edge detection techniques to detect edges and K-means clustering algorithm to cluster the input prawn image and detect diseases.

III. DATA COLLECTION

During the study of the historical 20 endangered fish species, IUCN has been initiating in Kaptai Lake. Among 20 species 16 were found to be available, 4 rarely available, and 34 were not available during the study period. Accessibility of three categories of disapprovingly endangered, threatened, and helpless fishes from Kaptai Lake are presented in considering exposed species, 30% were found accessible, 7% rarely available and 63% were not available during the study period. Available fish landing hubs and marketplaces would be helpful to provide primary knowledge of fish fauna, mostly to know about the fish fauna in the biggest artificial Kaptai Lake, Rangamati, Bangladesh.

In this study, Kaptai reservoir and fishery Ghat in Rangamati and Sunamganj Haor were selected to collect the different types of freshwater fishes. The data were also collected through the interview with fishermen, associated persons and cross-check interviews with informers. After that, the collected data were cross-checked with the landing data from Bangladesh Fisheries Development Corporation fish landing center especially Rangamati and Sunamganj, Bangladesh. Several large diverse fisheries are found in Rangamati town and Sunamganj

Community Based Resource Management Project (CBRMP). Different types of diseases and symptoms are described in Table I. After collecting different types of fishes from Rangamati Kaptai Lake and Sunamganj Haor, we created a database which was uploaded to kaggle.com [11]. We used different types of fish diseases and symptoms of the diseases in our database. In our study, we also used 126 fish species from 39 families initiated in the Sunamganj Haor area [9].

TABLE I. AVAILABILITY OF FRESHWATER FISH SPECIES KAPTAL LAKE AND SUNAMGANJ HAOR

Family	Lacal name	Common name	Scientific name
Schilbeidae	Bacha	Bacha	Eutropiichthys Vacha
Siluridae	Kani Pabda	pabo catfish	Ompok Bimaculatus
Cyprinidae	Mrigel, Mirka	Caurvery white carp	Cirrhinus Cirrhosis
Nandidae	Bheda	Mud perch	Nandus Nandus
Gobiidae	Bele	Tank goby	Glossogobius Giuris
Belonidae	Kaikka	Needle fish	Xenentodon Cancila
Sciaenidae	Poa	Pama	Otolithoides Pama
Clariidae	Magur	Air breathing catfish	Clarias Batrachus
Pangasidae	Pangas	Yellowtail catfish	Pangasius Pangasius
Cyprinidae	Rui	Rohu	Labeo Rohita
Notopterid	Chital	Humped featherback	Notopterus Chitala

IV. PROPOSED METHODOLOGY

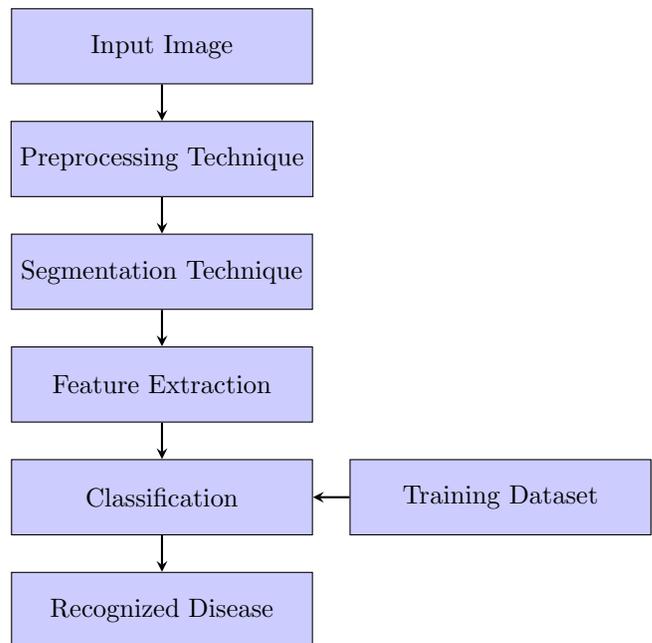


Fig. 1. Block Diagram of the Developed System.

“Fig. 1” describes the proposed block diagram.

A. Input Image

The system required a color image to start the procedure. An affected fish image passed as an input to the system.

B. Pre-processing

Noise Filtering is applied to remove several forms of noises and to filter the redundant information from the images [12]. Contrast Stretching is compulsory when the system needs to intensify or diminish the total number of pixels. Pre-processed results of the input images are shown in “Fig. 2”.

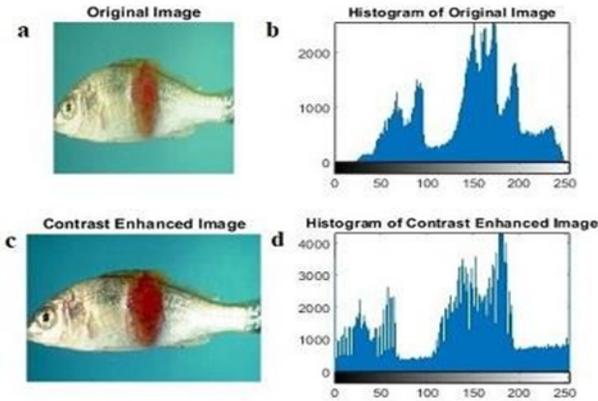


Fig. 2. (a) Input of the EUS Affected Original Fish Image (b) Histogram Result (c) Enhanced Result and (d) Histogram of Enhanced Image.

C. Image Segmentation

Fuzzy C-means (FCM) is a clustering method used to separate data into two or more clusters. We used FCM in our system. Firstly, it picks two initial cluster values from the input image, $I (r \times c)$. The system accesses every pixel as $I (i, j)$, where

$$0 \leq j < c \quad (1)$$

and

$$0 \leq i < r \quad (2)$$

Secondly, the average of median values of each row is calculated. Thirdly, the distance between every pixel of the image and cluster values calculated. Again, the distance between the pixels of the image and averaged medians are measured also. Fourthly, the final fuzzy field smooths the data obtained from cluster values and from averaged medians. Fifthly, the fuzzy rule is applied to segment the image between two clusters. Sixthly, the segmented region is reconstructed by updating cluster centers. Seventhly, steps third to sixth repeated for t times. The whole process is iteratively performed until the difference between two consecutive clusters reaches to a constant position and becomes minimum [13]. Finally, the segmented regions of both clusters for a test image are found.

K-means clustering is an unsupervised learning algorithm that is used enormously. The method is used to define k samples, one for each cluster [14]. The next step is to assign each point of input data into the nearest cluster. After selecting each data point, the first step is completed by re-calculating the new cluster centroid for the k -clusters. An iterative process has been generated until the new centroid reaches the constant state. As a result, the data points separate into k clusters step by step.

D. Feature Extraction

After segmentation, when the area of the interest region is detected then the significant features are extracted. GLCM and Gabor Filter are used to extract feature vectors from input fish disease image like texture. Texture element is extracted from the RGB colored image. The GLCM functions represent texture features of an image. Textures are the features whose focus on the distinct pixels that create an image. Texture features such as statistical feature is used which contains GLCM, grey level histogram, run length matrices and autocorrelation features for texture extraction. GLCM procedures extract 2nd-order statistical texture features [7]. Textural features include entropy, correlation, contrast, homogeneity and energy. The texture is particularly appropriate for this kind of study because of its possessions. The system also used Gabor filter features. The Gabor filters consist of parameters such as standard deviation, orientation and the radial center frequency [8]. The system combined the GLCM feature and the Gabor Filters into one feature set. This combination of the Gabor filters and the GLCM feature generate a better outcome on the fish texture dataset.

E. Classification Using Multi-SVMs

The proposed system used multi-SVMs to classify fish diseases. The extracted features of each fish disease image are stocked in the database. The SVM classifier will calculate the feature number of database images and the feature value of the input image, based on these values the classifier will separate the input image from that category. SVM will compare the test sample feature set to all training samples and select the shortest distance. Support vector machines were initially calculated for binary classification. Several approaches have been proposed where typically the system constructs a multi-class classifier by combining several binary classifiers. The SVM is learned by features given as an input to its training procedure. During training, the SVM identifies the appropriate boundaries in the 2-categories. Features are called according to class associative to a specific class. Using received features, the multi-SVM starts classification steps. The system uses multi-SVM as a two class SVM repeatedly. When the test image is passed to the classifier, it classifies it from the first two classes. After that classified class and a new class used to classify the test image. Similarly, it continued its classification steps until all categories were tested. From this last classification, the test image is classified as the true categorized class [14].

V. RESULT AND DISCUSSION

The proposed method is examined in MATLAB R2018a. The experimented database was uploaded to kaggle.com [11]. The database contains fish disease images of freshwater of Rangamati Kaptai Lake and Sunamganj Hoar area of Bangladesh. The system experimented using different fish diseases such as EUS, Red Spot, Argulus, Tail and Fin Rot, Broken antennae rostrum and Bacterial Gill rot. The system divided the input image into three clusters in both C-means fuzzy logic and K-means. All clusters of both C-means fuzzy logic and K-means clustering are given in “Fig. 3”.

From all of the clusters, one cluster chosen as the segmented cluster based the segmented region. All cluster regions

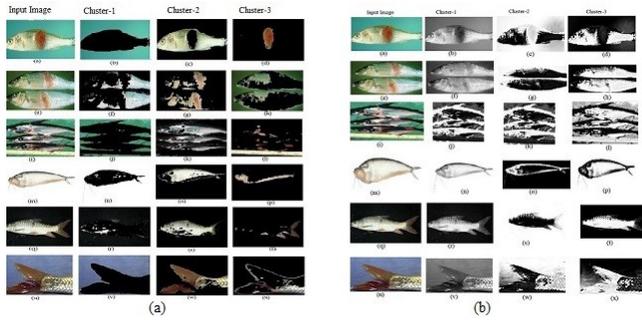


Fig. 3. a. C-means Clusters and b. K-means Clusters.

are calculated and the best cluster region chosen by user input. The segmented output is given in Fig. 4. In “Fig. 4”, the left side portion i.e., (a) represents the segmented image using C-means and the right side portion i.e., (b) represents the segmented images using K-means method.

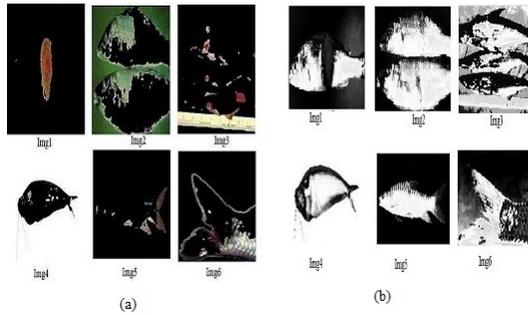


Fig. 4. Segmented Output (a. C-means and b. K-means).

TABLE II. GLCM FEATURE EXTRACTED FROM THE FISH IMAGES

Feature Name	Img1	Img2	Img3	Img4	Img5	Img6
Contrast	0.0473	0.0002	0.0001	0.0000	0.0001	0.0000
Correlation	0.0569	0.0002	0.0002	0.0001	0.0001	0.0001
Energy	0.0280	0.0001	0.0002	0.0000	0.0000	0.0000
Homogeneity	0.0577	0.0002	0.0002	0.0001	0.0001	0.0001
Mean	0.1127	0.0004	0.0005	0.0001	0.0002	0.0001
Standard Deviation	0.9833	0.0010	0.0010	0.0010	0.0010	0.0010
Entropy	0.9800	0.0010	0.0010	0.0010	0.0010	0.0010
RMS	0.9901	0.0010	0.0010	0.0010	0.0010	0.0010
Variance	0.9797	0.0010	0.0010	0.0009	0.0010	0.0010
smoothness	0.9604	0.0009	0.0009	0.0007	0.0010	0.0010
Kurtosis	0.2382	0.0001	0.0001	0.0007	0.0005	0.0002
Skewness	0.2339	0.0001	0.0001	0.0007	0.0005	0.0002
IDM	3.0432	0.0017	0.0023	0.0132	0.0042	0.0045

Table II illustrates the extracted feature of segmented images. The feature extracted using GLCM feature extractor. Combining those GLCM features and Gabor filtered features, passed to the classifier to create the desired classifier.

The results of the Multi-SVM classifier are shown in Table III. The performance of the classifier is calculated by the different procedures and The SVM classifier calculates the feature value of the database and of the test image. The SVM classifier uses that feature values to identify the test image class. Input feature is (1.0e+03) and maximum accuracy of EUS and Broken antennae and rostrum is (98.4127), the

TABLE III. MULTI-SVM CLASSIFIER RESULTS FOR THE FISH DISEASE IMAGES

S/N	Fish Image	Multi-SVM Classifier	Input feature	Maximum Accuracy	Row
1	Img1	EUS	1.0e+03	98.4127	500 iterations
2	Img2	Red spot	1.0e+03	95.2381	500 iterations
3	Img3	Broken antennae and rostrum	1.0e+03	98.4127	500 iterations
4	Img4	Argulus	1.0e+03	95.2381	500 iterations
5	Img5	Bacterial Gill rot	1.0e+03	96.8254	500 iterations
6	Img6	Tail and fin rot	1.0e+03	95.2381	500 iterations

accuracy of Linear Kernel is the highest 500 iteration.

TABLE IV. A COMPARISON BETWEEN K-MEANS CLUSTERING AND C-MEANS FUZZY LOGIC

S/N	Image	Fish disease	Accuracy of K-means	Number of Iteration	Accuracy of C-means
1	Img1	EUS	98.4127	500 iterations	99.25
2	Img2	Red Spot	95.2381	500 iterations	96.50
3	Img3	Broken antennae	98.4127	500 iterations	99.45
4	Img4	Argulus	95.2575	500 iterations	97.23
5	Img5	Bacterial Gill rot	96.8254	500 iterations	98.70
6	Img6	Tail and fin rot	94.7450	500 iterations	96.25

Table IV represents the proportional outcomes of the performance of between K-means clustering and C-means fuzzy logic segmentation. The system gives 96.48% accuracy using C-means fuzzy logic where K-means gives 97.90% accuracy.

VI. CONCLUSION

The paper proposed an image processing based approach and automatically detects different types of fish diseases. The testing of the proposed system has been experienced on the actual infected fish images collected from Rangamati Kaptai Lake and Sunamganj Hoar, Bangladesh. The experimental results indicate that the proposed approach is a valuable approach and can significantly support accurate and automatic detection of fish disease and the proposed combination gave better accuracy after applying the machine learning algorithm. The infected fish images were then classified with an average accuracy rate of K-means clustering and C-means fuzzy logic gives 96.48% and 97.90% respectively. In the future, it will present some guidelines such as establishing a better segmentation system, choosing improved classification algorithms and deep learning algorithms that can apply on different feature descriptors.

REFERENCES

- [1] FRSS. "Yearbook of Fisheries Statistics of Bangladesh." (2017).
- [2] Austin, Brian. "Methods for the diagnosis of bacterial fish diseases." Marine Life Science & Technology 1, no. 1 (2019): 41-49.
- [3] Juel Sikder, Utpol Kanti Das and Rana Jyoti Chakma, "Supervised Learning-based Cancer Detection" International Journal of Advanced Computer Science and Applications (IJACSA), 12(5), 2021. <http://dx.doi.org/10.14569/IJACSA.2021.01205101>
- [4] Malik, Shaveta, Tapas Kumar, and A. K. Sahoo. "Image processing techniques for identification of fish disease." In 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), pp. 55-59. IEEE, 2017.
- [5] Waleed, Ahmed, Hadeer Medhat, Mariam Esmail, Kareem Osama, Radwa Samy, and Taraggy M. Ghanim. "Automatic Recognition of Fish Diseases in Fish Farms." In 2019 14th International Conference on Computer Engineering and Systems (ICCES), pp. 201-206. IEEE, 2019.
- [6] Chacon, Mario I., Luis Aguilar, and Abdi Delgado. "Definition and applications of a fuzzy image processing scheme." In Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop., pp. 102-107. IEEE, 2002.

- [7] Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar. "Image texture feature extraction using GLCM approach." *International journal of scientific and research publications* 3, no. 5 (2013): 1-5.
- [8] Li, Wei, and Qian Du. "Gabor filtering based nearest regularized subspace for hyperspectral image classification." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, no. 4 (2014): 1012-1022.
- [9] Lyubchenko, Valentin, Rami Matarneh, Oleg Kobylin, and Vyacheslav Lyashenko. "Digital image processing techniques for detection and diagnosis of fish diseases." (2016).
- [10] Sankar, M. Muni, CH Nageswar Rao, G. Sailaja, N. Bhuvaneshwary, and P. Gunasekhar. "White spot syndrome virus detection in shrimp images using image segmentation techniques." *International Journal of Advanced Research in Computer Science and Software Engineering* 3, no. 9 (2013).
[11] <https://www.kaggle.com/Futpoldas/freshwaterfishdisease>.
- [12] Mahmud, Tanjim, Juel Sikder, Rana Jyoti Chakma, and Jannat Fardoush. "Fabric Defect Detection System." In *International Conference on Intelligent Computing & Optimization*, pp. 788-800. Springer, Cham, 2020.
- [13] Christ, MC Jobin, and R. M. S. Parvathi. "Fuzzy c-means algorithm for medical image segmentation." In *2011 3rd International Conference on Electronics Computer Technology*, vol. 4, pp. 33-36. IEEE, 2011.
- [14] Sikder, Juel, Utpol Kanti Das, and AM Shahed Anwar. "Cancer Cell Segmentation Based on Unsupervised Clustering and Deep Learning." In *International Conference on Intelligent Computing & Optimization*, pp. 607-620. Springer, Cham, 2020.

Improving Data Services of Mobile Cloud Storage with Support for Large Data Objects using OpenStack Swift

Aslam B Nandyal¹, Mohammed Rafi², M Siddappa³, Babu B. Sathish⁴

Dept. of Computer Science & Engg, University BDT College of Engineering, Davanagere, Karnataka, India^{1,2}

Dept. of Computer Science & Engg, Sri Siddhartha Institute of Technology Maralur Tumkur, Karnataka, India³

Department of Computer Science & Engg, R.V. College of Engineering Bengaluru, 560059, India⁴

Abstract—Providing data services support for large file upload and download is increasingly vital for mobile cloud storage. There is an increase in mobile users whose data access trends show more access and large file sharing. It is a challenging task for Mobile Application Developers to handle upload and retrieve large files to/from a mobile app because of difficulties with latency, bandwidth, speed, errors, and disruptions to service in a wireless mobile environment. Some scenarios require these large files to be used offline, sometimes to be updated by a single user, and sometimes be shared among all other users. The Wireless mobile environment must consider mobile user's constraints, such as frequent disconnections and low bandwidth, which affect the ability to handle data and transactions management. The primary objective of this study is to propose a cloud-based Mobile Sync service (sometimes referred as Mobile backend as a Service) with OpenStack Swift object storage to manage large objects efficiently using two main techniques of segmentation and object chunking with compression in a mobile cloud environment. This work further contributed to a prototype implementation of the proposed framework and provides Application Programming Interface (API) consisting of Create, Read, and Delete queries and chunking operations and a lightweight sync protocol that can manage large file synchronization and access. The experimental findings with object-chunking tested size settings show that the proposed Mobile Sync framework can accommodate large files ranging from 100MB to 1GB and provides a decrease in upload/download synchronization times of 63.203% / 92.987% percent as compared to other frameworks.

Keywords—Mobile cloud computing; mobile backend as a service; large files; distributed systems

I. INTRODUCTION

The development of enterprise mobile services can use variety of cloud computing models. The predominant service models for cloud computing are mainly classified as Infrastructure as a Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS) [1]. With the introduction of a new service model known as Mobile Backend as a Service (MBaaS), making it easy to integrate cloud-based applications with mobile platforms has become possible.

The model of Sync framework offers a cloud server infrastructure to store application data and facilitate easy configuration. The developer needs to do significant work for the application to remain responsive during interruptions in communication due to poor or no network. The Sync framework offers a solution for the unreliable connection problem with customized synchronization and replication processes and

helps synchronize with multiple clients. An intelligent Sync framework allows enterprise data to take offline and facilitate sync operation by syncing data across multiple mobile devices with the backend systems, detect and resolve the conflicts with configurable, standards-based rules, setting precedence based on policies [2] [3]. Ideally, the Sync framework should provide a consistent state at all times (strong consistency). However, the CAP theorem [4] for the distributed systems enforces the Sync framework to guarantee immediate availability and tolerate network partitions to provide a weak form of consistency, commonly known as eventual consistency [5].

Data Services is one of the critical capabilities of an MBaaS platform and provide the following features:

- **Data Management:** A quality MBaaS framework will provide the services to create, save, manage and sync application data and files within the framework itself, in addition to a mechanism for connecting into both public and private systems of record and abstraction layers [3].
- **Online/Offline Workflow:** The mobile apps can operate in online and offline modes, and hence MBaaS can support offline/online database synchronization [6].
- **Sync:** An intelligent MBaaS framework allows enterprise data to take offline, syncing data across multiple devices with the backend systems, detect and resolve the conflicts with configurable, standards-based rules, setting precedence based on policies [2].
- **Caching:** Various caching methods are offered as a part of the MBaaS platforms to reduce latency and boost app performance. The caching can be provided as an inbuilt feature in the framework or at a connection point to systems of record or even cross-platform on-device caching strategies (via client SDKs) [7].

Due to the rising numbers of people making their files shareware accessible via mobile devices, the reliability and efficiency of service for large files is a more crucial feature in mobile file-sharing than previously thought [8] [9]. Developers have to deal with complex issues of latency, speed, timeouts, and interruptions during the uploading and retrieving from mobile apps [10] [11] [12].

Recently, some efforts are focused on finding out the trend

of mobile users to access data on both public and private cloud storage services and several different personal cloud storage systems. One such study in [9] aimed to analyze data access trends in large-scale mobile cloud storage [9].

This research aimed to analyze the database of 350 million HTTP transaction log files from mobile applications to find the trend in how quickly cloud storage is being used for collaborative and large file storage. This study concluded that for retrieving one file in multiple sessions, the average volume was about 70 MB.

An empirical study was conducted by a cloud storage vendor [8], which involved the services of media file uploading, transformations, and storage in the cloud. This research studied a data set of one million mobile applications rendering these services. Their observation examined the statistical information regarding the number of files, including different file sizes and formats, which have been uploaded (from the year 2015 to 2016). Their research showed that the amount of growth files with size 100MB and above is 170%, while files of other sizes increased by 50% year over year. According to this study, it appears that the file sizes are increasing as mobile users make more frequent use of the files or share files and have larger storage requirements (above 100MB).

Some of the mobile operating systems limit the size of the file over which over-the-air (OTA) or app-store downloads are not allowed [13]. For example, Apple's iOS platform [14] limits the Cellular Data downloads to a file size of 100 MB. Android OS limits the size of downloads via cellular data to 150MB [15]. Based on the above studies and mobile operating system guidelines, it can be concluded that a file with a size greater than 100MB is considered a large file.

The rest of this paper is structured in the following manner: Next Section II provides a brief background and problem formulation. Section III analyzes some of the Mobile Sync frameworks in the literature along with support for large objects. Section IV provides the background information of the OpenStack cloud platform and support for large file storage in the Swift module that handles object storage. Section V describes the details of the proposed framework for data services to handle large files. The data management at both mobile Client and Cloud server-side is discussed. Section VI discusses the performance of the proposed Mobile Sync framework, followed by a conclusion and future work in Section VII.

II. BACKGROUND AND PROBLEM FORMULATION

More recently, there has been an exponential growth in mobile devices with requirements for seamless personal data synchronization and availability across devices for mobile users. Different methods, such as Chunking, Deduplication, Segmentation, and Delta-encoding, have been used by cloud storage providers to maximize storage space and reduce transmission time [10]. In addition to custom features, various Mobile Cloud Storage providers have developed and implemented services to incorporate techniques of Chunking, Deduplication, Segmentation, and Delta-encoding. Despite all the efforts, there is still much space for improvement in handling syncing data in the mobile cloud storage, as the sync time is much longer than expected under some circumstances. Executing

data synchronization is a challenging task in a mobile/wireless environment with frequent disconnections.

As commercial storage systems are primarily closed source with encrypted data, the public remains unclear regarding their designs and operating processes. Exhaustive research of the sync protocol of specific frameworks can be time-consuming to determine the cause of sync difficulty and maybe inefficient [16].

Furthermore, while some existing services attempt to integrate multiple capabilities to increase sync speed and efficiency in mobile/wireless environments, whether these strategies are viable is still unclear [11].

Ultimately, as a mobile cloud storage system will need storage and network technologies, storage techniques must be flexible and function effectively in a mobile/wireless environment. Communications in such environments are vulnerable to high delay or interruption due to mobility and changing channel conditions [17].

Although several mobile sync frameworks support mobile customer data replication and management systems, they lack support for large artifacts (more than 10MB to Gigabytes) [11] [17] [18] [19] [20] [21]. The main observation from the literature study in the papers [22] [23] [24] revealed that out of 19, only nine frameworks (47.36 percent) support large objects, including commercial frameworks; additionally, few have limitations (in terms of maximum file upload size, chunking support option, and handling techniques for better performance of large objects). In the case of local storage and updates on the cloud and on other client mobile devices, managing large data and maintaining consistency becomes difficult.

Deduplication techniques, in particular, do not always lead to sync efficiency by reducing redundant data transfers. Reasonable attempts to implement delta encoding algorithms are hampered by the distributed nature of storage infrastructure and may lead to high overhead traffic due to the lack of incremental sync. When synchronizing files across a slow network is necessary, the iterative synch scheme suffers from low throughput [16].

To tackle the above challenges, Chunking, Segmentation, and Compression techniques are suggested to improve the sync performance for large objects in modern mobile cloud storage systems, focusing on Data management for large objects.

III. RELATED WORK AND LARGE OBJECT HANDLING TECHNIQUES IN FRAMEWORKS

Cloud storage providers use different techniques to optimize the storage space and speed up data transmissions [10]. The main techniques used in different synchronization frameworks are Chunking, Bundling, Segmentation, Compression, Deduplication, and Delta-encoding.

- 1) **Chunking:** For each piece of the large file uploaded to the Server, some frameworks will break up the upload into multiple pieces and upload the parts one at a time. The process of dividing the file into several smaller files or sections is called file chunking.

- 2) Bundling: During uploading multiple files together, some frameworks of cloud providers combine them into larger bundles for the sake of efficiency before storing them in the cloud. In file bundling, the transmission latency is reduced because cloud servers have fewer connections to mobile clients.
- 3) Segmentation: This method consists of creating a file that takes in a large object and divides it into smaller, self-contained portions. Because Segmentation allows for virtually unlimited segment uploads in a single object with faster and parallel segment uploads, it is a favorite technique used by Internet service providers and application developers.
- 4) Compression: Prior to transmitting to the cloud, the file data can be compressed. With a bit of overhead of processing, data compression can minimize the traffic and reduce storage requirements [11].
- 5) Deduplication: In the case of an identical copy of a file being uploaded by the user or another user in the cloud storage service, that file can be deduplicated. In such cases, instead of sending the file over the network again, it maintains only a unique link to avoid network traffic and reduce storage requirements. Some frameworks support chunk level deduplication if the provider's data storage unit is a chunk object rather than a file.
- 6) Delta-encoding: This process transmits only the modified portion of a file with Compression. If the previous version of the file already exists on the server, the transmission will contain only the modified parts of the file compared previous version.

A. Analysis of Large Object Support in Frameworks

For mobile cloud services, supporting large file uploading and retrieval is critical as data sizes of sharing content are increasing, and mobile users are accessing or sharing enormous size files [9] [8]. However, only Desktop customers and not mobile apps can use valuable large object services. Open-source and commercial cloud storage services for mobile devices are analyzed for large file object support. The studies are mainly classified into three categories: open source frameworks (Parse Server [25], BaasBox [21], Simba's [11] and Open Data Kit 2.0 [26]), academic research reference frameworks (SwiftCloud [18], QuickSync [16]) and commercial mobile cloud frameworks (Dropbox [27], Google Drive [28], Amazon Dynamo [29], CloudKit [30]).

Table II in the Appendix summarizes the large object support and techniques used to optimize the storage space and speed up data transmissions in the various reference frameworks. Many commercial cloud-based frameworks support large objects, but not every framework addresses large files, unfortunately.

Further investigation on frameworks revealed the various techniques of handling large objects and the maximum files supported with options of resuming uploads if interrupted due to network disconnections. A brief explanation of each framework with respect to handling large files is given below.

In the category of open-source reference frameworks, the Parse Server [25] service platform uses MongoDB as the

backend datastore. It can only handle files up to 10MB in size. The 'ParseFile' is a particular data type that makes it easier for Developers to store application files in the cloud. Parse provides another option with a data type known as 'ParseObject' to upload an array of up to 10MB in bytes or as a series of Streams. Implementing the 'SaveAsync' method, the users can save the file to the Parse framework.

Another MBaaS open-source framework based on the Play framework but does not support large files by default is BaasBox [21]. However, based on the component design architecture, some custom implementations are needed to support uploading and downloading data up to 300MB. It is built on the Play web application framework, which is a lightweight, stateless, open architecture. BaasBox requires configuration modifications for the maximum payload size in POST operations. By default, the value of POST request size is 100KB, which can be changed based on the server configuration. Special HTTP requests known as Body parsers in the Play framework are used for POST or PUT operations.

Simba [11] is a recent framework aimed to expedite the development and deployment of data-centric mobile apps and enable them to store data into the cloud storage. Simba extended the table interface of Izzy [17] but the sync protocol of Simba [11] does not support streaming APIs that can handle large files like Media or Video.

SwiftCloud [18] is another middleware framework that is based on the technique of Conflict-Free Replicated Data Types (CRDTs) with the Riak [31] key store and does not support storing objects over 50MB for performance reasons.

Mobius [7] is an application platform that enables real-time cloud-based data replication and messaging for mobile devices. Mobius is focused on addressing the development challenges of data management and messaging for data-centric mobile apps and does not deal with large files. Also, large size handling is not considered in the case of special CRDT cloud types in libraries like TouchDevelop [32] [33]. On the other hand, high-level ideas of sets and maps enable support for large files, but the storage providers do not support the storage of large objects (greater than 50 MB) for performance reasons.

Open Data Kit 2.0 [26] supports the Android operating system and enables the data to be stored in the cloud and handle offline data management. The default size limit on remote procedure calls in Android service is 1MB. To overcome the limitation of 1MB, the ODK Kit exposes higher-level features using a transport-level interface to developers. Using a client-side proxy, the ODK Kit implements a chunking interface.

Dropbox [27] is a commercial sync service provider, and files uploaded via the dedicated REST APIs can be up to 150MB in size. There is a maximum file size of 150 MB only when uploading with the files put API. Dropbox exposes chunked upload API to upload large chunks of data. A chunk of any size between 150 MB and 4 MB can be chosen. Dropbox has a built-in support to resume uploads if the upload is affected because of network disconnections.

QuickSync [16] is a framework that focuses on addressing the sync performance issues considering the network conditions. The framework integrates Seafile/Dropbox APIs and allows large data up to 180MB. It allows the big files to

be uploaded using the "chunk" API support. The APIs also supports automatic presumable data upload. A chunk of any size between 150 MB and 4 MB can be chosen.

Google Drive [28] is another file sync service provider that provides SDK to upload/retrieve data to/from cloud storage. To achieve resumable uploads, more than 5MB of data can be uploaded by making one request or via multiple requests with Google Drive SDK. In order to make the upload as fast as possible, larger chunk sizes are recommended. The upload request in Google Drive recommends chunk size to be in multiples of 256 KB.

Apple has extended the service of iCloud [32] with Cloud Kit [30], a new way of storing and accessing data stored in iCloud storage. The iCloud storage allowance is dependent on user type (paid or premium subscriptions allow larger storage). There is no explicit description document size limitation and Core Data (iOS local) storage limit. However, uploads depend on the device's or iCloud user storage limit. The operating system (iOS) starts and manages the upload and download of data from devices on the iCloud account. The iCloud app only needs to adopt the lifecycle of document management and need not communicate directly with iCloud servers. There is no need to invoke data upload or download operations in most cases.

Amazon uses DynamoDB [29] as the backend data store for cloud storage, focusing on high availability. Mobile SDKs from Amazon provide the way to interact with cloud services via REST APIs for DynamoDB. DynamoDB allows safe update operation of data even in the face of network partitions or server failures. In the DynamoDB, the total maximum item size is 400 KB, which includes both the name attribute's binary length and value. Suppose the application requires more storage space than that allowed by the Amazon DynamoDB limit. In that case, the developer may try compressing large attributes, or the app may store data in AWS Simple Cloud Storage (S3) [33] and associate the Amazon S3 ID with the Amazon DynamoDB entity using the Amazon object identifier.

The key observation from the study is that out of 19, only nine frameworks (47.36%) support large objects, and a few also has limitations. Because current Mobile Sync frameworks do not support large objects and have restrictions (in terms of maximum file upload size, chunking support, configuration, and large object handling strategies for better performance), this article provides an enhanced cloud-based Mobile Sync framework.

IV. OPENSTACK SWIFT AND SUPPORT FOR LARGE OBJECTS

OpenStack platform (developed by NASA and Rackspace) offers a combination of open-source tools for the management of the core cloud computing services in the areas of computing, identity, storage, networking, and image services. This platform can be customized and integrated with additional packages based on the requirements. The framework proposed in this paper is based on Swift, the object storage of OpenStack. The primary design hierarchy of Swift is based on a *tenant/container/object* structure to efficiently, safely, and cheaply store files.

The flow of request processing in Swift is shown in Fig. 1. When a client submits a request to retrieve an object A, an intermediary proxy server retrieves the object. The proxy is regarded as a stateless single entry point to the storage cluster and allows it to scale to arbitrary clients. The proxy is also responsible for determining the appropriate object server for the client's requested object and eventually returning the response object to the client. Apart from the object catalog, the cluster has a running container catalog that stores data about objects grouped within containers. Alternatively, hash functions are used to locate containers. Swift is based on the Web Server Gateway Interface (WSGI), which allows frameworks to define a pipeline. This pipeline comprises one or more middlewares that can pre-process requests before they reach the main web server component.

With regard to the enterprise architecture, Objects correspond to files and are arranged in Containers, i.e., directories. Tenants form the highest level of hierarchy to set up an organization assigning a set of containers. Swift defines two types of ACLs: tenant-level and container-level. A tenant-level ACL allows administrative access to the tenants. Container level ACLs define the permissions on the container for reading, writing, and listing. OpenStack Swift limits the association of any ACL for objects, and the ACL of a container applies to all objects in it.

This study recommended the techniques of Segmentation and chunking to deal with large object files. The principal goal is to attain faster reading and writing speeds using a low object-to-node ratio with a lesser number of objects having large chunk sizes. In addition to effective data reduction, the method uses effective bandwidth reduction techniques [11] [16].

Using the Chunking Mechanism, large objects can be split into smaller parts of a certain data unit when uploading data to cloud storage services without raising the resulting file size issue (in the user interface of memory-constrained devices). Hence when a user attempts to upload a large file, some frameworks divide the input file into smaller parts and then upload these smaller chunks asynchronously later. The proposed chunking mechanism can handle large object data management and end-to-to-end consistency in a mobile cloud environment [16].

It is also a common technique to decrease network traffic by using chunks [34]. The proposed Mobile sync framework in this article provides data structures that can accommodate both table and object data. A single Table ($Table_{DS}$) can contain several rows (Row_{DS}). In the proposed cloud synchronology design, in the case of a Row_{DS} with one or more objects, the changeset for the sync will only consider the modified chunks. Individual chunks are not versioned.

The technique of Segmentation consists of creating a file that takes in a large object and divides it into smaller, self-contained portions. These smaller segments are then transmitted as one object together. Segmentation allows a nearly infinite object size, with more segments that can be uploaded nearly simultaneously in parallel for quick transfer. The Open Stack Swift [35] supports large object uploads by utilizing this segmentation technique. Such a framework would provide cloud-based services with support for large files of any size, from megabytes to gigabytes, and allow the Developers to

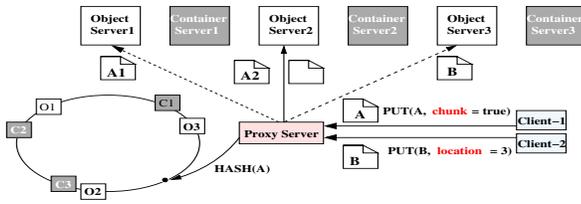


Fig. 1. Request Processing in OpenStack Swift.

```

• Large object support APIs
swift upload test_container -S 1073741824 large_file
swift download test_container large_file

# First, upload the segments
curl -X PUT -H 'X-Auth-Token: <token>' http://<storage_url>/
container/myobject/00000001 --data-binary '1'
curl -X PUT -H 'X-Auth-Token: <token>' http://<storage_url>/
container/myobject/00000002 --data-binary '2'
curl -X PUT -H 'X-Auth-Token: <token>' http://<storage_url>/
container/myobject/00000003 --data-binary '3'

# Next, create the manifest file
curl -X PUT -H 'X-Auth-Token: <token>' -H 'X-Object-Manifest:
container/myobject/' http://<storage_url>/container/myobject/
--data-binary ''

# And now we can download the segments as a single object
curl -H 'X-Auth-Token: <token>' http://<storage_url>/
container/myobject/
    
```

Fig. 2. Open Stack Swift APIs for Large Object Support.

arbitrarily complex, synchronized large objects to be built and maintained in the cloud.

Fig. 1 and 2 shows the sequence of request in segmentation process of OpenStack Object Storage. While the data is being uploaded, the mobile client can specify how much of it to chunk, and the proxy server will divide it into smaller sections. The different blocks are given internal names according to their position in the cluster. Creation of order lists all object file names also creates a manifest file (see Fig. 2). The proxy reads the manifest file to fetch the parts from the GET request of client. With the integration of OpenStack Object Storage, the proposed Mobile Sync framework can support files that are as large as 5 GB in size.

In addition to the techniques of Chunking and Segmentation, this research also incorporates Compression so that the data will be transmitted to the cloud in a smaller, more compressed form. Google's SPDY [36] with Google Protocol buffers [37] are employed, which uses multiplex to HTTP extensions to save on network overhead and perform better Compression for multiplexing requests over a single connection. Google Protocol buffers provide an expandable serialization mechanism for structured data, which are language and platform-neutral.

V. PROPOSED FRAMEWORK FOR DATA SERVICES TO HANDLE LARGE FILES

A. Prototype Implementation

There are two main modules to the proposed architecture of the Mobile Sync framework: one for executing on the handheld device (Data Service) and the other for data in the cloud. Together, these two modules assist enables mobile application development with the Software Development Kit (SDK) provided by the framework. Fig. 3 shows the basic block diagram of the proposed Mobile Sync framework.

The 'Framework Data Service' (FDS) is a module that connects the mobile apps and the Cloud Data Server by using a custom sync protocol and works as a mediator for transferring data and messages. Each client application is built with the

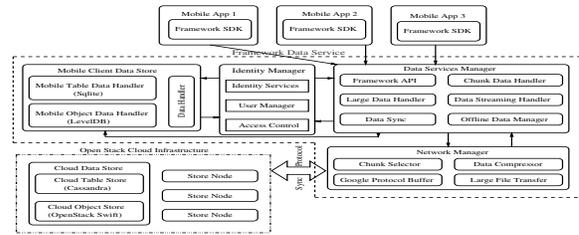


Fig. 3. Block Diagram of Proposed Mobile Sync Framework.

Framework Data Service API provided by the SDK and talks to the system-wide data service (FDS) via streaming and CRUD APIs. Mobile Local database store consisting of Sqlite and LevelDB, enable FDS to save all the application data and metadata in the local store. For the Mobile client apps that run on mobile devices, the local database is not accessible directly since FDS manages it.

The primary modules used to manage large objects are 'Chunk Data Handler' and handler modules like 'Large Data and Streaming Data', which are included within the framework SDK. These modules use a simplified data storage model for apps to use the chunking process proposed in this research article. 'Mobile Client Data Store' is a local storage module to store tabular data and large app objects in the memory of a mobile device (typically the internal flash memory or the external SD card). On the client-side, the framework integrates SQLite for table and LevelDB for object data (in addition to handling large files).

The data model of the proposed Mobile Sync framework facilitates the storage of data for all client applications and hides the complex details of storing and synchronizing data. A custom Data Model consists of chunk layout support with additional support for larger objects.

The 'Cloud Data Store' is the other main server-side module responsible for data management and interacts with the 'Framework Data Service' (FDS) via the custom Sync Protocol. The primary responsibility of this module is to manage data across multiple mobile clients of the framework and implement the chunking & segmentation method at the server-side. For supporting large files in the cloud, through Object Storage, the framework uses the Infrastructure of OpenStack Swift, which has built-in support for the Segmentation process. The framework API must handle the request and responses from the OpenStack Swift server.

B. Framework APIs

The proposed Mobile Sync framework API is designed in the same way as the well-known CRUD interface, allowing apps to set Table/Object properties, access their data, push new data, and resolve conflicts. The framework provides a streaming API abstraction that allows objects to be written to or read from, making it ideal for dealing with large object. It also allows to read or write only a portion of a huge object locally. Table I lists only the APIs that are specific to chunk processing.

TABLE I. CHUNK PROCESSING APIS IN THE PROPOSED MOBILE SYNC FRAMEWORK.

API	Purpose
UploadObjectChunk[]	Creating chunk
GetChunkObjectRange[]	Retrieving chunk
RemoveChunkObject[]	Remove a chunk
RemoveChunks(table, chunkref)	Removing multiple chunks
UploadManifest[dataptr]	Creating large objects using Chunk
DownloadManifest[]	Retrieving large objects using Chunk
RemoveObjects[objptr]	Deleting large objects

C. Mobile-side Large Objects Handling Method through LevelDB with LSM

This proposed Mobile Sync framework integrates SQLite for the table data structure and LevelDB for the file object storage on the client-side. LevelDB is integrated into the SDK of the proposed Mobile Sync framework for handling large files. LevelDB uses an advanced data structure known as Sorted String Table [38] and Log-Structured Merge (LSM) [39] to handle large workloads with gigabytes of data. LevelDB's append and update performance meet the throughput criteria for the mobile client-side layer. LevelDB also has atomic snapshot in-like functionality, which is used for synchronization.

Sorted String Table (SSTable) [38] is a valuable and practical data structure for storing key-value pairs in large numbers and high throughput sequential access. SSTable offers flexibility for sequential read/write requests with workloads consisting of data sets that are Gigabytes in size.

The Log-Structured Merge (LSM) [39] architecture adds various new behaviors to the SSTable. Write operations are always fast no matter the size of the data set (append-only) because the LSM permits all write requests directly to the MemTable index. In addition, random reads can be obtained quickly or quickly served from memory (Initially search MemTable and then the indexes in SSTable). SSTable periodically flushes the MEMTables to the disk.

LevelDB architecture uses the SSTable and MemTable processing schemes to form an efficient database engine with powerful algorithms. Many other related products follow the same architecture include Apache Cassandra, Google BigTable, and Hadoop's HBase.

Fast write operations are allowed in LevelDB regardless of the data-set size, as all write operations are directly executed to the log and the MemTable. The logs of up to 2MB are periodically written to disk assorted string table files (SST) into a database. Each piece of SST data is compressed into single-writable 4K sections. Entries are positioned so that an end-marker block designates the beginning of each data set, and the most recently processed section of the list points to the start of the next. Bloom filters perform lookups more quickly and enable fast search of indexed blocks.

For an improved reading speed, LevelDB breaks SST into sets or levels (see Fig. 4). Each level in LevelDB has ten times the size of the previous one.

D. Server-side Large Object Support with Segmentation

This proposed Mobile Sync framework integrates Cassandra [40] for the table data structure and OpenStack Swift [35]

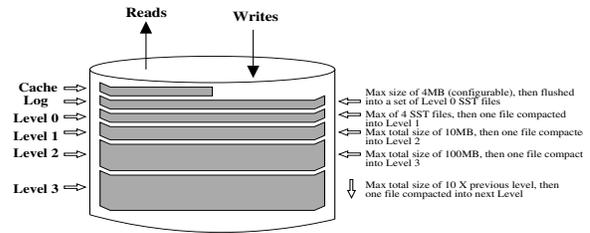


Fig. 4. Architecture of LevelDB with SSTable and MemTable.

for the file object storage on the cloud server-side.

OpenStack Swift [35] employs the process of Segmentation to enable large file uploads. The goal of the Segmentation process is to create a single file that divides the object into segments. With Segmentation, it is possible to upload a single object of virtually any size while taking advantage of the double uploads and the ability to upload multiple segments in parallel.

Unless the size of an object exceeds the maximum value (5 GB) set for the Swift cluster, each object is considered as a single file and stored in the disk. This restriction of the maximum file size of 5 GB avoids one object taking up all of the storage while half of the disk is empty. If the item to be stored is enormous, it is typically stored in several segments to allow future reassembly.

VI. RESULTS AND DISCUSSION

The proposed Mobile Sync framework is designed for use in large file sync scenarios, from a couple of hundred megabytes up to several gigabytes. Multiple modules serve both the clients and the servers from in same architecture.

The evaluation process discusses the performance of the proposed Mobile Sync framework for the following factors:

- 1) The performance of Application Programming Interface (API) consisting of Create, Read and Delete queries and chunking operations [11] [41] [42]
- 2) The efficiency of Sync Protocol [43] [44] [45].

A. Efficiency of Chunking and Data Access APIs

The Application Programming Interface (API) of the proposed Mobile Sync framework for cloud storage is designed to handle requests from thousands of mobile clients. For evaluating the performance of the cloud storage interface, a Linux Test client is implemented. The prototype included a test application to issue requests of subscriptions for reading or writing to a table data structure exposed by the SDK by generating a configurable number of threads. The test application will then generate required read/write (I/O) requests. Both file object and tabular data sizes can be configured as per the requirement during the API testing. The chunk size for objects and consistency scheme is also configurable.

A series of tests are conducted with an individual API call. A combination API invokes the Create-Read-Delete to carry out the performance testing of exposed APIs from the proposed Mobile Sync framework. Test files ranging in varying sizes from 1MB up to 1 GB are generated for testing. A test

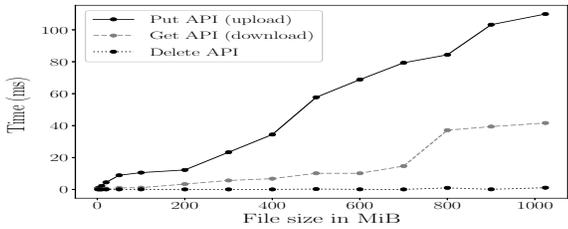


Fig. 5. Latency for Put, Get and Delete APIs in Proposed Mobile Sync Framework.

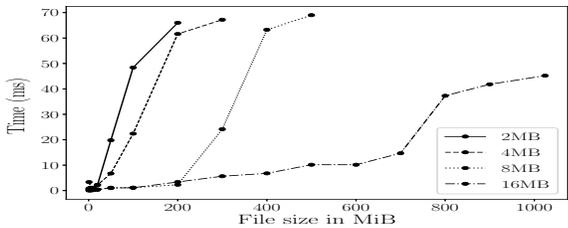


Fig. 6. Latency for Upload Operation with Different Chunk Sizes in Proposed Mobile Sync Framework.

suite is created for large file reading to analyze if the Mobile Sync framework can handle 1MB to 1GB of data. In addition, testing is also done to delete individual files. The Create-Read test is done with 8MB of the default chunk size. A graphical representation in Fig. 5 shows the latency for Put, Get and Delete APIs in the proposed Mobile Sync framework.

As illustrated in Fig. 5, latency increases with the file size. Upload APIs also took longer than download APIs because the upload operation includes data acknowledgments and processing time. It is observed that the time variation of upload data for a size greater than 500MB is less since the framework sync protocol employs data compression during the optimized transfer of data in the network.

The Delete operation is quicker Mobile Sync framework, as the operation completes instantly by marking objects as deleted, instead of removing them. The objects are permanently removed in the scheduled deletion cycle through configuration. The experimental evaluations show that the proposed Mobile Sync framework APIs can handle reading, and removing large files with excellent efficiency.

The third measurement evaluated the performance of object chunking. A method for modifying the object-to-to-node ratio is achieved by altering the testing file size with varying chunk size [11] [41] [42]. A large file size with a larger chunk size results in a low object-to-node ratio, enabling fast reads and writes. Fig. 7 illustrate the impact of configuring different object chunk size (2MB, 4MB, 8Mb and 16MB) on Upload or Put query for proposed Mobile Sync framework. Having a large chunk size demonstrated to be more effective since the large chunks only require a few partitions but can transmit more data efficiently and quickly. For the proposed Mobile Sync framework chunk size of 16MB is recommended subjected to the

B. Sync Protocol Efficiency

The main goal of the proposed Mobile Sync framework is the efficient synchronization of mobile device data through high-level abstractions. It is, therefore, essential to ensure that the synchronization process is not significantly affected by overhead by the proposed framework. So, it has to be shown that the Sync protocol of the Mobile Sync framework is lightweight. To do so the overhead of synchronization of rows with 1 byte of table data with six scenarios is calculated as follows: First with no data, second with one-byte data, and four other cases with 64 KB, 200 KB, 100 MB, and 300 MB objects. In order to minimize compressibility, random bytes are produced for the payload.

Fig. 7 and Fig. 8 show the overhead of sync protocol for a single message with 1 row and ten rows, respectively, for different payloads.

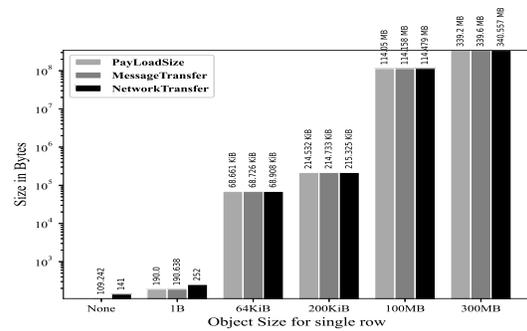


Fig. 7. Overhead of Sync Protocol for a Single Message with 1 Row with Different Payloads.

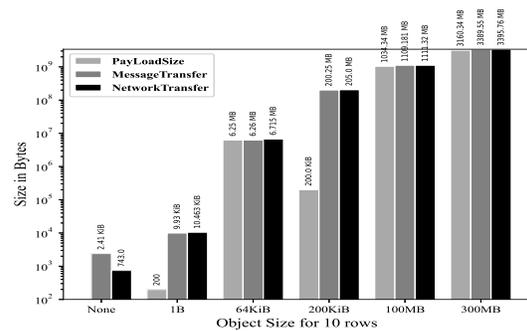


Fig. 8. Overhead of Sync Protocol for a Single Message with 10 Rows with Different Payloads.

The test results indicated that the Sync Protocol of the Mobile Sync framework produces a total message overhead of approximately 109 bytes. There is no object in this request but a single row with 1 byte of tabular data. There will be a reduced overhead for per-row baseline requests with the integration of data compression and batch operations of 10 rows turning into one sync request. Thus, with an increase in the payload (table or object) size, the data transfer overhead ultimately becomes negligible.

To sum up, the network overhead is reduced for the batched row or multiple rows operations because of data compression used in sync protocol. By working with single rows instead of

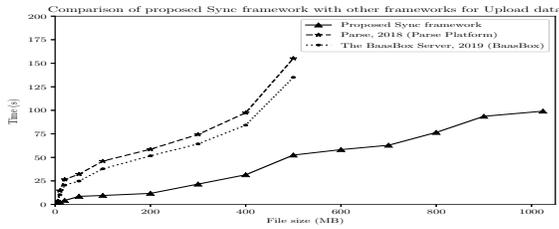


Fig. 9. Comparison of Proposed Mobile Sync Framework with other Frameworks for Upload Data.

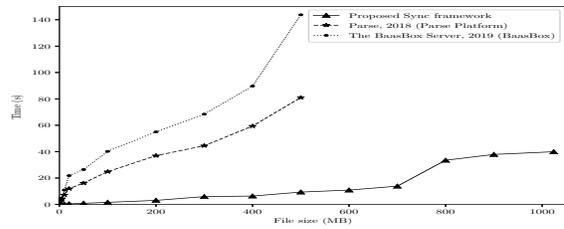


Fig. 10. Comparison of Proposed Sync Framework with Other Frameworks for Download Data.

batches, the Mobile Sync framework incurs a little overhead. Hence the sync protocol is lightweight and efficient with batching or group operations. More or less, empirical testing supports the claim that synchronization protocol is lightweight.

C. Performance Comparison with Other Sync Frameworks

The proposed Mobile Sync framework local performance is compared with two other open source mobile frameworks namely ParseServer [25] and BaasBox [21]. A file size of up to 500MB is tested in both Parse Server and BaasBox with the prebuilt virtual machine setup available. Some custom changes are required in Parse Server and BaasBox to support files upto 500MB .

The proposed Mobile Sync framework local performance is compared with two other open-source mobile frameworks, namely ParseServer and BaasBox, with a chunk size of 16MB. Fig. 9 shows latency comparison of Upload API (Put query) for the proposed Mobile Sync framework. The BaasBox upload process is handled by the REST APIs in the Play framework and takes more time than the proposed Mobile Sync framework due to the time taken for HTTP buffer processing and acknowledgment. Also, BaasBox does not support dedicated APIs through the Play framework to handle large files. BaasBox seems to be better in efficiency than the Parse Server framework. Overall the proposed framework reduces synchronization time with object chunking by 65.4% for upload on average when compared ParseServer and BaasBox.

Fig. 10 shows latency comparison of Get API (download query) for the proposed Mobile Sync framework, Parse Server, and BaasBox. The performance of Parse Server is comparatively better than BaasBox for files up to 500MB, and it should be noted that the experiment for download run on mobile with only a single file and no other application running on the device. The latency, which is the measure of response time between the device and a service's Server, must be considered. Since the BaasBox, Parse Server, and proposed Mobile Sync framework run on the virtual network setup in the testing network, the performance is better. The proposed Mobile Sync framework runs better than Parse Server and BaasBox since while downloading, the tests are configured with a chunking feature of 16MB to retrieve data in the device. The download tests are also dependent on the memory available for downloading and processing in the client device, depending on the RAM size. Overall the proposed framework reduces synchronization time with object chunking by 93.7% for download on average when compared to ParseServer and BaasBox.

Delete API performance in the proposed Mobile Sync framework and other two frameworks Parse Server, and Baas-Box is almost identical. These frameworks follow a lazy deletion policy wherein objects are marked deleted and physically removed after the completion of the sync operation.

VII. CONCLUSION AND FUTURE WORK

It is essential to develop Mobile applications to store and access data from backend enterprise systems. Certain usage patterns require storing and access data in large files. Uploading or downloading large files is a complex and time-consuming process for developers because of difficulties with latency, bandwidth, speed, errors, and disruptions to service in wireless mobile environment. In this article different techniques (Chunking, Bundling, Segmentation, Compression, Deduplication and Delta-encoding) for large objects (ranging from hundreds of megabytes to 5GB) in different mobile cloud storage solutions are analyzed. A Mobile Sync framework is proposed with both flexible chunking support and high throughput feature with segmentation technique to store and access large objects for a mobile cloud storage framework. The prototype implementation of the framework supported upload or download larger objects from the cloud storage, with support for tunable chunking configuration at the mobile side and local caching or data transfer only for a part of large objects. The extensive evaluations under the representative data-set show that the Mobile Sync framework works can quickly and effectively store large files and effortlessly keep minimal traffic burden on large workloads with reduced synchronization time. As a future work it is desired to extend the Mobile Sync framework with support for consistency schemes (like Eventual, Strong, Sequential consistency and others).

REFERENCES

- [1] P. Mell and T. Grance, "The nist definition of cloud computing recommendations of the national institute of standards and technology," *Nist Special Publication*, vol. 145, p. 7, 2011.
- [2] M. Satyanarayanan, "Fundamental challenges in mobile computing," *Annual ACM Symposium on Principles of Distributed Computing*, pp. 1-7, 1996.
- [3] A. Gheith, R. Rajamony, P. Bohrer, K. Agarwal, M. Kistler, B. W. Eagle, C. Hambridge, J. Carter, and T. Kaplinger, "Ibm bluemix mobile cloud services," *IBM Journal of Research and Development*, vol. 60, no. 2-3, pp. 7-1, 2016, doi:https://doi.org/10.1147/JRD.2016.2515422.
- [4] S. Gilbert and N. Lynch, "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services," *ACM SIGACT News*, vol. 33, no. 2, p. 51, 2002. [Online]. Available: http://dl.acm.org/citation.cfm?id=564585.564601

- [5] N. Agrawal, A. Aranya, and C. Ungureanu, "Mobile data sync in a blink," in *Presented as part of the 5th USENIX Workshop on Hot Topics in Storage and File Systems*, 2013.
- [6] H. Wu, L. Hamdi, and N. Mahe, "Tango: a flexible mobility-enabled architecture for online and offline mobile enterprise applications," in *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. IEEE, 2010, pp. 230–238.
- [7] B.-G. Chun, C. Curino, R. Sears, A. Shraer, S. Madden, and R. Ramakrishnan, "Mobiuz: unified messaging and data serving for mobile apps," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 141–154, doi:<https://doi.org/10.1145/2307636.2307650>.
- [8] F. Shanon Montelongo, "How to upload large files," <https://blog.filestack.com/thoughts-and-knowledge/how-to-upload-large-files/>.
- [9] Z. Li, X. Wang, N. Huang, M. A. Kaafar, Z. Li, J. Zhou, G. Xie, and P. Steenkiste, "An empirical analysis of a large-scale mobile cloud storage service," in *Proceedings of the 2016 Internet Measurement Conference*. ACM, 2016, pp. 287–301.
- [10] I. Drago, M. Mellia, M. M. Munafò, A. Sperotto, R. Sadre, and A. Pras, "Inside dropbox: understanding personal cloud storage services," in *Proceedings of the 2012 ACM conference on Internet measurement conference*. ACM, 2012, pp. 481–494, doi:<https://doi.org/10.1145/2398776.2398827>.
- [11] D. Perkins, N. Agrawal, A. Aranya, C. Yu, Y. Go, H. V. Madhyastha, and C. Ungureanu, "Simba: Tunable end-to-end data consistency for mobile apps," in *Proceedings of the Tenth European Conference on Computer Systems*. ACM, 2015, p. 7, doi:<https://doi.org/10.1145/2741948.2741974>. [Online]. Available: <https://github.com/SimbaService/Simba>
- [12] Y. Bai and Y. Zhang, "Stoarranger: Enabling efficient usage of cloud storage services on mobile devices," in *Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on*. IEEE, 2017, pp. 1476–1487.
- [13] T. Ketola, "Quantifying software development: Applying mobile monetization techniques to your software development process," in *2014 Computer Games: AI, Animation, Mobile, Multimedia, Educational and Serious Games (CGAMES)*. IEEE, 2014, pp. 1–4.
- [14] A. Inc, "ios app ota limit in cellular network," <https://github.com/baasbox/baasbox/>, accessed: 2021-05-01.
- [15] Google, "Reduce your app size," <https://developer.android.com/topic/performance/reduce-apk-size>, accessed: 2021-05-01.
- [16] Y. Cui, Z. Lai, X. Wang, and N. Dai, "Quicksync: Improving synchronization efficiency for mobile cloud storage services," *IEEE Transactions on Mobile Computing*, vol. 16, no. 12, pp. 3513–3526, 2017, doi:<https://doi.org/10.1109/TMC.2017.2693370>.
- [17] S. Hao, N. Agrawal, A. Aranya, and C. Ungureanu, "Building a delay-tolerant cloud for mobile data," in *2013 IEEE 14th International Conference on Mobile Data Management*, vol. 1. IEEE, 2013, pp. 293–300, doi:<https://doi.org/10.1109/MDM.2013.43>.
- [18] N. Preguiça, M. Zawirski, A. Bieniusa, S. Duarte, V. Balegas, C. Baquero, and M. Shapiro, "Swiftcloud: Fault-tolerant geo-replication integrated all the way to the client machine," in *2014 IEEE 33rd International Symposium on Reliable Distributed Systems Workshops (SRDSW)*. IEEE, 2014, pp. 30–33.
- [19] V. Balegas, S. Duarte, C. Ferreira, R. Rodrigues, N. Preguiça, M. Najafzadeh, and M. Shapiro, "Putting consistency back into eventual consistency," in *Proceedings of the Tenth European Conference on Computer Systems*. ACM, 2015, p. 6, doi:<https://doi.org/10.1145/2741948.2741972>.
- [20] Parse, "Parse," 2016, <http://parse.com>.
- [21] "The baasbox server," <https://github.com/baasbox/baasbox/>, accessed: 2019-01-26.
- [22] Y. P. Faniband, I. Ishak, F. Sidi, and M. A. Jabar, "A review of data synchronization and consistency frameworks for mobile cloud applications," *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, vol. 9, no. 12, pp. 601–611, 2018.
- [23] —, "Netmob: A mobile application development framework with enhanced large objects access for mobile cloud storage service," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 7, 2019. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2019.0100784>
- [24] —, "Enhancing mobile backend as a service framework to support synchronization of large object," in *Proceedings of the 2017 International Conference on Information Technology*, ser. ICIT 2017. New York, NY, USA: ACM, 2017, pp. 383–387, doi:<http://doi.acm.org/10.1145/3176653.3176719>.
- [25] P. Platform, "Parse platform," 2016, <https://parseplatform.github.io/>.
- [26] W. Brunette, S. Sudar, M. Sundt, C. Larson, J. Beorse, and R. Anderson, "Open data kit 2.0: A services-based application framework for disconnected data management," in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2017, pp. 440–452, doi:<https://doi.org/10.1145/3081333.3081365>.
- [27] Dropbox, "Build your app on the dropbox platform," 2016, <https://www.dropbox.com/developers>.
- [28] G. Drive, "Google drive," 2016, <https://developers.google.com/drive/>.
- [29] "Amazon dynamodb - best practices for storing large items and attributes," <https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/bp-use-s3-too.html>, accessed: 2019-01-26.
- [30] A. Shraer, A. Aybes, B. Davis, C. Chrysafis, D. Browning, E. Krugler, E. Stone, H. Chandler, J. Farkas, J. Quinn *et al.*, "Cloudkit: structured storage for mobile applications," *Proceedings of the VLDB Endowment*, vol. 11, no. 5, pp. 540–552, 2018.
- [31] R. Klopphaus, "Riak core: Building distributed applications without shared state," in *ACM SIGPLAN Commercial Users of Functional Programming*. ACM, 2010, p. 14, doi:<https://doi.org/10.1145/1900160.1900176>.
- [32] A. Inc, "icloud for developers," 2016, <http://developer.apple.com/icloud/>.
- [33] A. W. S. Mobile, "Aws sdk," 2016, <https://aws.amazon.com/mobile/>.
- [34] A. Muthitachareon, R. Morris, T. M. Gil, and B. Chen, "Ivy: A read/write peer-to-peer file system," *ACM SIGOPS Operating Systems Review*, vol. 36, no. SI, pp. 31–44, 2002, doi:<https://doi.org/10.1145/844128.844132>.
- [35] "Openstack swift object storage service," 2018, <http://swift.openstack.org>.
- [36] Google, "Research & Drafts — SPDY — Google Developers." [Online]. Available: <https://developers.google.com/speed/protocols>
- [37] "Protocol buffers," <https://developers.google.com/protocol-buffers/>, accessed: 2019-01-26.
- [38] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber, "Bigtable: A distributed storage system for structured data," *ACM Transactions on Computer Systems (TOCS)*, vol. 26, no. 2, p. 4, 2008.
- [39] P. O'Neil, E. Cheng, D. Gawlick, and E. O'Neil, "The log-structured merge-tree (lsm-tree)," *Acta Informatica*, vol. 33, no. 4, pp. 351–385, 1996.
- [40] A. Lakshman and P. Malik, "Cassandra: structured storage system on a p2p network," in *Proceedings of the 28th ACM symposium on Principles of distributed computing*. ACM, 2009, pp. 5–5, doi:<https://doi.org/10.1145/1582716.1582722>.
- [41] H. Wada, A. Fekete, L. Zhao, K. Lee, and A. Liu, "Data Consistency Properties and the Trade-offs in Commercial Cloud Storage: the Consumers' Perspective." *Cidr*, pp. 134–143, 2011. [Online]. Available: http://www.cidrdb.org/cidr2011/Papers/CIDR11{_}_}Paper15.pdf
- [42] L. Rupperecht, R. Zhang, B. Owen, P. Pietzuch, and D. Hildebrand, "Swiftanalytics: Optimizing object storage for big data analytics," in *2017 IEEE International Conference on Cloud Engineering (IC2E)*. IEEE, 2017, pp. 245–251.
- [43] D. Bermbach, E. Wittern, and S. Tai, *Cloud service benchmarking*. Springer, 2017.
- [44] M. Klems, D. Bermbach, and R. Weinert, "A runtime quality measurement framework for cloud database service systems," in *Quality of Information and Communications Technology (QUATIC), 2012 Eighth International Conference on the*. IEEE, 2012, pp. 38–46.
- [45] D. Bermbach and S. Tai, "Benchmarking eventual consistency: Lessons learned from long-term experimental studies," in *2014 IEEE International Conference on Cloud Engineering*. IEEE, 2014, pp. 47–56.

- [46] S. Burckhardt, "Bringing touchdevelop to the cloud," 2013, <https://www.microsoft.com/en-us/research/blog/bringing-touchdevelop-to-the-cloud/>.
- [47] S. Burckhardt, M. Fähndrich, D. Leijen, and B. P. Wood, "Cloud types for eventual consistency," in *European Conference on Object-Oriented Programming*. Springer, 2012, pp. 283–307, doi:https://doi.org/10.1007/978-3-642-31057-7_14.
- [48] D. Bermbach, J. Kuhlenkamp, B. Derre, M. Klems, and S. Tai, "A middleware guaranteeing client-centric consistency on top of eventually consistent datastores." in *IC2E*, 2013, pp. 114–123, doi:<https://doi.org/10.1109/IC2E.2013.32>.
- [49] "Kony mobilefabric," http://docs.kony.com/7_0_PDFs/sync/kony_sync_orm_api_guide.pdf, accessed: 2019-01-26.
- [50] "Evernote system limits," <https://help.evernote.com/hc/en-us/articles/209005247>, accessed: 2019-01-26.
- [51] Kinvey, "Kinvey baas," 2016, <https://www.kinvey.com/>.

APPENDIX

TABLE II. SUMMARY OF VARIOUS TECHNIQUES EMPLOYED TO SUPPORT LARGE OBJECTS IN DIFFERENT REFERENCE DATA SYNCHRONIZATION FRAMEWORKS.

Framework	Large Object Support	Chunking	Bundling	Segmentation	Compression	Deduplication	Delta Encoding	Open-Source
Parse Server [25]	X	X	X	X	X	X	X	✓
BaaSBox [21]	X	X	X	X	X	X	X	✓
Open Data Kit 2.0 [26] *	X	X	X	X	-	X	X	✓
Simba [11]	X	-	X	X	✓	X	X	✓
SwiftCloud [18]	X	-	✓	X	X	X	X	✓
Indigo [19]	X	-	✓	X	X	X	X	X
Izzy [17]	X	-	X	X	X	X	X	X
Mobius [7]	X	-	X	X	X	X	X	X
TouchDevelop [46] [47]	X	-	✓	X	-	X	X	X
Middleware for client-centric consistency [48]	X	-	X	X	-	-	-	X
QuickSync [16]	✓	✓	✓	X	X	X	X	X
NetMob [23]	✓	✓	✓	✓	✓	X	X	X
Bluemix Mobile Cloud Service [3]	✓	✓	✓	✓	-	-	-	X
Dropbox [27]	✓	✓	✓	X	✓	X	X	X
Amazon DynamoDB [29]	✓	X	X	X	-	-	-	X
Google Drive [28]	✓	✓	X	✓	-	✓	✓	X
iCloud with CloudKit [32]	✓	✓	-	-	-	-	✓	X
Kony [49]	✓	✓	✓	-	-	-	-	X
Evernote [50]	✓	✓	-	-	-	✓	X	X
Kinvey [51]	✓	-	-	-	-	-	-	X

A Comparative Study of Stand-Alone and Hybrid CNN Models for COVID-19 Detection

Wedad Alawad¹, Banan Alburaidi², Asma Alzahrani³, Fai Alfaj⁴
Department of Information Technology, College of Computer,
Qassim University, Buraydah, 51452, Saudi Arabia¹
Department of Information Technology, College of Computer,
Qassim University, Buraydah, 51452, Saudi Arabia^{2,3,4}

Abstract—The COVID-19 pandemic continues to impact both the international economy and individual lives. A fast and accurate diagnosis of COVID-19 is required to limit the spread of this disease and reduce the number of infections and deaths. However, a time consuming biological test, Real-Time Reverse Transcription–Polymerase Chain Reaction (RT-PCR), is used to diagnose COVID-19. Furthermore, sometimes the test produces ambiguous results, especially when samples are taken in the early stages of the disease. As a potential solution, machine learning algorithms could help enhance the process of detecting COVID-19 cases. In this paper, we have provided a study that compares the stand-alone CNN model and hybrid machine learning models in their ability to detect COVID-19 from chest X-Ray images. We presented four models to classify such kinds of images into COVID-19 and normal. Visual Geometry Group (VGG-16) is the architecture used to develop the stand-alone CNN model. This hybrid model consists of two parts: the VGG-16 as a features extractor, and a conventional machine learning algorithm, such as support-vector-machines (SVM), Random-Forests (RF), and Extreme-Gradient-Boosting (XGBoost), as a classifier. Even though several studies have investigated this topic, the dataset used in this study is considered one of the largest because we have combined five existing datasets. The results illustrate that there is no noticeable improvement in the performance when hybrid models are used as an alternative to the stand-alone CNN model. VGG-16 and (VGG16+SVM) models provide the best performance with a 99.82% model accuracy and 100% model sensitivity. In general, all the four presented models are reliable, and the lowest accuracy obtained among them is 98.73%.

Keywords—COVID-19; convolutional neural network; hybrid models; chest X-Ray; deep learning

I. INTRODUCTION

The continuous outbreak of the novel coronavirus was first reported in Wuhan, Hubei Province, China. In a preliminary report, it was revealed that the virus shares an 88% serial identity with two coronaviruses derived from bats, similar to SARS. The new coronavirus was preliminarily named nCov-2019. In February 2020, a study group from the International Committee on Taxonomy of Viruses classified the virus as SARS-CoV. Shortly thereafter, the World Health Organization (WHO) formally named the disease caused by the novel coronavirus “COVID-19” [1].

The most common symptoms of this disease are coughing, headaches, fatigue, shortness of breath, loss of smell, pain in the throat, and a high temperature. COVID-19 continues to have a destructive impact on global health and commerce, as

well as individuals’ lives. The number of infections and deaths increases day by day. As of April 18, 2021, the total number of infected persons around the world has reached 141 million with a mortality rate of 3 million.

Therefore, there is a need to cooperate across disciplines and integrate resources to defend against COVID-19 and prevent it from further spread. Because this virus has the ability to spread fast between people, the first and most important step in COVID-19 defense is early detection. The most common method used to detect COVID-19 is the Transcription-Polymerase Chain Reaction (RT-PCR) test. The RT-PCR test usually takes up to 6 hours to produce results [1]. This waiting time is long when we consider the urgent need to test millions of samples and receive the results as fast as possible to prevent those who are infected from spreading the disease to others. Additionally, sometimes patients who have already been exposed to the virus and show severe symptoms could still get false negative results in the (RT-PCR) test [1].

Thus, the development of a fast and efficient alternative is necessary to improve the process of diagnosing COVID-19. One possibility is to exploit the fact that the disease can be diagnosed using chest X-Ray images. In contrast to other types of medical imaging, X-Ray images are available at most hospitals, and have been routinely used in COVID-19 diagnoses thus far. Moreover, they are cost-efficient and quickly produce results [2]. Machine learning and deep learning algorithms are promising in this case. Potential automated COVID-19 detection models can be provided by training such algorithms on chest X-Ray image datasets. These models may aid in producing test results within a few seconds as well as reduce inaccurate results.

This paper addresses the following research question: How would replacing the fully connected layer with a machine learning classifier affect the classification model performance? To answer this question, we examined the efficiency of two types of machine learning models used to detect COVID-19. Using one of the largest available chest X-Ray image datasets, we have built four different models to study and compare between the stand-alone CNN model and hybrid machine learning models with regard to their general effectiveness, giving particular attention to their COVID-19 classifying abilities.

The rest of this paper is structured as follows: Section II provides background and analysis of pre-existing relevant studies. Section III discusses the details of datasets and methodologies used in this study. Section IV presents and discusses

the results of our experiments. Section V concludes the paper by mentioning the most prominent points of this study. Finally, section VI provides ideas that may be implemented in the future.

II. RELATED WORK

Several studies have investigated the use of machine learning techniques to detect COVID-19. Among machine learning algorithms, most researchers used CNN techniques, e.g., Inception (GoogleNet), ResNet (Residual Networks), and DenseNet (Dense Networks), to build the detection models. From the dataset perspective, chest X-Ray images were used more frequently than Computed Tomography (CT) images to develop those models.

Most of the related studies have faced challenges due to a lack of available datasets. Researchers have applied methods such as augmentation and K-fold cross validation to overcome this setback. In [3], the dataset size reached 1,592 images after augmentation, and the authors used the ImageNet dataset and four CNN techniques to build detection models. The accuracy of VGG16 and VGG19 based models were the highest, having achieved a 99.38% accuracy. In contrast, K-fold cross validation was the method used to train the detection models on more data in [4]. In that study, VGG 16 and ResNet50 techniques were used to develop models that distinguish between COVID-19 and pneumonia. The dataset was comprised of 204 images, and their results showed that 89.2% and 80.39% of COVID-19 cases were identified correctly by these techniques, respectively.

In addition to the rarity of available datasets, the quality of obtainable images needs to be enhanced to improve the performance of detection models. To achieve this, the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm was applied in [5]. The authors conducted a comparison study to investigate the effect of using CLAHE to enhance covid diagnosis. They compared the detection accuracy of the model when the CLAHE was applied and when the original datasets were used to build the model without applying any image quality enhancement techniques. The accuracy of the developed model increased from 83.00% to 92.00% after the implementation of CLAHE.

Despite COVID-19 dataset challenges, many detection models with acceptable performance have been presented. The transfer learning concept was used to build detection models in previous studies. In [6], the researchers developed a model that detected COVID-19 from chest X-ray images. They used a modified version of VGG-19 by adding a MLP (multilayer perceptron) on top of the VGG-19 model. The accuracy of the model was 96.3%. Additionally, the authors in [7] used seven pre-trained models to develop a deep learning framework that identified COVID-19 cases. Their results showed that VGG19 and DenseNet201 achieved better performance when compared with other models. The accuracy was 90% and the sensitivity was 100% for COVID-19 and 80% for normal images.

Some papers have used hybrid models, meaning they combined multiple models to solve one problem. The authors in [8] used four different types of ensemble learning, feature ensembles, majority voting, feature classification, and class modification, to classify COVID-19 and pneumonia cases.

More specifically, SVM, Bagging Classifier, and AdaBoost were used as classifiers in the models they developed. The accuracy of the combination of Inception V3 and Bagging was 99.36

Table I shows a summary of previous studies that used CNN architectures to build COVID-19 detection models. Along with the performance of detection models, the table illustrates the size of datasets and the CNN technique used.

III. STAND-ALONE CNN MODEL VS. HYBRID MODELS

In this study we have conducted several experiments to investigate the abilities of stand-alone CNN and hybrid models to efficiently and accurately detect COVID-19 in patients. This section presents the datasets and methodologies used in our comparative study.

A. Dataset Collections

As aforementioned, one of the challenges that researchers have faced in previous studies is the limited repository of COVID-19 datasets. Moreover, the datasets that do exist are relatively small. Thus, we used images from four datasets to train and develop our models on COVID-19 cases. For normal cases, we used one dataset. Fig. 1 shows samples of COVID-19 and normal chest X-ray images and Table II illustrates a summary of these datasets.

TABLE I. SUMMARY OF CNN-BASED MODELS FOR COVID-19 DETECTION

Ref.	Dataset size	CNN Technique	Accuracy	Sensitivity	Specificity
A Deep Learning Approach to Detect COVID-19 Patients from Chest X-ray Images [7]	COVID-19 : 295 Normal : 659	Sequential CNN	98.3%	-	-
A Deep-Learning-Based Framework for Automated Diagnosis of COVID-19 Using X-ray Images [3]	COVID-19 : 790 Normal : 802	VGG19	99.38%	100%	98.77%
Using X-ray images and deep learning for automated detection of coronavirus disease [9]	Pneumonia : 4,273 COVID-19 : 231 Normal : 1583	Inception-Resnet-V2	92.18%	92.11%	96.06%
COVID-19 Detection in Chest X-ray Images using a Deep Learning Approach [5]	Pneumonia : 647 COVID-19 : 204 Normal : 649	VGG16	92%	94.92%	92%
Deep learning COVID-19 detection bias- accuracy through artificial intelligence [6]	COVID-19 : 181 Normal : 364	VGG19	96.3%	97.1%	-
COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks [10]	Pneumonia : 700 COVID-19 : 224 Normal : 504	VGG19	98.75%	92.85%	98.75%
Finding COVID-19 from Chest X-Rays using Deep Learning on a Small Dataset [4]	COVID-19 : 102 Pneumonia : 102	ResNet50	89.2%	-	-
Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning [11]	COVID-19 : 520 Normal : 5000	Framework of(ResNet18, ResNet50, SqueezeNet, DenseNet-121)	-	98%	90%
COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images [12]	COVID-19 : 25 Normal : 25	VGG19, DenesNet121	90%	100%	-
Deep Learning Based Hybrid Models for Prediction of COVID-19 using Chest X-Ray [8]	Pneumonia : 4273 COVID-19 : 1182 Normal : 1583	VGG-16	90.19%	94.16%	-
		VGG-16 & SVM	91.19%	93.15%	-
		VGG-16 & Bagging	90.19%	92.16%	-
		VGG-16 & Ada Boot	90.19%	90.10%	-

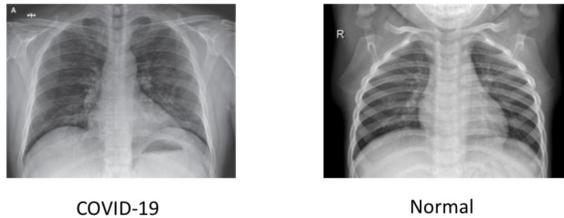


Fig. 1. Normal and COVID-19 Chest X-Ray Images.

1) *COVID-19 chest X-ray images datasets*: The details of the four COVID-19 chest X-ray image datasets are as follows:

COVID-19_dataset-1: We retrieved this dataset from the Github repository, and it is more popular than any other currently available datasets. It was created and collected by Joseph Paul Cohen, a postdoctoral fellow at Montreal University [13]. The dataset contains 930 images of chest X-ray and CT images of patients with diverse diseases, including both bacterial and viral illnesses, as well as COVID-19 and pneumonia. COVID-19 images alone account for 584 of the dataset's images, with the remainder classified as other. The chest X-ray images are classified into four views, Posterior Anterior (PA), Anterior Posterior (AP), AP Supine, and Lateral.

COVID-19_dataset-2: We obtained the second dataset from the Github repository. It was created by Linda Wang and colleagues from the University of Waterloo in Canada [14]. The dataset contains 238 chest X-ray images, 58 of which are images of patients infected with COVID-19. The COVID-19 images are classified with two views: 32 images are PA and 26 images are AP.

COVID-19_dataset-3: The previous team, Linda Wang and colleagues, created the third dataset as well. This dataset contains 55 chest X-ray images, 35 of which are COVID-19 images and the rest are either Pneumonia or not classified [15].

COVID-19_dataset-4: Dataset 4 is also from the Github repository. It was created by the Institute for Diagnostic and Interventional Radiology at the Hannover Medical School in Hannover, Germany [16]. It contains 243 images of COVID-19 chest X-ray images. Those images include two views: 49 are PA images and 194 images are AP.

2) *Normal Chest X-ray Images Datasets*: The details of the normal chest X-ray image datasets are as follows:

Normal-dataset: We obtained this dataset from the Kaggle website. It was created by Paul Mooney, Developer Advocate at Kaggle [17]. The dataset contains 5,863 chest X-Ray images with two classes, pneumonia and normal. The number of normal images is 1,583 and the number of pneumonia images is 4,273. All images in this dataset were in AP view. To balance our data, we took only 690 normal images from this dataset.

3) *Data Preprocessing*: We converted all images to JPEG format to provide ease by handling only one format type, and to reduce the dataset size to accelerate the training process. Furthermore, we applied normalization to improve image clarity and overall quality. Furthermore, we resized them to 224*224. We tried several rations to split data into train and

TABLE II. DETAILS OF DATASETS USED

Dataset	Medical Images Type	Classes	View	Count of COVID-19 Images	Dataset Size
COVID-chestxray-dataset [13]	X-ray , CT	COVID-19, viral, bacterial, pneumonias	AP, PA, AP Supine, Lateral	504	930
Actualmed-COVID-chest xray-dataset [14]	X-ray	COVID-19 , no finding	PA, AP	58	238
Figure1-COVID -chestxray-dataset [15]	X-ray	COVID-19, pneumonias , no finding	-	35	55
COVID-19-image-repository [16]	X-ray	COVID-19	AP ,PA	243	243
COVID-19-image-repository [16]	X-ray	Pneumonia, Normal	AP	0	5863

test sets. We found that using 60% of data for training the model and 40% for testing it achieved the best performance.

B. Experimental Environment

We used Keras to implement our COVID-19 detection models, which is an open-source library for deep learning applications written in Python [18]. The code was implemented by the Colab environment, a service hosted by Jupyter. The Colab environment provides free access to computing resources, including GPU, which is the most widely used computing technology in artificial intelligence [19]. Additionally, we used TensorFlow as the backend for the machine learning platform [20].

C. Stand-alone CNN Model

As part of our study, we developed a stand-alone CNN model to distinguish between COVID-19 and normal cases. Convolutional neural networks, also known as (CNN), are a specialized kind of neural network used in the computer vision field that contributes to automatic feature extraction and data processing with a known grid-like topology. The CNN architecture has three main layers: convolutional layer, pooling layer, and fully connected layer [21]. The main components of CNN architecture are illustrated in Fig. 2.

Convolutional layer : This layer applies several filters to the input to generate feature maps [21].

Pooling layer : This layer reduces the size of feature maps in terms of reducing the internal dimensions. Max pooling and average pooling are the two operations available in this layer [21].

Fully connected layer : This layer is also called the dense layer. In the fully connected layer, the inputs are connected to the output with a learnable weight and are assigned to the final outputs [21].

Activation function : This is a function used to facilitate knowledge of difficult and complex patterns. It includes sigmoid, tanh, and Rectified Linear Unit (ReLU). The most common is the corrected linear unit (ReLU) [21].

From 1989 to present, improvements have been made in the CNN architecture in terms of number of layers, parameters, and functions. These architectures vary from lightweight to heavyweight structures [23]. From among those architectures, we chose the Visual Geometry Group (VGG-16) architecture to implement our detection models based on information we

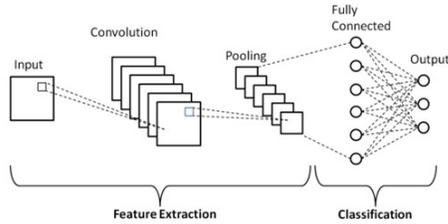


Fig. 2. Key Components of CNN Architecture [22].

garnered from the literature review. VGG architecture was proposed by Simonyan and Zisserman in 2014. It includes VGG-11, VGG-11-LRN, VGG-13, VGG-16, and VGG-19. The depth of the layers vary [24].

To develop our detection model, we exploited the principle of transfer learning. We used weights of a pre-trained VGG-16 that were trained on a large dataset called ImageNet. It learned a good representation of low level features like spatial, edges, rotation, lighting, shapes. These features can be integrated to enable the knowledge transfer and act as a feature extractor for new images in different computer vision problems [25].

In our model, we removed the top layer of the pre-trained model to train the model on a new chest X-ray images dataset. For optimization, we used Adam’s algorithm, an effective stochastic optimization method for training deep learning models with a 0.001 learning rate. The number of training epochs was 50 and the batch size for each epoch was 26. The activation function in the hidden layer was ReLU, and sigmoid was used in the last layer because our classification problem is binary. The parameters of our stand-alone CNN model are illustrated in Table III and the model implementation steps are shown in Fig. 3.

D. Hybrid Models

In addition to the stand-alone CNN model, we have developed three hybrid models. Each model consists of a CNN architecture (VGG-16) for feature extraction and one of the following classification algorithms for classification: Support Vector Machine (SVM), Random Forests (RF), and Extreme Gradient Boosting (XGBoost). Fig. 4 shows an illustrative diagram of the proposed hybrid models.

A brief description of the selected classification algorithms follows:

Support Vector Machines (SVM): Essentially, SVM is a classification algorithm which tries to find the plane that separates the classes with the widest margin in the sample space in the most convenient way [26].

Random Forest (RF): RF is a collective classification and regression algorithm that uses decision trees as a classifier. Each decision tree is trained using a random data set derived from the original data set. The majority voting is used for the final classification [26].

Extreme Gradient Boosting (XGBoost): This is a boosting algorithm for classification and regression tree models, which is derived from the gradient lifting decision tree [27].

TABLE III. PARAMETERS OF OUR STAND-ALONE CNN MODEL.

Number of training epochs	50
Batch size	26
Optimization	Adam
Learning rate	0.001
input shape	(224,224,3)
Number of Fully connected layers	2
Number of convolution layers	13
Number of pooling layers	5
Activation function in output layer	Sigmoid

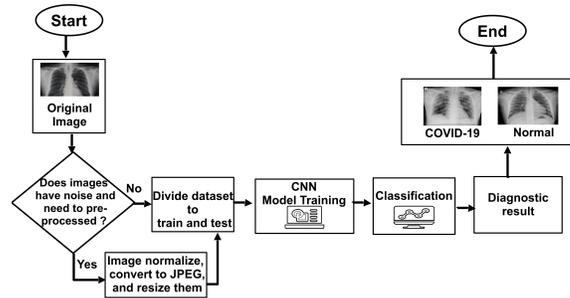


Fig. 3. Implementation Steps of the Stand-alone CNN Model.

We followed several steps to develop our hybrid models, which you can find illustrated in Fig. 5. First, we trained the VGG16 on our dataset. Then, we selected the last max pooling layer, which is the layer that comes after all convolution layers, to extract features. We added a flatten layer after the max pooling layer to handle the dimensionality issues, see Fig. 6. Because we use CNN 16 for the feature extraction part and not for classification, we discarded the fully connected and softmax layers, which are the dense layers after the flatten layer. After that, we used the extracted features to train the classification algorithms and develop models that can distinguish between COVID-19 and normal cases accurately. The number of extracted features is 25,088 features for every single image in the training dataset.

IV. RESULTS AND DISCUSSION

This section outlines the performance metrics we use to evaluate the developed models. Furthermore, it summarizes and discusses the main results of our experiments and presents a discussion related to previous studies.

A. Performance Metrics

We used various evaluation metrics to evaluate the proposed models. These metrics are as follows:

Confusion matrix: This is a technique that summarizes the performance of the classifier used. It presents true positive (TP) and true negative (TN) values, which means the number of correctly rated positive and negative instances. It also shows false positive (FP) and false negative (FN) values, which means the number of misclassified negative and positive instances [28].

Accuracy: Accuracy is the percentage of the test set

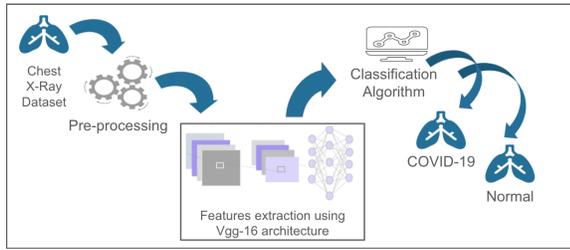


Fig. 4. An illustrative Diagram of the Proposed Hybrid Models.

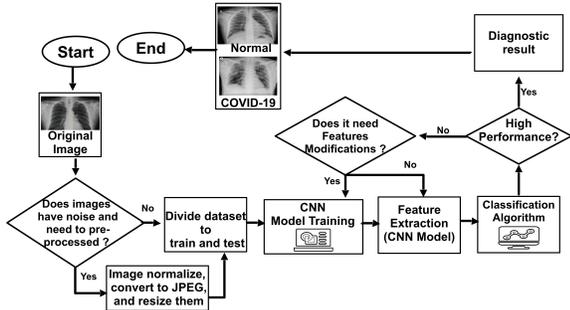


Fig. 5. Implementation Steps of the Hybrid CNN Models.

instances that are correctly classified by the classifier.

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (1)$$

Sensitivity (Recall): Recall is also referred to as the true positive rate, which means the proportion of positive cases that are correctly identified.

$$Sensitivity = \frac{TP}{TP+FN} \quad (2)$$

Specificity: This is the true negative rate. It refers to the proportion of negative instances that are correctly identified.

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

Precision: Precision is defined as the proportion of the true positives against all the positive results (both true positives and false positives).

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

F1 score: It represents the harmonic mean between recall and precision values.

$$F1score = \frac{2*Precision*Recall}{Precision+Recall} \quad (5)$$

B. Tests Results

In general, the results showed that all of our developed detection models are sufficient, especially the CNN and hybrid (CNN+SVM) models. This finding indicates that both stand-alone and hybrid models could achieve high performances when used to detect COVID-19 cases. Even though all our models are adequate, the CNN+XGBoost model is considered

Model: "sequential"

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 7, 7, 512)	14714688
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 4096)	102764544
dense_1 (Dense)	(None, 4096)	16781312
dense_2 (Dense)	(None, 2)	8194

Fig. 6. The Added Flatten Layer to the VGG-16 Architecture.

the least effective. The values of its accuracy, sensitivity, specificity, precision, and F1-score were 0.9873%, 0.9928%, 0.9874%, 0.9821%, 0.9819%, respectively.

Furthermore, the sensitivity of three models (CNN, CNN+RF, and CNN+SVM) was 100%; however, the CNN+RF model achieved less specificity compared to the two other models. The specificity of the CNN and CNN+SVM models was 0.9964% and it was 0.9855% for the CNN+RF model. Just like the specificity, the accuracy, precision, and F1-score of the CNN and CNN+SVM models were better than those of the two other models. The accuracy, precision, and F1-score of the CNN and CNN+SVM models were 0.9982%, 0.9964%, and 0.9982%, respectively. Also, the accuracy, precision, and F1-score of the CNN+RF model were 0.9928%, 0.9857%, and 0.9929%, respectively. Fig. 7 shows the confusion matrices of the presented models, and Table IV-A summarizes their performances.

Overall, there is not a substantial difference in the performance of our models; all of them are excellent in terms of their ability to detect COVID-19. Besides the stand-alone CNN model, two of the hybrid models, the CNN+SVM and CNN+RF, classified 100% of COVID-19 cases correctly. Nonetheless, there are two main limitations to the generalization of our findings. The first one is the lack of available Covid-19 datasets. The second limitation is that this study focused only on one of the CNN techniques and three machine learning classifiers because of the following constraints: Authors used the Google Colab platform to perform experiments due to the unavailability of powerful computing resources. However, the restrictions in the provided resources that the platform applies prevented them from conducting further investigations. Additionally, this research was not financially supported, and thus researchers only had limited Internet data, and they used personal computers with bounded processing power and memory capacity.

TABLE IV. PERFORMANCE METRICS FOR THE PRESENTED MODELS

	CNN	CNN-RF	CNN-XGBoost	CNN-SVM
Accuracy	0.9982	0.9928	0.9873	0.9982
Sensitivity(Recall)	1.0000	1.0000	0.9928	1.0000
Specificity	0.9964	0.9855	0.9874	0.9964
Precision	0.9964	0.9857	0.9821	0.9964
F1-score	0.9982	0.9929	0.9819	0.9982

C. Discussion Related to Previous Studies

This subsection compares our models with pre-existing models presented in previous studies. These prior models

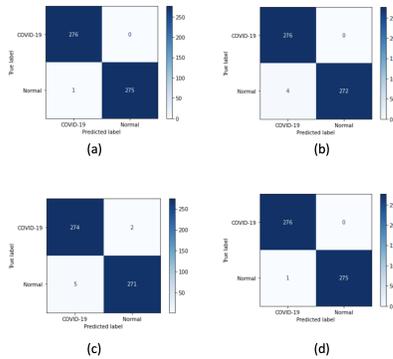


Fig. 7. Generated Confusion Matrices for a) Stand-alone CNN Model, b) CNN-RF model c) CNN-XGBoost Model, d) CNN-SVM Model

were developed using chest X-ray images and the VGG-16 model as a feature extractor or a classifier. Table V shows the performance for our models as compared to previous models.

In [8], the authors presented a modified CNN based model to detect COVID-19. The model accuracy and sensitivity before modification were 90.19% and 94.16%, respectively. After modifying the last layer of the original model, its performance improved. Its accuracy and sensitivity became 99.52% and 97.93%, respectively. Still, the accuracy and sensitivity of our models are higher. In [5], even though the researchers have used the CLAHE method to enhance their detection model, its performance is still low. Its accuracy was 83%, and then became 92%.

Compared to previous models, a better model has been presented in [3]. Its accuracy was 99.3% and its sensitivity was 99.28%. Our stand-alone model outperformed it. Some factors that lead to that improvement in ours is that the number of COVID-19 images that we used to develop our model is larger compared to these studies, and we have used a balanced dataset as well.

Just like the stand-alone CNN model, our hybrid models outperform those that have been proposed in [8]. Two out of the three classification algorithms that have been used in that paper are different from ours. Authors have selected the following algorithms: SVM, Bagging, and AdaBoost, while we have selected these classification algorithms: SVM, random forest, and XGBoost. As shown in Table V, the SVM is the classifier that leads to developing models with high accuracy in both studies. Our two other algorithms, however, surpassed Bagging and AdaBoost.

V. CONCLUSION

This study addressed some of the challenges of the traditional COVID-19 test method. It exploited the power of machine learning to accelerate the process of detecting COVID-19 and to enhance its efficiency. We investigated the effectiveness of stand-alone CNN models and hybrid machine learning models in detecting the disease. We combined five chest X-Ray images datasets to develop four COVID-19 detection models: a stand-alone CNN model and three hybrid machine learning models. As a comparison to some of the previous studies that have been published in the early few months of the pandemic,

the count of chest X-Ray images that we used to develop our models is considered one of the biggest.

TABLE V. COMPARISON WITH STATE-OF-THE-ART MODELS

Ref.	Model	Accuracy	Sensitivity	Specificity	Precision	F1-Score
[5]	VGG-16	92%	94.92%	92%	-	-
[3]	VGG-16	99.33%	99.28%	99.38%	-	99.28%
[8]	VGG-16	90.19%	94.16%	-	94.12%	94.14%
	VGG-16 & SVM	91.19%	93.15%	-	94.12%	93.63%
	VGG-16 & Bagging	90.19%	92.16%	-	92.22%	92.19%
	VGG-16 & Ada Boot	90.19%	90.10%	-	89.16%	89.63%
Our study	VGG-16	99.82%	100%	99.64%	99.64%	99.82%
	VGG-16 & RF	99.28%	100%	98.55%	98.57%	99.29%
	VGG-16 & XGBoost	98.73%	99.28%	98.74%	98.21%	98.19%
	VGG-16 & SVM	99.82%	100%	99.64%	99.64%	99.82%

Our findings illustrate that all of the four proposed models are effective in detecting COVID-19. The lowest detection accuracy obtained was 98.73% which is the accuracy of the VGG16+XGBoost model. The highest accuracy was 99.82% which is the accuracy of both VGG-16 and(VGG16+SVM) models. Furthermore, one of the most promising findings is that the sensitivity of the VGG-16, (VGG16+SVM), and (VGG16+RF) models is 100%, meaning they have a zero false negative case rate. That means that from all the examined cases, 100% of the COVID-19 positive images were detected. This finding plays an important role in reducing the possibility of spreading the virus to more people.

VI. FUTURE WORK

Several experiments can be conducted to study the effect of different CNN architectures and optimizers on the model's performance. Furthermore, one of the areas that warrants additional study are COVID-19 mutations. It is necessary to develop machine learning and deep learning-based models capable of detecting new versions of COVID-19, such as B.1.1.7, B.1.1.207, P.1 and B.1.525 automatically. Additionally, it is important to build robust models that have the ability to distinguish between SARS, MERS, and COVID-19 accurately.

REFERENCES

- [1] Z. Y. Zu, M. D. Jiang, P. P. Xu, W. Chen, Q. Q. Ni, G. M. Lu, and L. J. Zhang, "Coronavirus disease 2019 (COVID-19): A perspective from china," *Radiology*, vol. 296, no. 2, pp. E15–E25, 2020.
- [2] L. Duran-Lopez, J. P. Dominguez-Morales, J. Corral-Jaime, S. Vicente-Diaz, and A. Linares-Barranco, "Covid-xnet: a custom deep learning system to diagnose and locate covid-19 in chest x-ray images," *Applied Sciences*, vol. 10, no. 16, p. 5683, 2020.
- [3] I. U. Khan and N. Aslam, "A deep-learning-based framework for automated diagnosis of covid-19 using x-ray images," *Information*, vol. 11, no. 9, p. 419, 2020.
- [4] L. O. Hall, R. Paul, D. B. Goldgof, and G. M. Goldgof, "Finding covid-19 from chest x-rays using deep learning on a small dataset," *arXiv preprint arXiv:2004.02060*, 2020.
- [5] F. A. Saiz and I. Barandiaran, "Covid-19 detection in chest x-ray images using a deep learning approach," *International Journal of Interactive Multimedia and Artificial Intelligence, InPress (InPress)*, vol. 1, 2020.
- [6] S. Vaid, R. Kalantar, and M. Bhandari, "Deep learning covid-19 detection bias: accuracy through artificial intelligence," *International Orthopaedics*, vol. 44, pp. 1539–1542, 2020.
- [7] K. F. Haque and A. Abdelgawad, "A deep learning approach to detect covid-19 patients from chest x-ray images," *AI*, vol. 1, no. 3, pp. 418–435, 2020.
- [8] R. K. Dubey *et al.*, "Deep learning based hybrid models for prediction of covid-19 using chest x-ray," *TechRxiv*, 2020.

- [9] K. Elasnoui and Y. Chawki, "Using x-ray images and deep learning for automated detection of coronavirus disease," *Journal of Biomolecular Structure and Dynamics*, no. just-accepted, pp. 1–22, 2020.
- [10] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, p. 1, 2020.
- [11] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deepcovid: Predicting covid-19 from chest x-ray images using deep transfer learning," *arXiv preprint arXiv:2004.09363*, 2020.
- [12] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.
- [13] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," *arXiv e-prints*, p. arXiv:2003.11597, Mar. 2020.
- [14] "Actualmed covid-19 chest x-ray dataset," 2020. Accessed on: Feb. 13, 2021. [Online]. Available: <https://github.com/agchung/Actualmed-COVID-chestxray-dataset>.
- [15] "Figure1-covid-chestxray-dataset," 2020. Accessed on: Feb. 13, 2021. [Online]. Available: <https://github.com/agchung/Figure1-COVID-chestxray-dataset>.
- [16] "Covid-19-image-repository," 2020. Accessed on: Feb. 13, 2021. [Online]. Available: <https://github.com/ml-workgroup/covid-19-image-repository>.
- [17] P. Mooney, "Chest x-ray images (pneumonia)," 2020. Accessed on: Feb. 13, 2021. [Online]. Available: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
- [18] "Keras documentation," accessed on: February. 20, 2021. [Online]. Available: <https://keras.io/about/>.
- [19] "Google colabory," accessed on: February. 20, 2021. [Online]. Available: <https://colab.research.google.com/notebooks/intro.ipynb>.
- [20] "Tensorflow," accessed on: February. 20, 2021. [Online]. Available: <https://github.com/tensorflow/tensorflow>.
- [21] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into imaging*, vol. 9, no. 4, pp. 611–629, 2018.
- [22] V. Phung and E. Rhee, "A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets," *Appl. Sci. (Basel)*, vol. 9, no. 21, p. 4500, 2019.
- [23] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020.
- [24] Z. Li, W. Yang, S. Peng, and F. Liu, "A survey of convolutional neural networks: analysis, applications, and prospects," *arXiv preprint arXiv:2004.02806*, 2020.
- [25] S. Tammina, "Transfer learning using vgg-16 with deep convolutional neural network for classifying images," *International Journal of Scientific and Research Publications*, vol. 9, no. 10, pp. 143–150, 2019.
- [26] P. Mooney, "Classification algorithm - an overview — sciencedirect topics," 2021. Accessed on: April. 28, 2021. [Online]. Available: <http://www.sciencedirect.com/topics/engineering/classification-algorithm>.
- [27] O. D. Gülgün and E. Hamza, "Medical image classification with hybrid convolutional neural network models," *Bilgisayar Bilimleri ve Teknolojileri Dergisi*, vol. 1, no. 1, pp. 28–41.
- [28] M. Hossain and M. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, no. 2, p. 1, 2015.

A Self-adaptive Algorithm for Solving Basis Pursuit Denoising Problem

Mengkai Zhu¹, Xu Zhang², Bing Xue³, Hongchun Sun⁴
School of Mathematics and Statistics
Linyi University
Linyi, China, 276005

Abstract—In this paper, we further consider a method for solving the basis pursuit denoising problem (BPDP), which has received considerable attention in signal processing and statistical inference. To this end, a new self-adaptive algorithm is proposed, its global convergence results is established. Furthermore, we also show that the method is sublinearly convergent rate of $O(\frac{1}{k})$. Finally, the availability of given method is shown via some numerical examples.

Keywords—Basis pursuit denoising problem; algorithm; global convergence; sublinearly convergent rate; sparse signal recovery

I. INTRODUCTION

The basis pursuit denoising problem (BPDP) is considered to be an important issue encountered in the fields of signal processing and statistical inference, which is to find a sparse signal $\bar{x} \in R^n$ from linear system $z = A\bar{x}$, and can be mathematically depicted as the following

$$\min_{x \in R^n} F(x) := f(x) + \rho\varphi(x), \quad (1)$$

where $f(x) := \frac{1}{2}\|Ax - z\|_2^2$, $\varphi(x) = \|x\|_1$, $A \in R^{m \times n}$ ($m \ll n$), $\rho > 0$ is a parameter, and the ℓ_1 -norm and ℓ_2 -norm of the vector x are defined by $\|x\|_1 = \sum_{i=1}^n |x_i|$ and $\|x\|_2 = (\sum_{i=1}^n x_i^2)^{1/2}$, respectively. In addition, we denote the solution set of the problem (1) by Ω^* , and $\Omega^* \neq \emptyset$.

Clearly, (1) is an unconstrained convex optimization problem, and some standard algorithms such as the Newton-type algorithms or the conjugate gradient methods to solve it. But, these methods are not suitable for large-scale cases of BPDP, and it even become invalidation as n increases. In recent years, there are a lot of algorithms for solving BPDP have been extensively developed since its appearance. He and Cai et al([1]) introduce a splitting method (MPRSM) for solving Dantzig selector problem, and the BPDP is a special case of this problem. Based on this theory, Sun and Liu et al ([2]) further investigate MPRSM for BPDP, and regularize its first subproblem by the proximal regularization. Yang and Zhang ([3]) investigate alternating direction methods for several ℓ_1 -norm minimization, including the basis pursuit problem, the basis-pursuit denoising problems, and so on. Yu et al.([4]) apply the primal Douglas-Rachford splitting method to solve equivalent transformation form of BPDP. In [5], the authors proposed some efficient methods to solve ℓ_1 -norm minimization problems, and are used in BPDP. Zhang and Sun ([6]) presented projection-type method to solve BPDP, its global convergence results of the new algorithm is established. BPDP can be transformed into a smooth optimization problem

by some splitting technique equivalent. some iterative algorithms which can solve smooth optimization problem are applicable to this problem. Xiao and Zhu ([7]) transformed BPDP into a convex constrained monotone equations, and presented a conjugate gradient method for the equivalent the forms. Sun and Tian ([8]) give a derivative-free conjugate gradient projection algorithms for non-smooth equations with convex constraints. Sun et al. ([9]) reformulated BPDP as variational inequality problem, and proposed a novelly inverse matrix-free proximal point algorithm. Base on the same transformation of ([9]), Feng and Wang ([10]) also proposed a projection-type algorithm. Although there are so many ways to solve it, the solving speed and accuracy are still need improved. In the paper, we further consider a new self-adaptive method to solve BPDP, which this method is sublinearly convergent rate, the motivation behind this is for the better numerical performance when the dimension increases.

The rest of this paper is organized below. In Section 2, some related properties are given, which are the basis of our analysis. We present a new self-adaptive algorithm with Armijo-like line search to solve BPDP, and show that this method is global convergence in detail. Furthermore, the sublinearly convergent rate of $O(\frac{1}{k})$ is presented. In Section 3, we give some numerical experiments on BPDP for sparse signal recovery to show availability of the presented algorithm. Finally, some results are described in Section 4.

In the end of this section, we give some notations used in this paper. Use \mathcal{R}^N to denote an N -dimensional Euclidean space with the standard inner product. For vectors $x, y \in \mathcal{R}^M$, we use $\langle x, y \rangle$ to denote the standard inner product. We denote the standard l_1 -norm and l_2 -norm by $\|\cdot\|_1$ and $\|\cdot\|$, respectively.

II. ALGORITHM AND CONVERGENCE

In this section, we will present a new iterative algorithm with Armijo-like line search to solve BPDP, and the global convergence and sublinearly convergent rate of new algorithm is proved in detail. To this end, we give some needed preliminaries which will be used in the sequel.

Definition 2.1: ([11]) Set $f : \mathcal{R}^N \rightarrow \mathcal{R}$ be convex. The subdifferential of f at x is defined as

$$\partial f(x) = \{\xi \in \mathcal{R}^N | f(y) \geq f(x) + \langle \xi, y - x \rangle, \forall y \in \mathcal{R}^N\}.$$

For $F(x)$ involved in (1), we establish quadratic approxi-

mation of $F(x)$ below:

$$Q_L(x, y) := \left[f(y) + \langle x - y, \nabla f(y) \rangle + \frac{L}{2} \|x - y\|^2 \right] + \rho\varphi(x), \quad \forall L > 0. \quad (1)$$

Let

$$p_L(y) = \arg \min_{x \in R^n} \{Q_L(x, y)\}, \quad (2)$$

and (2) can be further written as:

$$p_L(y) = \arg \min \left\{ \rho\varphi(x) + \frac{L}{2} \left\| x - \left(y - \frac{1}{L} \nabla f(y) \right) \right\|^2 \right\}. \quad (3)$$

Next, we recall Lemma 2.1 below, which is fundamental property for smooth function in the class $C^{1,1}$. It will be crucial for the convergence analyses of our algorithm below.

Lemma 2.1: (Lemma 3.2. [12],[13]) Let $f : \mathcal{R}^N \rightarrow \mathcal{R}$ be a continuously differentiable function with constant L_f . Then, for any $L \geq L_f$, one has

$$f(x) \leq [f(y) + \langle x - y, \nabla f(y) \rangle + \frac{L}{2} \|x - y\|^2], \forall x, y \in \mathcal{R}^N. \quad (4)$$

From Lemma 2.1, if $L \geq L_f$, then for any $y \in R^n$, one has

$$\begin{aligned} F(p_L(y)) &= f(p_L(y)) + \rho\varphi(p_L(y)) \\ &\leq [f(y) + \langle p_L(y) - y, \nabla f(y) \rangle \\ &\quad + \frac{L}{2} \|p_L(y) - y\|^2] + \rho\varphi(p_L(y)) \\ &= Q_L(p_L(y), y). \end{aligned} \quad (5)$$

Now, we formally state our algorithm for model (1) as follows.

Algorithm 2.1.

Step0. Choose $\beta > 0$, $\epsilon \geq 0$, some $\eta > 1$, $\gamma \in (0, 1)$, and an initial point $x^0 \in R^n$, and let $k := 0$.

Step1. For $k = 1, 2, \dots$, update the next iterate x^k via

$$(x^k)_i = \begin{cases} (\theta_k)_i + \frac{\sigma}{L_k} & \text{if } (\theta_k)_i < -\frac{\sigma}{L_k} \\ 0 & \text{if } |(\theta_k)_i| \leq \frac{\sigma}{L_k} \\ (\theta_k)_i - \frac{\sigma}{L_k} & \text{if } (\theta_k)_i > \frac{\sigma}{L_k} \end{cases} \quad (6)$$

where $\theta_k = x^{k-1} - \frac{1}{L_k} \nabla f(x^{k-1})$, $L_k = \eta^{m_k} \beta$, and m_k is the smallest integer $m \geq 0$ such that

$$f(x^k) \leq f(x^{k-1}) + \langle x^k - x^{k-1}, \nabla f(x^{k-1}) \rangle + \frac{\eta^{m_k} \beta}{2} \|x^k - x^{k-1}\|^2. \quad (7)$$

Step2. If $\|x^k - x^{k-1}\| \leq \epsilon$, stop. Then, x^k is a solution of (1). Otherwise, go to Step 1 with $k \triangleq k + 1$.

Remark 2.1: By the subdifferential of the absolute value function $|t|$, which be given as follows:

$$\partial(|t|) = \begin{cases} -1 & \text{if } t < 0, \\ [-1, 1] & \text{if } t = 0, \\ 1 & \text{if } t > 0. \end{cases}$$

Combining this with (7), we obtain the following results.

If $(\theta_k)_i < -\frac{\sigma}{L_k}$, then $(x^k)_i = (\theta_k)_i + \frac{\sigma}{L_k} < 0$, i.e., $(x^k)_i = (\theta_k)_i - \frac{\sigma}{L_k} \frac{\partial\varphi(x^k)}{\partial x_i}$, where $\frac{\partial\varphi(x^k)}{\partial x_i} = -1$.

If $(\theta_k)_i > \frac{\sigma}{L_k}$, then $(x^k)_i = (\theta_k)_i - \frac{\sigma}{L_k} > 0$, i.e., $(x^k)_i = (\theta_k)_i - \frac{\sigma}{L_k} \frac{\partial\varphi(x^k)}{\partial x_i}$, where $\frac{\partial\varphi(x^k)}{\partial x_i} = 1$.

If $|(\theta_k)_i| \leq \frac{\sigma}{L_k}$, then $|(\theta_k)_i| / \frac{\sigma}{L_k} \leq 1$. From $(x^k)_i = 0$, one has $(x^k)_i = 0 = (\theta_k)_i + \frac{\sigma}{L_k} (\theta_k)_i / \frac{\sigma}{L_k} = (\theta_k)_i - \frac{\sigma}{L_k} \frac{\partial\varphi(x^k)}{\partial x_i}$, where $\frac{\partial\varphi(x^k)}{\partial x_i} = (\theta_k)_i / \frac{\sigma}{L_k}$. By the above analysis, we have

$$x^k = \theta_k - \frac{\sigma}{L_k} \partial\varphi(x^k),$$

i.e.,

$$\sigma \partial\varphi(x^k) + \nabla f(x^{k-1}) + L_k(x^k - x^{k-1}) = 0. \quad (8)$$

Remark 2.2: Combining (7) with Lemma 2.2, we know that

$$\eta^{m_k} \beta = L_k \geq L_f = \|A^\top A\| \quad (9)$$

for some m . In addition, we know that L_k/η must violate (7), i.e., $L_k < \eta \|A^\top A\|$. Thus, we obtain

$$\|A^\top A\| \leq L_k < \eta \|A^\top A\|. \quad (10)$$

Using $L_k = \eta^{m_k} \beta$ and $\eta > 1$, one has $\beta < L_k$ for every $k \geq 1$. Hence, $\beta < \eta \|A^\top A\|$.

Next, we will discuss global convergence results and sub-linearly convergent rate of the proposed method. To this end, we present some lemmas below.

Lemma 2.2: The sequence $\{F(x^k)\}$ generated by Algorithm 2.1 is non-increasing.

Proof: For any $k \geq 1$, we have

$$\begin{aligned} F(x^k) &= f(x^k) + \rho\varphi(x^k) \\ &\leq f(x^{k-1}) + \langle x^k - x^{k-1}, \nabla f(x^{k-1}) \rangle \\ &\quad + \frac{L_k}{2} \|x^k - x^{k-1}\|^2 + \rho\varphi(x^k) \\ &= Q_{L_k}(x^k, x^{k-1}) \\ &\leq Q_{L_k}(x^{k-1}, x^{k-1}) \\ &= F(x^{k-1}), \end{aligned} \quad (11)$$

where the first inequality is obtained by using (4) with $y = x^{k-1}$, $x = x^k$ and $L = L_k$, the second inequality follows from (2) and (6). Thus, the desired result follows.

Lemma 2.3: For any $x \in R^n$, and $k \geq 1$, we have

$$F(x) - F(x^k) \geq \frac{L_k}{2} \|x^k - x^{k-1}\|^2 + L_k \langle x - x^{k-1}, x^{k-1} - x^k \rangle. \quad (12)$$

Proof: By (5), one has

$$F(p_{L_k}(x^{k-1})) \leq Q_{L_k}(p_{L_k}(x^{k-1}), x^{k-1}). \quad (13)$$

By (6), one has $x^k = p_{L_k}(x^{k-1})$. Combining this with (13), we obtain

$$F(x) - F(x^k) \geq F(x) - Q_{L_k}(x^k, x^{k-1}). \quad (14)$$

Since f and φ are convex, we can deduce

$$\begin{aligned} F(x) &= f(x) + \sigma\varphi(x) \\ &\geq [f(x^{k-1}) + \langle x - x^{k-1}, \nabla f(x^{k-1}) \rangle] \\ &\quad + [\sigma\varphi(x^k) + \sigma\langle x - x^k, \xi^k \rangle], \end{aligned} \quad (15)$$

where $\xi^k \in \partial\varphi(x^k)$. Using (1) with $x = x^k, y = x^{k-1}$, one has

$$\begin{aligned} Q_{L_k}(x^k, x^{k-1}) &= [f(x^{k-1}) + \langle x^k - x^{k-1}, \nabla f(x^{k-1}) \rangle] \\ &\quad + \frac{L_k}{2} \|x^k - x^{k-1}\|^2 + \sigma\varphi(x^k). \end{aligned} \quad (16)$$

Applying (15), (16) and (14), we have

$$\begin{aligned} F(x) - F(x^k) &\geq F(x) - Q_{L_k}(x^k, x^{k-1}) \\ &\geq \langle x - x^k, \nabla f(x^{k-1}) + \sigma\partial\varphi(x^k) \rangle \\ &\quad - \frac{L_k}{2} \|x^k - x^{k-1}\|^2 \\ &= L_k \langle x - x^k, x^{k-1} - x^k \rangle - \frac{L_k}{2} \|x^k - x^{k-1}\|^2 \\ &= L_k \langle (x - x^{k-1}) + (x^{k-1} - x^k), x^{k-1} - x^k \rangle \\ &\quad - \frac{L_k}{2} \|x^k - x^{k-1}\|^2 \\ &= L_k \langle x - x^{k-1}, x^{k-1} - x^k \rangle + \frac{L_k}{2} \|x^k - x^{k-1}\|^2. \end{aligned} \quad (17)$$

where the first equality holds by (8).

Theorem 2.1: Suppose that x^* be an arbitrary solution of (1), and $\{x^k\}$ be sequence generated by Algorithm 2.1. Then, for any $k \geq 1$, one has

$$F(x^k) - F(x^*) \leq \frac{\eta \|A^\top A\|}{2k} \|x^* - x^0\|^2. \quad (18)$$

Proof: Applying Lemma 2.3 with $x = x^*, k = m$, one has

$$\begin{aligned} &\frac{2}{L^m} (F(x^*) - F(x^m)) \\ &\geq \|x^m - x^{m-1}\|^2 + 2 \langle x^* - x^{m-1}, x^{m-1} - x^m \rangle \\ &= \|x^m - x^{m-1}\|^2 \\ &\quad + \langle (x^* - x^m) + (x^m - x^{m-1}), x^{m-1} - x^m \rangle \\ &\quad + \langle x^* - x^{m-1}, (x^{m-1} - x^*) + (x^* - x^m) \rangle \\ &= \langle x^* - x^m, x^{m-1} - x^m \rangle \\ &\quad + \langle x^* - x^{m-1}, x^{m-1} - x^* \rangle + \langle x^* - x^{m-1}, x^* - x^m \rangle \\ &= \langle x^* - x^m, (x^{m-1} - x^*) + (x^* - x^m) \rangle \\ &\quad + \langle x^* - x^{m-1}, x^{m-1} - x^* \rangle + \langle x^* - x^{m-1}, x^* - x^m \rangle \\ &= \|x^* - x^m\|^2 - \|x^* - x^{m-1}\|^2. \end{aligned} \quad (19)$$

Since x^* be a solution of (1), then one has $F(x^*) - F(x^k) \leq 0$. Combining this with (10), we obtain

$$\begin{aligned} &\frac{2}{\eta \|A^\top A\|} (F(x^*) - F(x^m)) \\ &\geq \frac{2}{L^m} (F(x^*) - F(x^m)) \\ &\geq \|x^* - x^m\|^2 - \|x^* - x^{m-1}\|^2. \end{aligned} \quad (20)$$

where where the second inequality is by (19). By (20), we can deduce

$$\frac{2}{\eta \|A^\top A\|} [kF(x^*) - \sum_{m=1}^k F(x^m)] \geq \|x^* - x^k\|^2 - \|x^* - x^0\|^2. \quad (21)$$

Applying (12) with $x = x^{m-1}, k = m$, we deduce

$$\frac{2}{L_m} (F(x^{m-1}) - F(x^m)) \geq \|x^m - x^{m-1}\|^2. \quad (22)$$

By Lemma 2.2, one has $F(x^{m-1}) - F(x^m) \geq 0$, combining (22) with (10), we obtain

$$\frac{2}{\|A^\top A\|} (F(x^{m-1}) - F(x^m)) \geq \|x^m - x^{m-1}\|^2, \quad (23)$$

i.e.,

$$\begin{aligned} &\frac{2}{\|A^\top A\|} [(m-1)F(x^{m-1}) - mF(x^m) + F(x^m)] \\ &\geq (m-1) \|x^m - x^{m-1}\|^2. \end{aligned} \quad (24)$$

By (24), we can deduce

$$\begin{aligned} &\frac{2}{\eta \|A^\top A\|} [-kF(x^k) + \sum_{m=1}^k F(x^m)] \\ &\geq \frac{1}{\eta} \sum_{m=1}^k (m-1) \|x^m - x^{m-1}\|^2. \end{aligned} \quad (25)$$

Adding (21) and (25), we have

$$\begin{aligned} &\frac{2k}{\eta \|A^\top A\|} (F(x^*) - F(x^k)) \\ &\geq \|x^* - x^k\|^2 \\ &\quad + \frac{1}{\eta} \sum_{m=1}^k (m-1) \|x^m - x^{m-1}\|^2 - \|x^* - x^0\|^2 \\ &\geq -\|x^* - x^0\|^2. \end{aligned} \quad (26)$$

Combining this with the fact $F(x^*) - F(x^k) \leq 0$, we obtain

$$\frac{2k}{\eta \|A^\top A\|} (F(x^k) - F(x^*)) \leq \|x^* - x^0\|^2. \quad (27)$$

Thus, the desired result follows.

Remark 2.3: Theorem 2.1 indicates that we can obtain an ϵ -optimal solution, denoted by \bar{x} , and requires the number of iterations at most $\lceil c/\epsilon \rceil$ such that $F(\bar{x}) - F(x^*) \leq \epsilon$, where $c = \frac{\eta \|A^\top A\| \|x^0 - x^*\|^2}{2}$.

Theorem 2.2: Suppose that Ω^* is bounded. Then, the $\{x^k\}$ generated by Algorithm 2.1 converges globally to a solution of (1).

Proof: By (11), using $F(x) \geq 0$, we know that $\{F(x^k)\}$ be convergent. Combining this with (23), one has

$$\lim_{k \rightarrow \infty} \|x^k - x^{k-1}\| = 0. \quad (28)$$

Applying (19) and the fact $F(x^*) - F(x^k) \leq 0$, we have

$$\|x^* - x^k\| \leq \|x^* - x^{k-1}\|. \quad (29)$$

By (29), then the nonnegative sequence $\{\|x^k - x^*\|\}$ is decreasing, so it converges. Since the solution set of (1) is bounded. Thus, $\{x^k\}$ is bounded, and let $\{x^{k_i}\}$ be a subsequence of $\{x^k\}$ and converges toward \bar{x} , combining this with (28), one has

$$\begin{aligned} \lim_{i \rightarrow \infty} \|x^{k_i-1} - \bar{x}\| &\leq \lim_{i \rightarrow \infty} \|x^{k_i} - x^{k_i-1}\| \\ &\quad + \lim_{i \rightarrow \infty} \|x^{k_i} - \bar{x}\| \\ &= 0. \end{aligned} \quad (30)$$

From (8), one has $\sigma\partial\varphi(x^{k_i}) + \nabla f(x^{k_i-1}) + L_{k_i}(x^{k_i} - x^{k_i-1}) = 0$. Combining this with (30) and (28), one has

$$\begin{aligned} \|\sigma\partial\varphi(\bar{x}) + \nabla f(\bar{x})\| &= \lim_{i \rightarrow \infty} \|\sigma\partial\varphi(x^{k_i}) + \nabla f(x^{k_i-1})\| \\ &= \lim_{i \rightarrow \infty} L_{k_i} \|x^{k_i} - x^{k_i-1}\| \\ &= 0. \end{aligned} \quad (31)$$

Since the function $F(x)$ be convex, combining this with (31), we have \bar{x} is a solution of (1). As a result, the \bar{x} can be used as x^* to discussion of Theorem 2.1 above. Thus, we obtain that the sequence $\{\|x^k - \bar{x}\|\}$ also converges, combining $\lim_{i \rightarrow \infty} \|x^{k_i} - \bar{x}\| = 0$, we have $\lim_{k \rightarrow \infty} \|x^k - \bar{x}\| = 0$. i.e. $\{x^k\}$ converges globally toward \bar{x} .

Remark 2.4: By (28), we know that the termination criteria in Step2 of Algorithm 2.1 is reasonable.

III. NUMERICAL RESULTS

In this section, we present some numerical experiments about BPDP to show availability for Algorithm 2.1. All of codes are written by MATLAB 9.2.0.538062 and performed on a Windows 10 PC with an AMD FX-7500 Radeon R7,10 Computer Cores 4C+6G CPU, 2.10GHz CPU and 8GB of memory. In these experiments, let

$$\mu = 0.001, n = 2^{11}, m = \text{floor}(n/4), k = \text{floor}(m/8),$$

and sensing matrix A is generated by MATLAB scripts below:

$$[Q, R] = \text{qr}(A', 0); A = Q'.$$

The initial signal \bar{x} is generated by

$$p = \text{randperm}(n); x(p(1:k)) = \text{randn}(k, 1).$$

We set the stop criterion is

$$\frac{\|F_k - F_{k-1}\|}{\|F_{k-1}\|} < 10^{-10},$$

where $F_k = F(x_k)$. The relative error is calculated by

$$\text{RelErr} = \frac{\|\hat{x} - \bar{x}\|}{\|\bar{x}\|},$$

where the recovery signal be denoted by \hat{x} .

A. Test on additive Gaussian white Noise

In this subsection, apply Algorithm 2.1 to recover a simulated sparse signal which observation data is corrupted by additive Gaussian white noise. We set

$$n = 2^{11}, m = 2^9, k = 2^6.$$

The original signal, the measurement and the reconstructed signal (marked by red point) by Algorithm 2.1 are given in Fig.1. Obviously, from the first and the third plots in Fig.1, all elements in the original signal are circled by the red points, which indicates that the Algorithm 2.1 can recover the original signal quite well.

On the other hand, use a same technique in [8] to create another type of matrix A . Using the parameters above, the original signal, the measurement and the reconstructed signal (marked by red point) by the Algorithm 2.1 is given in Fig.2. It can be concluded that our algorithm is can also reconstruct the original signal in [8].

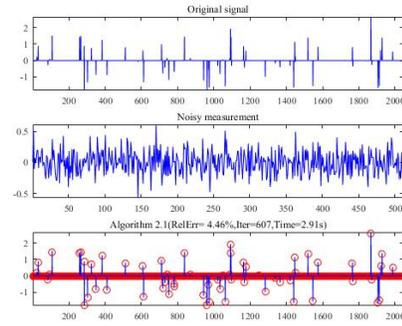


Fig. 1. Signal Recovery Result.

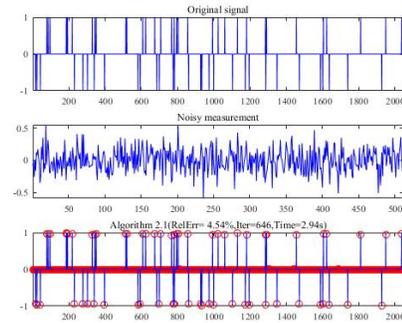


Fig. 2. Signal Recovery Result (Matrix A is Generated by [8]).

B. Compare with Different k -Sparse Signal ($n = 2^{11}, m = 2^9$)

In this subsection, we compare the CPU Time and the RelErr among Algorithm 2.1, Algorithm 3.1 in [2] (PPRSM) and Algorithm 3.1 in [10] (LAPM). All algorithms have run 5 times, respectively, and the average of the the CPU Time and the RelErr are obtained. Some parameters about algorithms above are listed as follows:

Algorithms 2.1: $\beta = 4, \eta = 3, \sigma = 0.01, \gamma = 0.5$.
PPRSM: $\gamma = 0.2, \sigma = 0.1$.
LAPM: $\beta = 0.25; \tau = 0.6$;

The numerical results are listed in Table I. From the table, we can drive that CPU time of Algorithm 2.1 are obviously less than other algorithms in different k -Sparse signal whether it is Free noise or Gaussian noise. In addition, we can not only know that the running speed is faster than other algorithms, but also that our algorithm is more accurate than other algorithms, which shows that Algorithm 2.1 is batter than PPRSM and LAPM.

IV. CONCLUSION

In this paper, we consider a new self-adaptive method to solve the basis pursuit denoising problem (BPDP), which has received considerable attention in signal processing and statistical inference. Global convergence result of this method is given in detail. Furthermore, the global sublinearly convergent rate of the method also is shown. Finally, some numerical

TABLE I. COMPARE WITH DIFFERENT K-SPARSE SIGNAL WITH FREE NOISE OR GAUSSIAN NOISE

k-Sparse signal	Methods	No noise		Gaussian noise	
		CPU Time	RelErr	CPU Time	RelErr
60	Algorithm 2.1	2.4215	4.6268	2.2801	4.6119
	PPRSM	3.8147	4.6324	3.9575	4.9572
	LAPM	5.9058	4.4218	6.8491	4.6555
80	Algorithm 2.1	3.8112	4.1989	3.4419	4.3016
	PPRSM	5.0971	4.4387	5.4456	4.9157
	LAPM	8.6797	4.4898	9.0344	4.3922
100	Algorithm 2.1	4.1165	4.8924	4.3346	4.5957
	PPRSM	6.7260	4.7547	7.3171	4.0318
	LAPM	11.7922	4.6787	12.6948	4.7655

results illustrate that this algorithm is valid for the given tests on sparse signal recovery.

According to its limitations, this work has several possible extensions. Firstly, the parameters of Algorithm 3.1 is adjusted dynamically to further enhance the efficiency of the corresponding method. Secondly, we may established error bound for (1) just as was done for GLCP in [14], [15], [16], [17], and may use the error bound estimation to establish quick convergence rate of the new Algorithm for solving (1). This is a topic for future research.

ACKNOWLEDGMENT

The authors gratefully acknowledge the valuable comments of the editor and the anonymous reviewers. This work was supported by Natural Science Foundation of China (Nos. 12071250, 11801309), and Shandong Province college students' innovation and entrepreneurship training program (S202010452115).

REFERENCES

[1] H. He, X. Cai, D. Han, "A fast splitting method tailored for Dantzig selector," Computational Optimization and Applications, vol.62, no.2, pp. 347-372, May, 2015.
[2] M. Sun, J. Liu, "A proximal Peaceman-Rachford splitting method for compressive sensing," Journal of Applied Mathematics and Computing, vol.50, no. 1-2, pp.349-363, February, 2016.
[3] J.F. Yang, Y. Zhang, "Alternating direction algorithms for ℓ_1 -problems in compressive sensing," SIAM Journal on Scientific Computing, vol.33, no. 1, pp.250-278, February, 2011.

[4] Y.C. Yu, J.G. Peng, X.L. Han and A.A. Cui, "A Primal Douglas-Rachford Splitting Method for the Constrained Minimization Problem in Compressive Sensing," Circuits, Systems, and Signal Processing, vol.36, no.10, pp.4022-4049, January, 2017.
[5] W. Yin, S. Osher, D. Goldfarb, J. Darbon, "Bregman iterative algorithms for ℓ_1 - minimization with applications to compressed sensing," SIAM Journal on Imaging Sciences, vol.1, no.1, pp.143-168, March, 2008.
[6] B.H. Zhang, H.C. Sun, "A New Projection Type Algorithm for Compressive Sensing," Journal of Advances in Mathematics and Computer Science, vol.30, no.1, pp.1-11, December, 2018.
[7] Y.H. Xiao, H. Zhu, "A conjugate gradient method to solve convex constrained monotone equations with applications in compressive sensing," Journal of Mathematical Analysis and Applications, vol.405, no.1, pp.310-319, September, 2013.
[8] M. Sun, M.Y. Tian, "A Class of Derivative-Free CG Projection Methods for Nonsmooth Equations with an Application to the LASSO Problem," Bulletin of the Iranian Mathematical Society, vol.46, pp.183-205, February, 2020.
[9] H.C. Sun, M. Sun, B.H. Zhang "An inverse matrix-free proximal point algorithm for compressive sensing," ScienceAsia, vol.44, no.5, pp.311-318, September, 2018.
[10] D.X. Feng, X.Y. Wang, "A linearly convergent algorithm for sparse signal reconstruction," J. Fixed Point Theory Appl, vol.20, pp.1-13, October, 2018.
[11] Y. Gao, Nonsmooth Optimization (in Chinese), Science Press, Beijing, 2008.
[12] Y. Nesterov, "A method for solving a convex programming problems with convergence rate $O(1/k^2)$," Soviet Math. Dokl, vol.269, pp.543-547, 1983.
[13] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Classics in Applied Mathematics, 30, SIAM, Philadelphia, 2000.
[14] H.C. Sun, Y.J. Wang, L.Q. Qi, "Global error bound for the generalized linear complementarity problem over a polyhedral cone," Journal of Optimization Theory & Applications, vol.142, pp.417-429, March, 2009.
[15] H.C. Sun, Y.J. Wang, "Further discussion on the error bound for generalized linear complementarity problem over a polyhedral cone," Journal of Optimization Theory & Applications, vol.159, no.1, pp.93-107, February, 2013.
[16] H.C. Sun, Y.J. Wang, S.J. Li, M. Sun, "A sharper global error bound for the generalized linear complementarity problem over a polyhedral cone under weaker conditions," Journal of Fixed Point Theory & Applications, vol.20, no.2, pp.1-19, April, 2018.
[17] H.C. Sun, Y.J. Wang, S.J. Li, M. Sun, "An improvement on the global error bound estimation for ELCP and its applications," Numerical Functional Analysis and Optimization, <https://doi.org/10.1080/01630563.2021.1919897>, May, 2021.

Wheelchair Control System based Eye Gaze

Samar Gamal Amer¹, Sanaa A. kamh³
Marwa A. Elshahed⁴
Department of Physics
Faculty of Women
Ain Shames University, Cairo, Egypt

Rabie A. Ramadan²
Computer Engineering Department
Faculty of Engineering
Cairo University
Cairo, Egypt

Abstract—The inability to control the limbs is the main reason that affects the daily activities of the disabled which causes social restrictions and isolation. More studies were performed to help disabilities for easy communication with the outside world and others. Various techniques are designed to help the disabled in carrying out daily activities easily. Among these technologies is the Smart Wheelchair. This research aims to develop a smart eye-controlled wheelchair whose movement depends on eye movement tracking. The proposed Wheelchair is simple in design and easy to use with low cost compared with previous Wheelchairs. The eye movement was detected through a camera fixed on the chair. The user's gaze direction is obtained from the captured image after some processing and analysis. The order is sent to the Arduino Uno board which controls the wheelchair movement. The Wheelchair performance was checked using different volunteers and its accuracy reached 94.4% with a very short response time compared with the other existing chairs.

Keywords—Dilip; numpy; gaze ratio; facial landmarks points; deep learning

I. INTRODUCTION

Wheelchairs have become essential for the elderly and the disabled nowadays. Electric wheelchairs are available on the market and are typically controlled using a joystick [1]. In some cases, many people can be injured unexpectedly during various events, including falls, accidents, violence, or even injuries while playing sports. These events may result in devastating neuromuscular disorders, resulting in significant disabilities as a result of injuries to the spinal cord, which may be caused by a lack of nerve supply, paralyzing every affected part of the human body. As a result, there is an emerging need to assist these people in improving their abilities to carry out their daily routine without the assistance of others. [2]. This was a strong motivation for issuing this research paper, different systems were used to control wheelchairs, such as the voice control system [3], brain control system [4], and eye control system [5]. In the voice control system, the Wheelchair is controlled by the sound so that the user gives voice commands such as forward, right, left, and stop, and the system will recognize the word then send the command to move the Wheelchair accordingly. The Wheelchair controlled by the brain signals depends on capturing the patient's so-called EEG signal to move the chair. The downsides of these systems include a reduced immunity to noise, which might cause the system to become distracted and respond incorrectly.

This research focuses on designing a wheelchair that doesn't depend on any physical movement of the user but tracks the motion of the eye with achieving better accuracy than others also, minimizing the delay response time of the detection. The proposed eye-tracking electric wheelchair differs from others with good validation parameters of the controller which uses simple equations depend on counting the white pixels in both eyes to identify the eye pupil position and direction in the user's face. The main idea of this work depends on capturing the user's image using a webcam. some image processing techniques were performed to determine the location of the face, the eye, and its direction, and then control the Wheelchair via the Arduino board [6]. The face detection technique is a computer technology used to detect the human face from a digital image based on facial landmark points. It is considered as a set of key points on human face images [7]. These points are described in the image by their actual coordinates (x, y). A pictorial image has been used to model the facial landmarks relations, as shown in Fig. 1.

This figure is an example of a facial landmark 68 points model (from 1 to 68). After face detection, the system can detect the eye gaze. Then, the process of detecting the position of the eye sclera is performed by calculating the value of the gaze ratio. The gaze ratio is the ratio of the white pixels between each side in both eyes. This ratio allows determining the area at which the person is looking [8,9,10].

The proposed smart wheelchair contributes to facilitating the lives of the disabled and making their lives easier and reducing their financial burden. Also, improving the accuracy and reducing the time between the user's eye and chair movements than other wheelchairs.

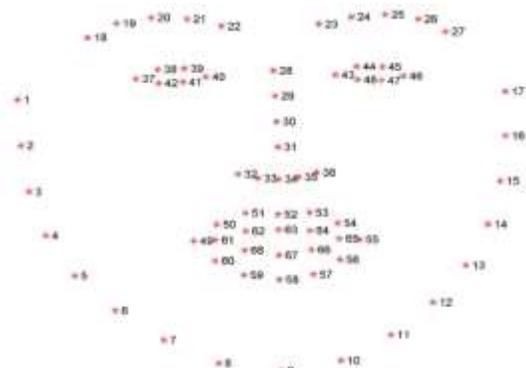


Fig. 1. Facial Landmark Points.

II. RELATED WORK

The technology that allows extensive use of eye-tracking principles includes sectors in the automotive industry, medical science, exhaustion simulation, automobile simulators, cognitive tests, computer vision, behavior recognition, etc. Over a while, the importance of eye recognition and monitoring in industrial applications grew. This importance of eye-tracking applications leads to more effective and durable designs required in many modern appliances. An extensive review of the literature relating to the eye-tracking system has been carried out in healthcare applications.

The authors in [11] uses machine learning to extract the iris by segmentation algorithm. All experiments were implemented using the MATLAB software, the algorithm accuracy reached to 86% which is not satisfactory, also the authors mentioned that the speed of detection increased but did not mention its value.

A wheelchair is designed in [12], which is based on eye movement. This chair consists of three main parts: imaging processing module, wheelchair-controlled module, SMS manager module, and appliance-controlled module. The imaging processing module consists of a camera installed with glasses. The captured image is transferred to the Raspberry Pi microcontroller will be processed using OpenCV to derive the direction of the two-dimensional eyeball. The eyeball movement will be transmitted wirelessly to a unit controlled by a wheelchair to control movement. But the Calibration time is not calculated. Their proposed chair is relatively high cost and lack of components. A microcontroller called Raspberry Pi were used which may cost up to 70\$, But our proposed system used the Arduino Uno, which doesn't exceed 10 \$ and it reaches the same result.

In [13], the researcher designed a wheelchair that moves by eye-tracking, first the camera capture the user image and detect the eye position by edge detection technique (Laplacian), Then detect the pupil direction by segmentation algorithm (Haar cascade algorithm) then, controlling the wheelchair DC motor with the help of a PID controller. The accuracy of detection was 90%. But Calibration time is not calculated, also, the Laplacian technique is unacceptably sensitive to noise. Haar cascade is an algorithm that depends on edge or line detection so it's suitable for face and eye detection but not suitable to detect eye movement.

The main idea in [14] is that the Wheelchair is controlled by the eyeball movement with the aid of a webcam, which further undergoes multiple image processing techniques. Continuous image is captured to detect the eye pupil's location; then, the Haar cascade algorithm is applied with the wheelchair image processing techniques. The DC motor is mounted on the wheels for quick wheelchair travel. The ultrasonic sensor is mounted on the Wheelchair to locate any obstacles in the direction of its movements and prevents the movement of the Wheelchair as per sensor order. The chair accuracy was 91% which needs improvement. Also, Haar cascade is multiple weak classifiers as mentioned before in this application and the wheelchair move in 4 direction, doesn't move in 360° as our proposed wheelchair move in 360° in the direction of the eyes.

In [15], a wheelchair-based eye-tracking system is presented, which is based on record a video to user's face using an infrared camera, the video is then fed into the proposed algorithm that of six stages: (1) eye area extraction, (2) iris boundary detection, (3) keyframe detection, (4) pupil localization, (5) deviation estimation, and (6) evaluation of strabismus. A database was created that included cover test data from both strabismus and normal subjects. The results of the experiments show that our proposed method can accurately evaluate strabismus deviation. The accuracy was over 91%, in the horizontal direction, and it was over 86% in the vertical direction. The only drawback is that measuring accuracy can improve.

The authors in [16] propose a system that enables a person to have a control wheelchair across the eye gaze. The system consists of a wheelchair, eye-tracking glasses, and a depth camera, Ambient Space Engineering, Portable Computer Sit Flexible Stand for Maximum comfort, and a safety off switch to turn off the system when needed. The author uses the CNN algorithm to capture the eye and calculate its direction. The accuracy of their study reached 92% and the calibration time reached 36 s. CNN algorithm is not preferred in face and eye detections as it is complex. Also, from the results that the calibration time is very long 36 s.

The proposed system is based on the deep Learning base method which has better accuracy in all face parts detection and an easy method to extract eyes Using the dlib and his Face Landmark points.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed system is simple and has low-cost components, as shown in Fig. 2. It is based on taking the user's face image. Getting the eye's location and then determining the pupil's direction which has a different value. The obtained values are transmitted to the Arduino board connected to a wheelchair to control its moving directions. The designed Wheelchair consists of an integrated circuit called L293d and a motor of the type of DC that works from 3 to 6 V.

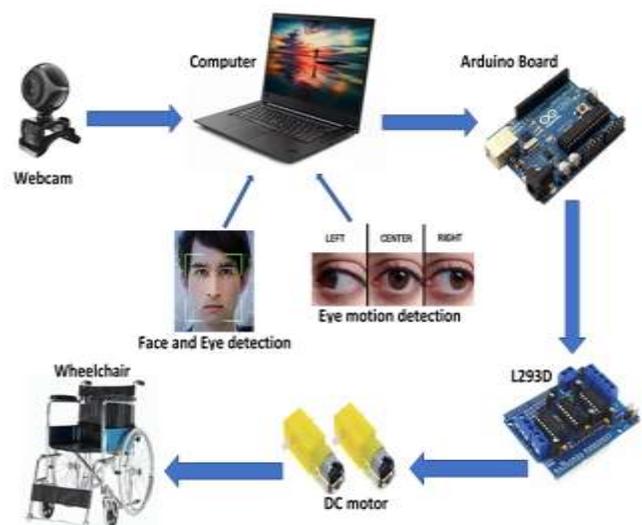


Fig. 2. Proposed System Architecture.

Face and eye detection are determined; then calculate the gaze ratio, which determines the eye direction is (left or right or center). The Arduino Uno is used with an L293D integrated circuit, which is a typical Motor driver or Motor Driver IC. L293D is a 16-pin IC that can control two DC motors simultaneously in any direction and then control in Wheelchair. 5 adult volunteers aged 15 to 50 years old are examined in this study.

This study presents a Technique depend on an algorithm is called "One Millisecond Face Alignment with an Ensemble of Regression Trees" developed by two Swedish Computer Vision researchers Kazemi and Sullivan in 2014[17,18], This detector is built in the dlib library and it detects facial landmarks very quickly and accurately and this algorithm is based on dividing the face into 68 points, Beside the pre-trained facial landmark detector inside the dlib library is used to estimate the location of 68 (x, y)-coordinates that map to facial structures on the face, This occurs by calling a frontal face detector from dlib library. This is a pre-trained detector based on Histogram of Oriented Gradients (HOG) features and a Linear SVM object detector [19]. And the dataset on which the dlib facial landmark was trained in the shape predictor 68 face landmarks" dataset. The facial landmarks with dlib technique don't need any preprocessing or filters but need a high-efficiency camera with strong lenses and high focus, The more efficient the camera, the higher accuracy of eye-tracking.

IV. METHOD

This work used a computational efficient technique to precisely predict the position of face landmarks. A cascade of regressors is used in our proposed strategy.

A. The Cascade of Regressors

To begin we introduce some notation. Let $x_i \in \mathbb{R}^2$ be the x, y-coordinates of the i^{th} facial landmark in an image I. Then the vector $S = (x_1^T, x_2^T, \dots, x_p^T)^T \in \mathbb{R}^{2p}$ denotes the coordinates of all the p facial landmarks in I. In this paper, the vector S is referred to as the shape. And $S^{\wedge}(t)$ is used to denote our current estimate of S. Each regressor, $r_t(\cdot, \cdot)$, in the cascade predicts an update vector from the image and $S^{\wedge}(t)$ that is added to the current shape estimate $S^{\wedge}(t)$ to improve the estimate:

$$S^{\wedge}(t+1) = S^{\wedge}(t) + r_t(I, S^{\wedge}(t)) \quad (1)$$

The regressor r_t makes its predictions based on features, such as pixel intensity values, computed from I and indexed relative to the current shape estimate $S^{\wedge}(t)$ [16]. We will explain how this indexing is performed in detail.

1) *Learning each regressor in the cascade:* Assume we have training data $(I_1, S_1), \dots, (I_n, S_n)$. To learn the first regression function r_0 in the cascade we create from our training data triplets of a face image, an initial shape estimate is obtained by,

$$(I_{\pi_i}, s_i^{\wedge 0}, \Delta s_i^{\wedge 0}) \text{ where } \pi_i \in \{1, \dots, n\} \quad (2)$$

$$s_i^{\wedge 0} \in \{S_1, \dots, S_n\} \setminus S_{\pi_i} \quad (3)$$

$$\text{And } \Delta s_i^{\wedge 0} = S_{\pi_i} - s_i^{\wedge 0} \quad (4)$$

For $i = 1, \dots, N$, $N = nR$ where R is the number of initializations used per image I, And by algorithm 1. The next regressor r_i in the cascade by setting (with $t = 0$) is obtained by:-

$$s_i^{\wedge(t+1)} = s_i^{\wedge(t)} + r_i(I_{\pi_i}, s_i^{\wedge(t)}) \quad (5)$$

$$\Delta s_i^{\wedge(t+1)} = S_{\pi_i} - s_i^{\wedge(t+1)} \quad (6)$$

This process is iterated until a cascade of T regressors r_0, r_1, \dots, r_{T-1} are learned which when combined give a sufficient level of accuracy.

2) Learning r_i in the cascade:

Have training data $\{(I_{\pi_i}, s_i^{\wedge(t)}, \Delta s_i^{\wedge(t)})\}_{i=1, N}$ and the learning rate

(shrinkage factor) $0 < v < 1$

Initialize

$$f_0(I, S^{\wedge(t)}) = \arg \min \sum_{i=1}^N \|\Delta s_i^{\wedge(t)} - \gamma\|^2$$

$\gamma \in \mathbb{R}^{2p}$

2. for $k = 1, \dots, K$:

(a) Set for $i = 1, \dots, N$

$$r_{ik} = \Delta s_i^{\wedge(t)} - f_{k-1}(I_{\pi_i}, s_i^{\wedge(t)})$$

(b) Fit a regression tree to the targets r_{ik} giving a weak regression function $g_k(I, S^{\wedge(t)})$.

(c) Update

$$f_k(I, S^{\wedge(t)}) = f_{k-1}(I, S^{\wedge(t)}) + v g_k(I, S^{\wedge(t)})$$

3. Output $r_i(I, S^{\wedge(t)}) = f_k(I, S^{\wedge(t)})$

B. Tree based Regressor

Now we'll go through the most important implementation details for training each regression tree.

Shape invariant split tests

We make a decision based on thresholding the difference between the intensities of two pixels at each split node in the regression tree. When defined in the coordinate system of the mean shape, the pixels used in the test are at positions u and v. Let k_u be the index of the facial landmark in the mean shape that is closest to u and define its offset from u as:

$$\delta x_u = u - x^{-k_u} \quad (7)$$

Then for a shape, S_i defined in image I_i , the mean shape image is given by

$$u^- = x_{i, k_u} + \frac{1}{s_i} R_i^t \delta x_u \quad (8)$$

Where s_i and R_i are the scale and rotation matrix of the similarity transform which transforms S_i to S^- , the mean shape. The scale and rotation are found to minimize.

$$\sum_{j=1}^p \|x_j' - (s_i R_i x_{i,j} + t_i)\|^2 \quad (9)$$

The sum of squares between the mean shape's facial landmark points, x_j^- 's, and those of the warped shape. V' is similarly defined. Each split is formally a decision involving three parameters $\theta = (\tau, u, v)$ and is applied to each training and test example as

$$h(I_{\pi i}, s_i^{\hat{t}}, \theta) = \begin{cases} 1 & I_{\pi i}(u') - I_{\pi i}(v') > \tau \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Where u' and v' are defined using the scale and rotation matrix which best warp $s_i^{\hat{t}}$ to S^- according to equation (7).

1) Choosing the node splits

The training of the regression tree

To train the regression tree we randomly generate a set of candidate splits, that is θ 's, at each node. We then greedily choose the θ^* , from these candidates, which minimizes the sum of square error.

This corresponds to minimizing Q , which is the set of indices of the training examples at a node.

$$E(Q, \theta) = \sum_{s \in \{l, r\}} \sum_{i \in Q_{\theta, s}} \|r_i - \mu_{\theta, s}\|^2 \quad (11)$$

where $Q_{\theta, l}$ is the indices of the examples that are sent to the left node due to the decision induced by θ , r_i is the vector of all the residuals in the gradient boosting algorithm, computed for an image I and

$$\mu_{\theta, s} = \frac{1}{\|Q_{\theta, s}\|} \sum_{i \in Q_{\theta, s}} r_i \text{ for } s \in \{l, r\} \quad (12)$$

The optimal split can be found very efficiently because if one rearranges equation (10) and omits the factors not dependent on θ then one can see that

$$\arg \min_{\theta} E(Q, \theta) = \arg \max_{\theta} \sum_{s \in \{l, r\}} \|Q_{\theta, s}\| \mu_{\theta, s}^T \mu_{\theta, s} \quad (13)$$

Here we only need to compute $\mu_{\theta, l}$ when evaluating different θ 's, as $\mu_{\theta, r}$ can be calculated from the average of the targets at the parent node μ and $\mu_{\theta, l}$ as follows:

$$\mu_{\theta, r} = \frac{\|Q\|\mu - \|Q_{\theta, l}\|\mu_{\theta, l}}{Q_{\theta, r}} \quad (14)$$

2) *Feature selection*: The thresholding of the difference in intensity values at a pair of pixels is used to make the decision at each node. Unfortunately, the drawback of using pixel differences is the number of potential splits (feature) candidates is quadratic in the number of pixels in the mean image. This makes it difficult to find good θ 's without searching over a very large number of them. However, this limiting factor can be eased, to some extent, by taking the structure of image data into account by the equation:

$$P(u, v) = \alpha e^{-\lambda \|u - v\|} \quad (15)$$

Over the distance between the pixels used in a split to encourage the selection of closer pixel pairs. We found using this simple prior reduces the prediction error on several face datasets.

V. EXPERIMENT

A. The Proposed System Flowchart

The proposed system is built based on Python programming language in Pycharm cross-platform. Building the system depends on OpenCV and Dlib libraries, it has been used dlib get frontal face detector in a frame or image. Dlib predictor is a tool that takes in an image containing some objects and outputs a set of point locations that define the pose of the objects. Here the shape predictor 68 face landmarks. The data model is used to create the predictor object.

The implementation of the system can be summarized in the following four points as in Fig. 3:

- The face and the eyes detection
- Calculating the gaze ratio to detect the eye direction
- The wheelchair design.
- The Connection between OpenCV and Arduino.

In this study, we can configure facial landmark points according to our specifications (detect the face and the eyes) as in Fig. 1. For instance, face landmarks that are enough can be from point 1 to point 27. Besides, left eye landmarks can be defined based on points 37-42, and then the right eye landmarks could be defined based on points 43-48.

B. Calculating the Gaze Ratio to Detect the Eye Direction

It is considered that the eye is divided into two parts white part (sclera) and the color part (pupil). The idea is to find out which of the two parts is more sclera visible as the sclera expresses the direction the eye is looking, as shown in Fig. 4.

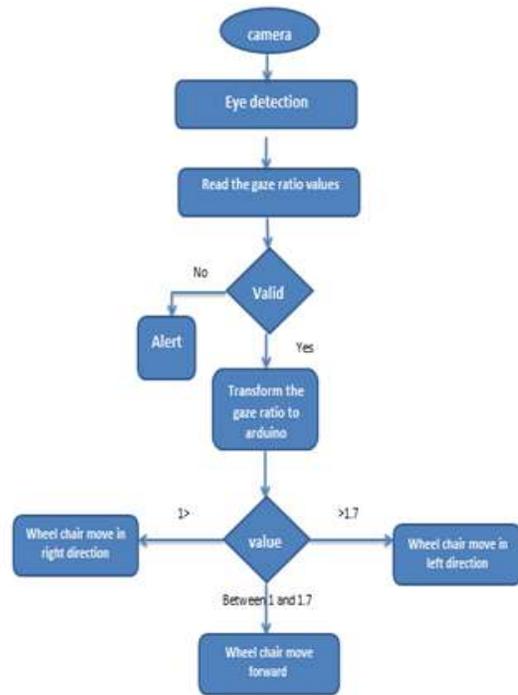


Fig. 3. The Proposed System Flowchart.

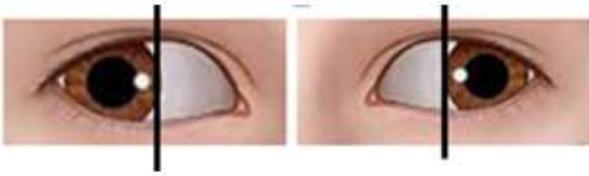


Fig. 4. Facial Detecting Eye Directions.

To detect the location of the sclera, convert the image into grayscale and calculate the number of white pixels on the left side and right side to both eyes. Then, calculate the gaze ratio to the right and left eye, which is the ratio between the left and white pixels and the right side of white pixels to each eye. Then, calculate the final gaze ratio to both eyes, which is the average value of the right and left gaze ratio. The values of the final gaze ratio could be computed using the following equation.

$$Gaze\ Ratio = \begin{cases} < 1, \text{Sclera in right direction} \\ > 1 \text{ and } < 1.7, \text{sclera in center} \\ > 1.7, \text{sclera in left direction} \end{cases} \quad (16)$$

These values are observed for all volunteers.

C. The Wheelchair Design

The smart Wheelchair consists of three main components, as shown in Fig. 5.

Arduino Uno board.

L293D integrated circuit.

Two DC motors.

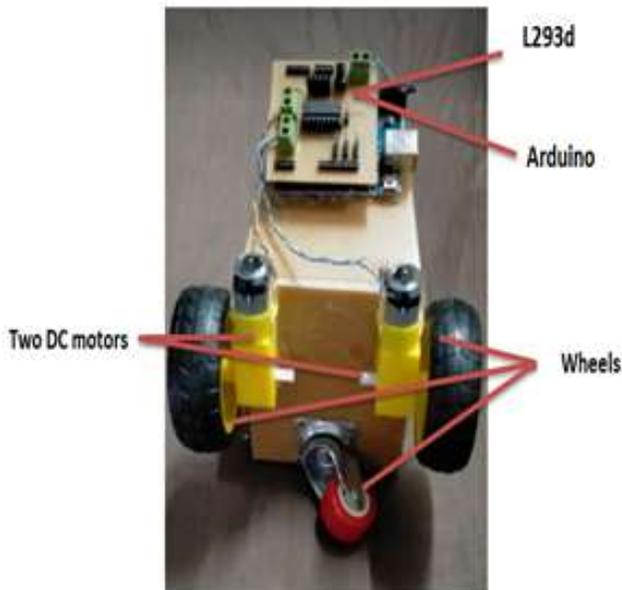


Fig. 5. Wheelchair Back Architecture.

1) *Arduino Uno board*: -is used to revise the eye direction from OpenCV and it is connected to the L293D integrated circuit.

2) *L293D integrated circuit*: -Using this L293D motor driver IC is very simple. The IC operates on the Half H-Bridge principle, a setup that is used to operate motors in both clockwise and anti-clockwise directions is the 1, H bridge. As stated earlier, this IC can run the two motors in any direction at the same time. In this experiment, L293D revise the control order (eye direction) from Arduino and control DC motors' control command [20].

3) *Two DC motors*: - All types of DC motors have some function, either electromechanical or electronic, to change the motor periodically's current path. In the experiment, each wheel in the chair is connected to the dc motor, and each motor receives the control order from the L293D integrated circuit to move the wheels in all directions. As shown in Fig. 6, Wheelchair forward architecture consists of a place to sit and a battery.



Fig. 6. Wheelchair Forward Architecture.

VI. RESULTS AND ANALYSIS

The performance measure criteria in this study are: -

- 1) Calibration time.
- 2) Accuracy.

According to calibration time: During the testing time, it is very fast in the order of a millisecond. According to accuracy: Accuracy was calculated by counting the gaze detection for 3 persons (P1, P2, and P3). For each person, 30 gazes are tested, 10 for the right eye, 10 for the left eye, and 10 for the blinking eye, as shown in Table I.

Where:

$$Accuracy = \frac{\text{no of gaze detected}}{\text{total no od detection}} \quad (17)$$

TABLE I. THE CALCULATION OF THE SYSTEM PERFORMANCE

S.No	P1	P2	P3
Total Detections No	30	30	30
Total no. of max left Detection	8/10	8/10	10/10
Total no. of max Right Detection	10/10	10/10	10/10
Total no. of max blinking Detection	10/10	10/10	9/10
Total no. of max Detection	28/30	28/30	29/30
Accuracy	93.3%	93.3%	96.6%
Error	6.5%	6.5%	3.4%

The obtained average of accuracy is 94.4%. It is evident from the results that this study provided a noticeable improvement in eye tracking compared to previous studies. First, a Facial landmark with the Dlib technique one of the latest and more accurate techniques that are used in determining the face and eye was issued in 2014, besides it is clear from the results that the calibration time doesn't exceed a millisecond and accuracy is 94.4%, While the best results were previously monitored, the calibration time was 33.82 s and the accuracy was 90%. Table II presents all the proposed system components' costs.

TABLE II. THE COST OF THE MOST IMPORTANT TOOLS USED

Tool	Price
Arduino Board	10 \$
LD293D	2 \$
2 Dc Motor	1 \$
Wheelchair Model Design	5 \$
Total	18\$

VII. CONCLUSION

Eye-tracking is an interesting field of research, especially to help the completely disabled to communicate better with the community and practice all life activities without anyone's help. For the proposed eye-tracking system, a successful circuit was built between a wheelchair and a webcam which was used to capture the user's face. A face detection algorithm was used using the landmarks of the face and the eyes based on Facial Landmarks points which divide the face into 68 points. The gaze ratio was calculated by tracking the pupil area. We have found that when this ratio is less than 1, the pupil of the eye is in the right direction and when it is greater than 1. The pupil is in the left direction, otherwise, it is in the middle, then these values were passed to the Arduino board, then to an L293D circuit, and then to the wheelchair wheels. The experiment was prepared using Python programming language and OpenCV Library. The experiment achieved better results than other wheelchairs, as the accuracy of the gaze detection was about 94.4%, the time was one microsecond, and the cost didn't exceed 15\$.

REFERENCES

- [1] S. Kundu, O. Mazumder, P. K. Lenka and S. Bhaumik, "Hand Gesture Recognition Based Omnidirectional Wheelchair Control Using IMU and EMG Sensors", Journal of Intelligent & Robotic Systems, vol. 91, no. 3, pp. 529-241, 2018.
- [2] Shahin, M. K., Tharwat, A., Gaber, T., Hassanien, "A Wheelchair Control System Using Human-Machine Interaction: Single-Modal and Multimodal Approaches". Journal of Intelligent Systems, April-2019.
- [3] Putri Madona, Husna Khairun Nisa, Yusmar Palapa Wijaya and Amnur Akhyan, "The Design Of Wheelchair Systems With Raspberry Pi 3-Based Joystick Analog And Voice Control", Published under licence by IOP Publishing Ltd, Indonesia, November 2019.
- [4] Mamunur Rashid, Norizam Sulaiman, Mahfuzah Mustafa, Sabira Khatun, Bifta Sama Bari, Md jahid Hasan, Mamunur Rashid, "Recent Trends and Open Challenges in EEG Based Brain-Computer Interface Systems", Malaysia, 24 March 2020.
- [5] Dahmani, M., Chowdhury, M. E. H., Khandakar, A., Rahman, T., Al-Jayyousi, K., Hefny, A., Kiranyaz, S., "An Intelligent and Low-Cost Eye-Tracking System for Motorized Wheelchair Control". Sensors, July 2020.
- [6] <https://opencv.org>.
- [7] Fwa Hua Leong, "Deep learning of facial embeddings and facial landmark points for the detection of academic emotions", International Conference on Information and Education Innovations, July 2020.
- [8] Samar Gamal Amer, Sanaa A. kamh, Marwa A. Elshahed, "WRITING BASED EYE GAZE ESTIMATION SYSTEM", Sci.Int.(Lahore), 2020.
- [9] Anagha Dwajan B., Sowmya M S., Bhavani S., Priya Reddy N, Usha M R, "Eye Gaze Controlled Wheelchair", INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT), May 2020.
- [10] Mahendran Subramanian, Pavel Orlo, Suhung Park, "Gaze-contingent decoding of human navigation intention on an autonomous wheelchair platform", International IEEE/EMBS Conference on Neural Engineering (NER), May 2021.
- [11] Rupanagudi, S. R., Bhat, V. G., Gurikar, S. K., Koundinya, S. P., M. S., S. K., R., S., Satyananda, "A Video Processing Based Eye Gaze Recognition Algorithm for Wheelchair Control". 10th International Conference on Dependable Systems, Services and Technologies. 2019.
- [12] Ragini Singh1, Harsha Rani2, Junaid Hasan Khan3, Kumari Komal, "Eyeball Controlled Wheelchair, Journal of Science and Technology", 2020.
- [13] Hadish Habte Tesfamikael, Adam Fray, Israel Mengsteab, Adonay Semere, "Simulation of Eye Tracking Control based Electric Wheelchair Construction by Image Segmentation Algorithm", Journal of Innovative Image Processing, 2021.
- [14] Nutthanan Wanluk, Aniwat Juhong, Sarinporn Visitsattapongse, C. Pintavirooj, "Smart wheelchair based on eye tracking", The 2020 Biomedical Engineering International Conference, 2020.
- [15] Zheng, Y., Fu, H., Li, R., Lo, W.-L., Chi, Z., Feng, D., Wen, D. "Intelligent Evaluation of Strabismus in Videos Based on an Automated Cover Test". Applied Sciences, 2019.
- [16] MOHAMAD A. EID1, NIKOLAS GIAKOU MIDIS, "A Novel Eye-Gaze-Controlled Wheelchair System for Navigating Unknown Environments: Case Study With a Person With ALS", IEEE, 2016.
- [17] V. Kazemi J. Sullivan, "One millisecond face alignment with an ensemble of regression trees", IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [18] X. Cao, Y. Wei, F. Wen, J. Sun, "Face alignment by explicit shape regression", 2013.
- [19] Zhou, W., Gao, S., Zhang, L., Lou, X., "Histogram of Oriented Gradients Feature Extraction from Raw Bayer Pattern Images", IEEE Transactions on Circuits and Systems, 2020.
- [20] <https://www.instructables.com/Using-Motors-With-L293D-IC/>.

Design and Implementation of a Most Secure Cryptographic Scheme for Lightweight Environment using Elliptic Curve and Trigonohash Technique

BhaskarPrakash Kosta¹, Dr. PasalaSanyasi Naidu²
Computer Science and Engineering Department
GIT, GITAM Deemed to be University
Vishakhapatnam, AP, India

Abstract—The Internet of Things (IoT) is a rising development and is an organization of all gadgets that can be gotten to through the web. As a central advancement of the IoT, wireless sensor networks (WSN) can be used to accumulate the vital environment parameters for express applications. In light of the resource limitation of sensor devices and the open idea of remote channel, security has become an enormous test in WSN. Validation as an essential security service can be used to guarantee the authenticity of data access in WSN. The proposed three factor system using one way hash function is depicted by low computational cost, and limit overhead, while achieving all other benefits. Keys are made from secret key for meeting for improving the security. We differentiated the arrangement's security and execution with some lightweight plans. As shown by the examination, the proposed plan can give more prominent security incorporates low overhead of correspondence. Encryption and unscrambling is done using numerical thoughts and by using the possibility of hash function. Mathematical thoughts are lightweight and update the security up by a staggering degree by diminishing the chances of cryptanalysis. When contrasted with different calculations, the proposed calculation gives better execution results.

Keywords—Internet of Things (IoT); authentication; one way hash function; lightweight environment; secret key

I. INTRODUCTION

As far off correspondence headway has made, brilliant world individuals can work by utilizing impending Internet of Thing (IoT) paradigm[1]. The fundamental contemplated the Internet of things, or IoT, is a strategy of interrelated getting ready contraptions, mechanical and automated machines, things, creatures or individuals that are given great identifiers (UIDs) and the capacity to move information over an association without expecting that human should human or human-to-PC facilitated exertion. Numerous splendid application can be made by using the data delivered by radio-frequency identification(RFID) and remote sensor organization. A comparable idea can be applied in various mechanical field[2] where canny application can be made using IoT devices. For example in healthcare[3] some distinguishing devices or sensors are embedded or attached to customer which assemble some basic information about customer, for instance, ECG, pulse, temperature and blood oxygen. these fundamental information is then shipped off affirmed clinical master with the help of a section encouraged

by the affiliation who runs the center, taking into account which the treatment is overseen without the actual presence of clinical master (This technique can be used to perceive COVID-19 patient among the people who are in separate so the spread of sickness can be avoided).

Using WSN biological data can be accumulated. WSN are regularly involved customers, passages(gateway server) and sensors centers. By and large sensor centers have limited energy, computational force and limit power moreover it has been seen that the energy usage of sensor center depends upon correspondence detachment [4][28], the energy usage for sending or tolerating a message of l -bits over or from a distance (d) are assessed by Eqs. 1 and 2, independently. Here the free space model is used if (d) isn't by and large a cutoff d_0 ; something different, multi-way model is used.

$$E_r(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2 & \text{for } d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4 & \text{for } d \geq d_0 \end{cases}$$

$$E_R(l) = lE_{elec}$$

where E_{elec} is the energy required by the electronic circuit, ϵ_{fs} and ϵ_{mp} are the energy required by the amplifier in free space and multi-path model, respectively. One of the plan for correspondence in WSN is showed up in Fig. 1

In Fig. 1, the customer can start the correspondence by sending login message to the gateway(the customer is currently enrolled with the portal the passage(server) moreover keeps up its correspondence with different sensors), as of now the server examinations the sales of customer and makes correspondence with remarkable sensor to make the correspondence possible among customer and sensor. Using this philosophy the power use of sensor center point can be decreased and the life of WSN can be extended. Considering limited security highlight of remote channel and confined resources of IoT devices, IoT faces veritable security and assurance challenges [5].

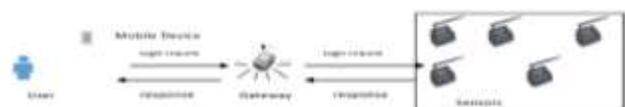


Fig. 1. Architecture for Communication.

Das in 2009 Das arranged a check scheme[6] for WSN, this arrangement was found having issue with inside attack, impersonation attack and server(gateway) center bypassing attack. In 2012 A.K Das [7] arranged powerful customer approval for progressive WSN using light weight calculation, for instance, hash work and symmetric cryptography, in any case it moreover had issues, for instance, inside attack, server primary key revelation attack[8]. In 2013, Xue et al. [9] proposed a lightweight client validation for WSN, and just hash activities are used to ensure the computational effectiveness. They asserted that their plan is secure against identity/secret key speculating assault and taken smart card assault. Be that as it may, their plan was discovered uncertain against stolen smart card assault, insider assault, following assault, and personality speculating assault [10].

As [8] reveals that most symmetric cryptography and hash based protocols are unreliable to user anonymity and smart card security breach assault, thus public key technique can be used like elliptic curve cryptography (ECC) which outfit same security as RSA with more diminutive key size. In 2011, Yeh et al. [11] considered affirmation convention for WSN reliant on ECC. This show had the issue customer namelessness (secrecy) and undetectability. Thereafter, Shi and Gong [12] presented an improved ECC based check convention for WSN. Regardless, these two shows can't give the features of customer namelessness (secrecy) and unrecognizability. In 2016, Jiang et al. [13] proposed an ECC based untraceable affirmation convention for WSN yet had some security defect as pointed by Li et al. [14].

In rest of this paper oversee proposed three-factor customer approval convention using one way hash function and contrasting encryption using one way hash function and disentangling calculation are similarly discussed and security discussion which are showed up in Sections II and III. Section IV oversees comparison of our show with other related show and discussion of security blemishes of Chang and Le's arrangement, Xiong- Li et al. scheme.

II. PROPOSED ANONYMITY AUTHENTICATION PROTOCOL FOR IIOT

In this part, we work with a three-factor validation protocol for modern web of things to accomplish client secrecy and oppose cell phone misfortune assault[15]. Here, we receive the fuzzy extractor [16] to handle biometrics data. More noteworthy degrees of safety are accomplished by biometric secret key, like fingerprints, retina examines. Notwithstanding, two biometric perusing are infrequently indistinguishable, despite the fact that they are probably going to be close. Notation used in full text is shown in Table I. Our plan contains four stages, for example initiation, registration, authentication and key agreement, and password change. We delineate these stages as follows.

TABLE I. NOTATIONS

Notation	Description
U_m, ID_m, PW_m	The Mobile Device(User), and the identification and secret Key
GN, X_{GN}	The gateway Server and its Server secret key
S_s, ID_s	The sensor hub and its identification
xx, XX	Private key and public key Gateway server
bm, Mm	Private key and public key of Mobile device(User)
X_{B-GN}	Shared secret key between Gateway Server and Sensor hub
PP	A point on the elliptic curve
$H1(\cdot), \oplus, \ominus$	The hash, XOR, and XNOR operation
A	An Attacker

A. Initiation Phase

To introduce the framework, Gateway Server picks a single direction hash work $H1(\cdot)$ which it imparts to every one of the clients. Additionally, it chooses an elliptic curve E dependent on a limited field F_p . From that point forward, Gateway server picks a subgroup GG of E with request huge prime n, and the generator is point PP. At that point Gateway worker produces a private key xx and figures the relating public key XX, where $xx \in Z^*n$ and $XX = xxPP$. Finally, GWN stores xx furtively and distributes the boundaries $\{E(F_p), GG, PP, XX\}$.

1) The Gateway Server(GN)

a) Gateway server chooses a secret key (X_{GN}) and one way hash function $H1(\cdot)$ which it shares with all the users through secured channel.

b) GN chooses an elliptic curve E dependent on a limited field F_p . From that point forward, GN picks a subgroup GG of $E(F_p)$ with request huge prime n, and the generator is point PP.

c) GN produces a private key xx and figures the comparing public key XX, where $xx \in Z^*n$ and $XX = xxPP$.

d) GWN stores xx covertly and distributes the boundaries $\{E(F_p), GG, PP, XX\}$.

e) Sensing Device(S_s) Registration Phase.

All the sensing hub(S_s) in IoT are registered offline by the GN as follows.

Step SD1. For each device S_s , the GN chooses a one way hash function $H1(\cdot)$ which it shares with all the sensor nodes in secured manner and a unique identity ID_s and calculates the corresponding shared secret key $X_{ID_s-GN} = (ID_s \oplus X_{GN})$.

Step SD2. The GN pre-loads $\{ ID_s, X_{IDS-GN}, H1(\cdot) \}$ in the memory of S_s

The Sensor Node(Ss)

The Server(GN)

Select IDs for Ss

compute $X_{IDS-GN} = X_{GN} \oplus IDs$

ID_s, X_{IDS-GN}



Store $IDs, X_{IDS-GN}, H1(\cdot)$

B. Registration Phase

To be associated with the framework as a legitimate client, the accompanying advances ought to be performed between client Um and GN, and afterward Um can get to the sensor information progressively by utilizing his/her cell phone.

Stage 1. A character ID_m , a secret word PW_m and an irregular number rm are picked by Um . Then, Um engraves the biometric data Bio_m on a cell phone with fuzzy extractor [16] and gets $Gen(Bio_m) = (\sigma_m, \tau_m)$. At that point Um ascertains the veiled secret key $RPW_m = PW_m \oplus rm$, (here \oplus stands for XOR operation) and submits $\{ID_m, RPW_m, \sigma_m\}$ to GN for enrollment.

Stage 2. While getting the enrollment demand, GN checks if ID_m in the data set. In the event that indeed, Um requested to present another character. Something else, GN figures $A_m = H1(ID_m \oplus X_{GN})$ and $B_m = A_m \oplus (RPW_m \oplus \sigma_m)$, and sends $\{B_m, XX\}$ to Um through a reliable manner.

C. Authentication Phase

When Um wants to access the sensor data of S_s , the following authentication steps should be performed among Um , GN and S_s , and a session key is shared among these three parties at the end of the authentication phase.

Step 1. Um inputs ID_m and PW_m on the mobile device, and imprints the biometric information Bio'_m on the mobile device with fuzzy extractor. Then the mobile device calculates.

$$\sigma_{m'} = Rep(Bio'_m, \tau_m),$$

$$RPW'_m = (PW_m \oplus rm)$$

$$A'_m = B_m \oplus (RPW'_m \oplus \sigma_{m'})$$

checks $A'_m \stackrel{?}{=} A_m$. Imbalance implies in any event one of three variables is invalid, and the login demand is rejected by the cell phone. Or disaster will be imminent, the cell phone plays out the following stage. The cell phone delivers an arbitrary numbers $bm \in Z_n^*$ and calculates.

$$MM1 = bm * PP,$$

$$MM2 = bm * XX$$

$$MM3 = ID_m \oplus MM2$$

$$MM4 = IDs \oplus MM2$$

$$MM5 = H1(A'_m)$$

(1)

Then the mobile device(User) submits the login request $\{MM1, MM3, MM4, MM5\}$ to GN.

Step 2. When receiving the request, GN first retrieves ID_m , IDs and then computes A'_m and then it calculates the hash value of A'_m that is $MM'5$.

$$MM'2 = xx * MM1$$

$$ID'_m = MM3 \oplus MM'2$$

$$ID's = MM4 \oplus MM'2$$

$$A''_m = (ID'_m \oplus X_{GN})$$

$$MM'5 = H1(A''_m)$$

GN server checks $MM'5 = MM5$ If yes then the GN Server get a assurance that the user is genuine then GN server calculates the shared secret key X'_{IDS-GN} between sensor node(Ss) (given by user) and itself then it calculates $MM6, MM7$ for performing authentication.

$$X'_{IDS-GN} = (ID's \oplus X_{GN})$$

$$MM6 = X'_{IDS-GN} \oplus ID'_m$$

$$MM7 = H1(X'_{IDS-GN}) \tag{2}$$

At last, GWN Server sends $\{MM6, MM7\}$ to sensor Node.

Step 3. After receiving the message from GN Server, Sensor node(Ss) first computes.

$$MM'7 = H1(X_{IDS-GN})$$

Sensor node checks $MM'7 = MM7$, if yes then sensor node(Ss) gets a assurance that the request came from correct server, then sensor node(Ss) calculates.

$$ID''_m = MM6 \oplus X_{IDS-GN}$$

$$MM8 = H1(X_{IDS-GN} \oplus ID''_m) \tag{3}$$

The Sensor node also gets the information about the user who wants to communicate, after that it sends responses $\{MM8\}$ to GN Server.

Step 4. After receiving the response from the sensor node(Ss) the GN Server retrieves computes.

$$MM'8 = H1(X'_{IDS-GN} \oplus ID'_m)$$

if $(MM'8 = MM8)$, GN Server gets an assurance that sensor node(Ss) is genuine and generates the session key to be given to both mobile device(user) and sensor node(Ss):

$$SKS = ID'_m \odot IDs \odot X_{GN} \odot A''_m \odot xx \tag{4}$$

$$MM9 = X'_{IDS-GN} \oplus SKS$$

$$MM10 = MM'2 \oplus SKS$$

$$MMM10 = H1(A''_m \oplus ID'_m)$$

Then GN Server sends $MM9$ to sensor node(Ss) and $(MM10, MMM10)$ to mobile device(user). Otherwise, server cuts off the communication.

Step 5. After receiving the message from GN Server, mobile device(User) computes MMM'10.

$$MMM'10 = H1(A'm \oplus IDm)$$

If $MMM'10 = MMM10$ then mobile device(user) gets a assurance that he is communicating with correct GN server then mobile device computes secret key(SKm).

$$SKm = MM10 \oplus MM2$$

$$MM11 = H1(MM10)$$

Then the mobile device(User) submits MM11 to GN Server.

Step 6. After receiving the message from GN Server, Sensor node(Ss) retrieves the secret key.

$$SKs = MM9 \oplus X_{IDs-GN}$$

$$MM12 = H1(MM9)$$

Then the Sensor node submits MM12 to GWN Server.

Step 7. After receiving the message from mobile device(user) MM11 and Sensor node MM12, GN Server, computes the following.

$$MM'11 = H1(MM'2 \oplus SKs)$$

$$MM'12 = H1(X'_{IDs-GN} \oplus SKs)$$

If $M'11 = M11$ and $M'12 = M12$, GN Server gets a assurance that Mobile Device(User) and sensor node(Ss) are is genuine and received the correct secret key and hence key synchronization ends.

If mutual validation and key exchange phase completes successfully, both mobile device(user) and Sensor hub(Ss) will transfer information by using the meeting key which will be computed by the GN server and given to both mobile device(user) and the sensor hub(Ss) in secured way. First source authentication is achieved, for this a one way hash function is used. The message digest is created for date(A_m) which is available at both mobile device(user) and GN server, so this message digest will act as a authenticator. In the same way sensor node(Ss) and GN server mutual authenticate each other by creating a authenticator using hash function (HA-160) by using the common data X_{IDs-GN} . The details of hash function(HA-160) is given below. After all the devices get assurance about other devices the GN server generates a secret key and passes it to both sensor hub(Ss) and mobile device(user) in a secured way.

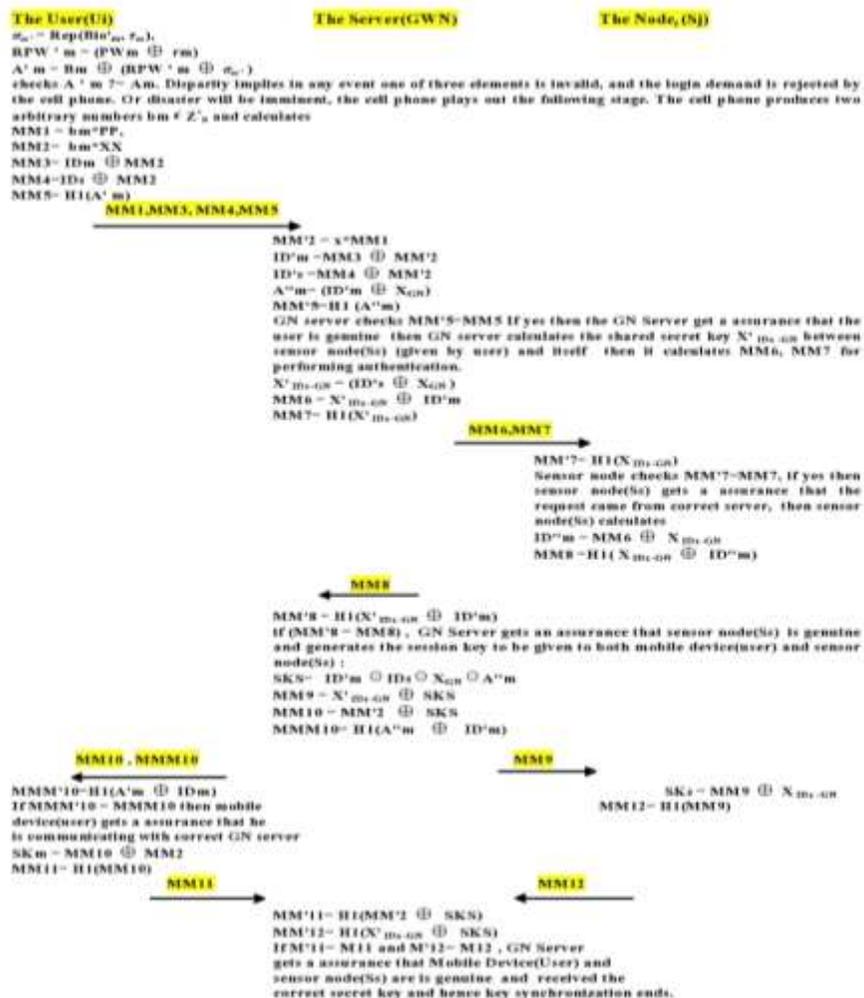


Fig. 2. Illustration of the Mutual Validation Stage and Key Synchronization Stage.

The Secret Key between User(Ui-Mobile) and Server(GN)

The Mobile device(User-Um)

private Key=bm

Public Key = MM1(MM1=bm*PP)

Secret key = bm*XX
= bm*xx*PP

The Server(GN)

private Key=xx

Public Key = XX(XX=xx*PP)

Secret Key =xx*MM1
= xx* bm * PP

The secret key for session synchronization is showed in Fig. 2. Hash algorithm and data encryption phase is described.

III. HASH ALGORITHM(HA-160)

The proposed calculation is named as HA-160[24][25][26]. This assessment HA-160 (Hash Algorithm) recognizes a message as responsibility with a most uncommon length of under 2^{128} pieces and makes a 160-piece message digest as yield. First the data message is isolated into squares of 1024 pieces. This 1024 pieces is taken as information, by then the information is diminished from 1024-digit squares to 512-cycle blocks. The Hash code creation work perceives two wellsprings of data which are 1024 pieces square of the message and the instate MD support (getting variable 160-bits). The cycle includes the going with advances.

A. Attach Padding Pieces

The message is padded so its length is reliable to 896 modulo 1024 (length = 896 mod 1024). Padding is continually done, whether or not the message is of needed length. Subsequently, the amount of padding pieces is in the extent of 1-1024. The padding contains alone 1 followed by the significant number of 0's for example 10... ..0.

B. Attach Length

A square of 128 pieces which contains the length of the message (going before cushioning) is joined to the message. This square is treated as an unsigned 128-digit number (most basic byte first).

C. Initialize MD Pad

A 160-piece cushion is utilized to hold impermanent and eventual outcome of the hash work. The assistance can be tended to as five 32-digit registers (SS0, SS1, SS2, SS3 and SS4). These registers are instated to the going with 32-bit numbers (Hexadecimal qualities):

SS0 = 67 45 23 01

SS1 = efc dab 89

SS2 = 98 ba dc fe

SS3 = 10 32 54 76

SS4 = c3 d2 e1 f 0

These attributes are same as the secret vector evaluations of SHA-1 which are normalized by Federal Information Processing Standards Publications (FIPS PUBS). These qualities are dealt with in big endian plan, which is the standard byte of a word in the low-address byte position.

D. Processes Message in 1024-Bitblocks

The message digest age strategy contains five sub capacity. This part is named (HA-160) in Fig. 1 and its reasoning is showed up in Fig. 2. Fig. 1 depicts the general getting ready of message to make a information survey. The consequence of the underlying two phases (after add padding pieces and append size) yields a information that is a number various of 1024-piece long. The all-inclusive information is addressed as the gathering of 1024-piece blocks $XX_0, XX_1, XX_2, \dots, XX_{L-1}$, so the total size of the all-inclusive information is $L \times 1024$ pieces (L = the amount of 1024 bit hinders), that is the outcome of various of 32-bit blocks. Here K addresses the real length of the message in pieces, "IV" is the fundamental vector which is used to present the five 32-cycle registers (SS0, SS1, SS2, SS3 and SS4). VC_i, VC_{i+1}, VC_q and VC_{L-1} address instate MD(carry vector) which holds widely appealing and inevitable result of the Hash work, separately. Each round takes two data sources one 1024-cycle block (XX_q) of the message and a 160 piece pass on vector (VC_q). Close to the completion of the L^{th} stage produces 160 piece message digest. From the start the given message is disengaged into 1024-digit blocks, and each square is passed to Hash Code making capacity (HA) as a commitment close by the 160-piece vector(as shown in Fig. 3 and Fig. 4).

The Hash work (H_{HA-160}) rationale is:

Remouldingtask(R_m) converts the given 1024-bit block into adjusted 512-bit block.

Progression task(P_m) converts the given 160-piece initial vector into extended 512-bit vector.

32-digit XOR operation(XOR_{op}) performs XOR procedure on each 32-pieces of altered 512-bit block and extended 512-bit vector.

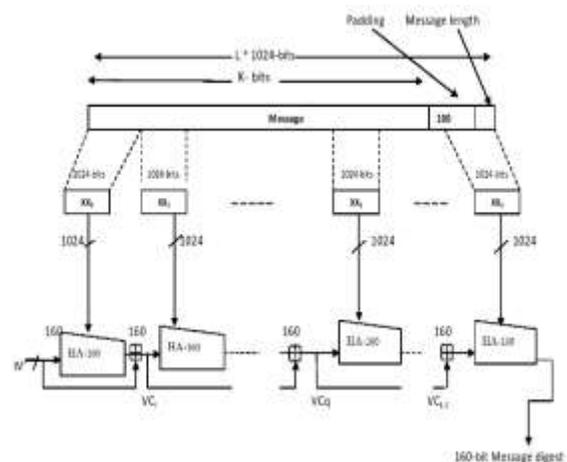


Fig. 3. Message Digest Creation using HA-160.

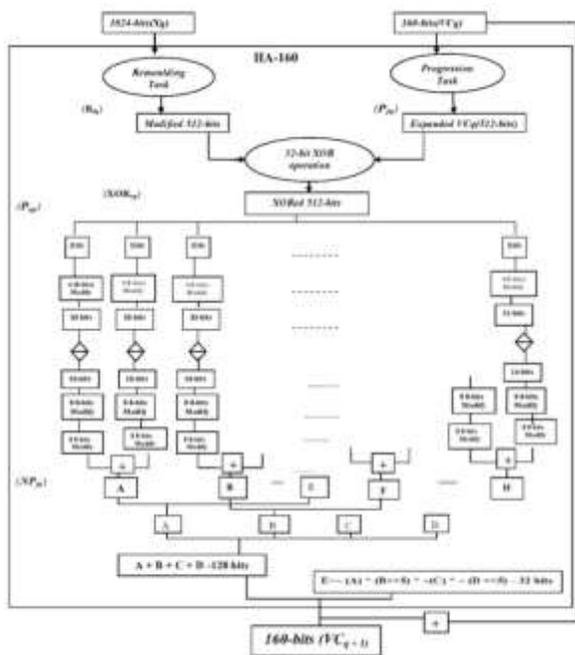


Fig. 4. The Rationale of Hash Work (Compression Work).

Partition and Modification operation (P_{op}) disconnects the 512-cycle block into 16 sub-square of 32-bits each and afterward each 32 digit block is isolated into four 8-bit block, after that it changes every eight bit block. This interaction is rehashed for every one of the 32 bit block.

New Point (on elliptic curve) estimation work (NP_{fn}) calculates new point on elliptic curve using 4 eight-cycle block (from each 32-bit block).

Where,

- XXq = the q th 1024-cycle square of the information.
- VCq = mooring variable arranged with the q th square of the information.
- Rfn = Remoulding task.
- $P=(xxx1,yyy1)$ and $Q=(xxx2,yyy2)$.
- Pfn = Progression task.
- $XORop=XOR(Exclusive-OR)$ development performed on each 32-cycle square of the adjusted 512-digit block (XXq) and relating 32-pieces of the comprehensive 512-bit block (VCq).
- Pop = 512-digit square can be withdrawn into 16 sub squares of 32-bits each and subsequently each 32 cycle block is confined into four 8-bit block, after that it changes each eight bit block.
- $NPfn$ = initial 32-cycle square further parceled into 2 sub parts of 16-piece each.

1stsub part = 16 bit, apportioned into two sub part (8-bits,8-bits) = $(xxx1, yyy1)$.

2nd sub part = 16 bit, divided into two sub part (8-bits, 8-bits) = $(xxx2, yyy2)$.

The characteristics ($xxx1, yyy1, xxx2,$ and $yyy2$ as showed in Fig. 5) are changed over into entire numbers followed by processing another point on elliptic curve. The above interaction is revamped for staying fifteen 32-cycle block and the outcome is 16 sub square every one of 16 pieces. Presently every one of this 8 sub square every one of 16 digit is changed, the initial 8-cycle of first square is XOR with a 8-bit sub square which are each of the zeros, the outcome is put away in initial 8-bit sub square. The subsequent 8-digit sub square of first square is added with initial 8-bit sub square and result is put away in second 8-cycle sub square and the interaction is rehashed for every one of the 8-piece sub square. The outcome is arranged 256 cycle block. Presently neighboring sub square every 16 pieces are added which brings about a sub square of 32 digit and we get eight 32 piece sub square. At that point this eight 32-digit sub square are converted to four 32-bit sub block by added first and fifth, second and sixth, third and seventh and fourth and eighth 32-digit sub square are added which brings about four 32-cycle sub square. By playing out some numerical procedure on over four processed 32 bit sub square and afterward XOR them brings about the fifth 32 cycle sub square. This five 32 piece sub square when concatenated adds up to 160 piece and when summed with introduced MD cradle shapes the hash code. This 160 piece hash code will be the contribution to next 1024 digit of message i.e it will go about as introduce MD cradle for next 1024 cycle of the info message. The last 160 piece code produced from last 1024 cycle of message will be the last message digest which will go about as confirmation code.

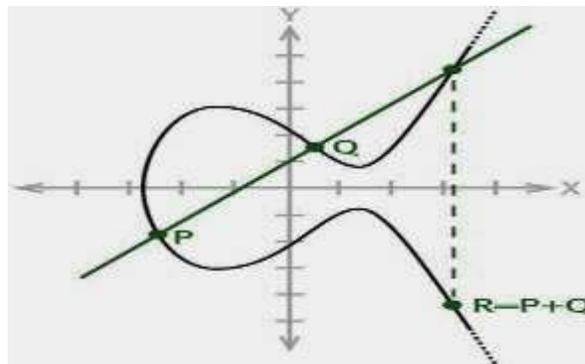


Fig. 5. Elliptic Curve Way to Represent Two Points.

1) *Remouldingtask* (R_{fn}): Each 1024-cycle square of data is isolated into 128 sub-blocks including 8-bits of each sub-block. One brief show of size 8 (Tempx8) is taken and instated with zeroes. The adjustment work includes two sub task as shown as follows.

Sub-work 1. Initially, the first 8-bits of 1024 block is XORed with Tempx8 variable i.e eight bit starting from least significant bit in Tempx8 is XORed with first 8-bits of 1024 block. the result is stored in first 8-bit of 1024 block (Modified Message) and Tempx8 is incremented by one. This extended Tempx8 is XOR with the going with 8-pieces of the message to make the going with 8-pieces of changed message as tended to in Fig. 6.

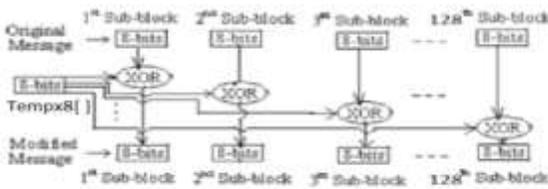


Fig. 6. Function of Sub-Work1.

Sub-work 2. The above result is again detached into sub square of 8 pieces and initial eight-cycle sub square and focus eight-bit sub square after bitwise complimenting(\sim) are XOR and the result is again bitwise complimented(\sim) and set aside in an alternate array(W2). The above cycle is reiterated for second 8-bit sub square and focus notwithstanding one 8-digit sub square and result is taken care of in second piece of disconnected array(W2). This is repeated for all extra 8-digit sub square. The result is 1024 cycle block is diminished to 512 digit block. The changed message as depicted under.

```
intMod_Inp(vector<int> mm, intnn)
{
    staticint W[130] ;//32 bits as a group
    int ii, jj=0,pp,qq,kk1=0,TT=0;
    pp=nn;
    nn*=128;
    for (ii=nn;ii<nn+128; ii++)
    {
        W[ii]=TT ^ mm[ii] ;
        TT++;
    }
    jj=((ii+nn)/2) ;
    qq=jj;
    for (i i=nn;ii<(qq+nn); ii++)
    {
        W2[kk1]=~(~(W[ii]) ^ ~(W[jj]));
        kk1++;
    }
    jj++;
}
return 0;
}
```

2) *ProgressionTask*(P_{fn}):The 160-piece Initial Vector (IV) is one of the commitments to the message digest creation work (HA-160). It will in general be stretched out to 512-bits by interfacing all hidden vector regards in indirect manner, which is called Concatenated Vector (VC). The association cycle is showed in Fig. 7.

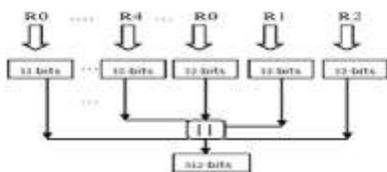


Fig. 7. Formation of 160-Bit Block (IV) to 512-Bit Block (CV).

3) *32-bit XOR operation*(XOR_{op}):XOR movement is done on starting 32-piece sub-block of modified 512-cycle square and expanded 512-piece square. For this initial four 8-digit block are extended and summed with the goal that it becomes 32 piece square. At that point this 32-piece square is XOR with initial 32-bit of extended or instate MD cradle of 512-piece square. This interaction is reshaped for rest of the pieces of adjusted 512-bit and extended 512-digit block. The outcome is 512-cycle block as demonstrated underneath.

```
for (tt1 = 0; tt1 < 16; tt1++)
{
    WW3[tt1] = (WW2[4 * tt1] << 24) + (WW2[4 * tt1 + 1] << 16) + (WW2[4 * tt1 + 2] << 8) + (WW2[4 * tt1 + 3]);
    if(tt1==0)
        WW3[tt1]=((WW3[tt1]) ^ (EE));
    if(tt1==1)
        WW3[tt1]=((WW3[tt1]) ^ (DD));
    if(tt1==2)
        WW3[tt1]=((WW3[tt1]) ^ (CC));
    if(tt1==3)
        WW3[tt1]=((WW3[tt1]) ^ (BB));
    if(tt1==4)
        WW3[tt1]=((WW3[tt1]) ^ (AA));
    if(tt1==5)
        WW3[tt1]=((WW3[tt1]) ^ (EE));
    if(tt1==6)
        WW3[tt1]=((WW3[tt1]) ^ (DD));
    if(tt1==7)
        WW3[tt1]=((WW3[tt1]) ^ (CC));
    if(tt1==8)
        WW3[tt1]=((WW3[tt1]) ^ (BB));
    if(tt1==9)
        WW3[tt1]=((WW3[tt1]) ^ (AA));
    if(tt1==10)
        WW3[tt1]=((WW3[tt1]) ^ (EE));
    if(tt1==11)
        WW3[tt1]=((WW3[tt1]) ^ (DD));
    if(tt1==12)
        WW3[tt1]=((WW3[tt1]) ^ (CC));
    if(tt1==13)
        WW3[tt1]=((WW3[tt1]) ^ (BB));
    if(tt1==14)
        WW3[tt1]=((WW3[tt1]) ^ (AA));
}
```

```
WW3[tt1]=(WW3[tt1]) ^ (AA);  
if(tt1==15)  
WW3[tt1]=(WW3[tt1]) ^ (EE);  
}  
}
```

4) *Partition and Modification operation*(P_{op}): Each 32-digit square of the above outcome is additionally allocated 2 sub-squares of 16-bits each. By then every 16-bit block is segregated into two 8-cycle sub square, all of these eight pieces are changed that is First the four pieces are taken from the 8-digit regard starting from least basic position then it is XOR with a variable1(mm) whose value is zero, the result is kept in a variable2(mm1) also the assessment of variable1(mm) is invigorated for instance it is given the assessment of variable2(mm1). Ensuing to completing the above advance again from 8-digit value(input), four pieces starting from most critical position is taken care of in variable3(mm2) and variable3 is changed by playing out a XOR with variable1(mm) . The data is reproduced for example 8-cycle regard is recomputed by taking care of the assessment of variable3(mm2)(four bits starting from least basic situation) in four pieces starting from most huge position and next four pieces of early on eight bit(input) communicating from least critical position gets the pieces from variable2(km1) (four pieces starting from least huge position) so this way the 8-digit of data is changed . A comparative procedure is embraced for remaining three 8-bit regards made from 32 cycle that is

$BB1 = (WW3[t] \gg 8) \& 0xff; \rightarrow 2^{nd}$ 8- piece number made from 32 bit number

$AA2 = (WW3[t] \gg 16) \& 0xff \rightarrow 3^{rd}$ 8 cycle number made from 32 bit number

$BB2 = (WW3[t] \gg 24) \& 0xff \rightarrow 4^{th}$ 8 digit number made from 32 cycle number

$AA1 \rightarrow yy1$

int Mod_Int2(int yy1)

```
{  
inti,j,mm=0,mm1=0,mm2=0;  
mm1=yy1 & 0xf;  
mm2=(yy1 >> 4) & 0xf;  
mm1= mm1 ^ mm;  
mm= mm1;  
mm2= mm2 ^ mm;  
y1 = (mm2<< 4) + (mm1);  
return yy1;  
}
```

A same strategy is adopted for remaining seven 32-digit blocks.

5) *New Point(on elliptic curve) estimation task* (NP_m): Each 32-digit block is separated into two sub squares of same length. These sub squares go probably as two motivations behind an elliptic curve which are used in new point appraisal in elliptic curve as exhibited in Fig. 5. First by using two point slope(λ) of line is resolved the this slant value(M) is used in the calculation of new X and Y center point on elliptic curve. The resultant new point is 16 cycle regard i.e X center is 8 pieces and Y center point is 8 pieces so both taken together is 16 digit sub square. An amount of sixteen 16-cycle sub square is created after new point computation. As of now this new point are changed by performing XOR on new point. Each 8-digit sub square is XOR with its adjoining 8-cycle sub square beside beginning 8-bit sub square which is XOR with a 8 bit sub square, everything being zero. After above action connecting sub squares are incorporated an especially that it becomes 32 cycle sub square. Repeating this, achieves eight 32-digit sub square.

Incline of line in elliptic curve (for real) is determined as:

$$\text{slope}(\lambda) = \frac{(yyy2 - yyy1)}{(xxx2 - xxx1)}$$

where (xxx1,yyy1) and (xxx2,yyy2) are points on elliptic curve. New point on elliptic curve for reals are calculated using the following formula

$$xxx_3(\text{new point}) = (\text{slope}(\lambda))^2 - xxx_1 - xxx_2$$

and

$$yy_3(\text{new point}) = \text{slope}(\lambda) * (xxx_1 - xxx_2) - yyy_1$$

the above calculation is for the case when $xxx_1 \neq$ (not equal to) xxx_2 (this is assumed that xxx_1 is not equal to xxx_2). The logic of new point estimation is demonstrated below.

```
for ( tt = 0; tt < 16; tt++)  
{  
AA1 = WW3[tt] & 0xff;  
AA1=(Mod_Int2(AA1));  
BB1 = (WW3[tt] >> 8) & 0xff;  
BB1=(Mod_Int2(BB1));  
AA2 = (WW3[tt] >> 16) & 0xff;  
AA2=(Mod_Int2(AA2));  
BB2 = (WW3[tt] >> 24) & 0xff;  
BB2=(Mod_Int2(BB2));  
MM=(BB2-BB1)/(AA2-AA1);  
AA3[ii]=(MM*MM)-AA1-AA2;  
AA3[ii]=AA3[ii]^TT1 ;  
BB3[ii]=MM*(AA1-AA2)-BB1;  
BB3[ii]=BB3[ii] ^ AA3[ii];  
TT1=BB3[ii];  
}
```

```
ii++;  
}
```

The above process results in A3 and B3 array . Now each of this array element is altered using the strategy

```
for(ii=0;ii<16;ii=ii+2)  
{  
if(ii<2)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=2 && ii<4)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=4 && ii<6)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=6 && ii<8)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=8 && ii<10)  
{
```

```
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=10 && ii<12)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=12 && ii<14)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
if(ii>=14 && ii<16)  
{  
AA3[ii] = (AA3[ii] << 24) + (BB3[ii] << 16) + (AA3[ii+1] << 8) + (BB3[ii+1]);  
BB3[ii] = (BB3[ii] << 24) + (AA3[ii] << 16) + (BB3[ii+1] << 8) + (AA3[ii+1]);  
AA3[ii+1] = (AA3[ii+1] << 24) + (BB3[ii+1] << 16) + (AA3[ii] << 8) + (BB3[ii]);  
BB3[ii+1] = (BB3[ii+1] << 24) + (AA3[ii+1] << 16) + (BB3[ii] << 8) + (AA3[ii]);  
}  
}
```

that is AA3[i] is recreated from AA3[i], BB3[i], AA3[i+1] and BB3[i+1] i.e. here AA3[i] is left shifted by 24 bits, BB3[i] is left shifted by 16 bits, AA3[i+1] is left shifted by 8 bits and BB3[i+1] is left shifted by 0(zero) bits and then AA3[i], BB3[i], AA3[i+1] and BB3[i+1] are added to generate AA3[i]. Similarly BB3[i] is recreated from AA3[i], BB3[i], AA3[i+1] and BB3[i+1] i.e. here BB3[i] is left shifted by 24 bits, AA3[i] is left shifted by 16 bits, BB3[i+1] is left shifted by 8 bits and AA3[i+1] is left shifted by 0(zero) bits and then AA3[i], BB3[i], AA3[i+1] and BB3[i+1] are added to generate BB3[i]. For recalculating AA3[i+1] again AA3[i],

BB3[i], AA3[i+1] and BB3[i+1] are used i.e. here AA3[i+1] is left shifted by 24 bits, BB3[i+1] is left shifted by 16 bits, AA3[i] is left shifted by 8 bits and BB3[i] is left shifted by 0(zero) bits and then AA3[i], BB3[i], AA3[i+1] and BB3[i+1] are added to generate AA3[i+1]. Finally for recalculating BB3[i+1] again AA3[i], BB3[i], AA3[i+1] and BB3[i+1] are used i.e. here BB3[i+1] is left shifted by 24 bits, AA3[i+1] is left shifted by 16 bits, BB3[i] is left shifted by 8 bits and AA3[i] is left shifted by 0(zero) bits and then AA3[i], BB3[i], AA3[i+1] and BB3[i+1] are added to generate BB3[i+1]. The same process is repeated for all the value of 'ii'. Here 'ii' is incremented by 2 when the statements inside looping structure are rotated using ii

6) *Result:*The 160-piece hash code is acquired by adding four 32 bit sub square i.e. first and fifth, second and sixth, third and seventh and fourth and eight 32-bit sub square are added which brings about four 32-digit sub square. Presently this four 32-cycle block got in above interaction are numerically controlled which brings about last 32-digit sub square as demonstrated beneath. The message digest (160 pieces) is gotten by connecting the five 32-cycle block got in above interaction.

$EE = \sim(AA) \wedge (BB \gg 5) \wedge \sim(CC) \wedge \sim(DD \ll 5);$

160-bit Hash Code = AA || BB || CC || DD || EE

All hexadecimal numbers are copied into pass on vector (or instate MD cushion), which is the message blueprint of a given message(if the message size is 1024-piece square) in any case the above cycle is emphasized until the last 1024-digit square of the message.

IV. ENCRYPTION ALGORITHM

Step 1: In trigonometry operation, for encryption we use a trigonometric formula made up of Cos function i.e. $Cos(x * \pi / 180.0) = y$ then Sin function i.e. $Sin(y * \pi / 180.0) = z$, here we require the value of π for Cos and Sin function. At the time of encryption and decryption this value of pi will be needed. The actual value of $\pi = 3.141592$, but in this case the Sensor Node and mobile device (user) will take the value of π be current key generated from the secret key for session SK_m or SK_s . For example, let the value of key be 77.

Here is the method how the key for encryption or decryption is generated from secret key for session (SKi).

1) First the secret key is stored in a variable and then this variable is rotated ten times, for each rotation it checks whether the rotation is even or odd. If even, the variable is right shifted by one and the shifted value is stored in a new array (xx), also the variable gets the current value of array x and the array is incremented. If odd, the variable is left shifted by one and the shifted value is stored in a new array (xx), also the variable gets the current value of array x and array is incremented. This process is repeated for a array of size ten, so at the end of first step we get ten different values generated from secret key. The process is shown below:

```
y=sks
jj=1;
for(ii=0;ii<10;ii++)
{
if(ii%2==0)
{
xx[ii]=y<<jj;
y=xx[ii];
}
else
{
xx[ii]=y>>jj;
y=xx[ii];
}
}
```

2) Now the new array is analyzed that is from secret key ten different keys are generated in stage1, then this ten keys are XOR and the result of XOR operation is kept in a variable (t). After computing the value of t, the new array is modified that is array xx first position will now hold the value of array xx second position and array xx second position will now hold the value of array xx third position and the process continues till ninth position tenth array xx position will hold the value of variable t. After completion of above step one position of array xx (tenth position) is altered, the same process is repeated for all the position (ninth to one position) of array xx. So after stage2 the ten different key generated in step one will be holding all new values

```
jj=1.
for(jj=0;jj<10;jj++)
{
t=0;
for(ii=1;ii<10;ii++)
t=(t + xx[ii]) ;
for(ii=0;ii<10;ii++)
{
if(ii==9)
tt2[ii]=t;
else
tt2[i]=xx[ii+1];
}
jj=9;
for(ii=0;ii<10;ii++)
{
xx[ii]=tt2[jj];
jj--;
}
}
```

3) The ten key generated in stage2 are reduced to five that is first key is XOR with tenth key and the result is stored in a separate array(kk1) in first position after a left shift by zero, then second key XOR with ninth key and result is stored in array(kk1) second position after a left shift by one and the process continues till stage2 ten keys are reduced to five. The process is shown below:

```

    jj=10;
    for(ii=0;ii>jj;ii++)
    {
        kk1[ii]=(kk1[ii] + kk1[jj])<<ii;
        jj=jj-1;
    }

```

4) The ten keys of stage2 are reduced to five in stage3 then in stage4 this five keys are XOR and the result is stored in a variable (kkmm). the output of stage4 will be the input to stage5.

```

    kkmm=0
    for(ii=0;ii<5;ii++)
    kkmm=(kkmm^kk1[ii]) ;

```

5) Finally the result of stage4 is reduced to a number less than 10 by applying modular operation. This result as shown in Table II will be acting a key.

Shared secret key for session = 77

TABLE II. GENERATED KEY

KKi	Value
KK1	1155
KK2	2156
KK3	4235
KK4	8316
KK5	16555
KK6	32956
KK7	65835
KK8	131516
KK9	262955
KK10	525756

Final value of Key =3.000000

Step 2: Take the input message, for example “24, 32, 56, 65, 73”. Here every element of input message will be encrypted by calling the trigonohash technique. The input(xx) is taken from input string to be encrypted, first the input is converted into radians by multiplying xx with pi/180.0 and the generated value is given to Cos function. The output of Cos function will be the input to Sin function. The output of Sin function is again manipulated which forms the final cipher text to be transmitted to destination. So the process of

encryption can be described as four step process as illustrated below:

```

for(ii=0;ii<nx;ii++) // nx → Number of input data
{
    yy11[ii]=cos(sz[i]*key/180.0); // sz[i] → ith Data
    yy12[ii]=sin(yy11[ii]*key/180.0);
        ss1=int(yy12[ii]);
        frac1= yy12[ii]-ss1;
        ss2 = Mod_Int3(ss1);
        yy12[ii]=ss2;
        yy12[ii]=yy12[ii]+frac1;
        ab1[ii]= hash1(sz[i]);
    }
int Mod_Int3(int xx)
{
    int static r1;
    int AA1,BB1,AA2,BB2;
    AA1 =xx & 0xff;
        BB1 = (xx >> 8) & 0xff;
        AA2 = (xx >> 16) & 0xff;
        BB2 = (xx >> 24) & 0xff;
    r1 = (AA1 << 24) + (BB1 << 16) + (AA2 << 8) + (BB2);
return r1;
}

```

These values(xx in radian) will go into the Cos (xx) , and the output of Sin (yy) after manipulating it with Mod_Int3 function form the cipher text. As we have changed the value of $\pi = SK_1 = 1.137201$ (for example) the result will be automatically changed as seen in the cipher text.

The general methodology for calculating cipher text is a four step process.

1) The first input is given to Cos(input*key/180.0) as Cos function takes input in radian form, the input is multiplied with $\pi/180$ here π is assigned the value of generated key.

2) The output of Cos function will be the input to Sin function i.e. the out will be given to Sin(cos output*key/180.0), as Sin also takes radian value the same procedure is adopted as above for generating output.

3) The output of Sin function is manipulated using Mod_Int3 function which rearranges the bytes in the input i.e in Sin function output least significant byte is placed most significant bytes place i.e interchanged then (least significant - 1) byte is interchanged with (most significant -1) this is repeated for all the bytes /2 in the output. The output of

Mod_Int3 is the final cipher text which is transmitted to destination along with the hash value(HA-160) of input.

Now if $yy12[ii]$ and encryption algorithm are known to the attackers then also they will not be able to breach the information. Because the value of π is not general like $\pi = 3.141592$ rather than it is 1.137201 (SK₁ or generated key) Experimental results are shown in Table III.

TABLE III. DEMONSTRATION OF ENCRYPTION PROCESS

Input Message	Key (pi)	CipherText	Hash value of Input
24	3.000000	0.015350	05c8778d2274ca4940aca31e32a71ebfc99b597d
32		0.014351	82817f1da244ca39d0dca32eb2771eaf4654610d
56		0.009919	2814db78a2388a55c37d1a4ab26adecbebe7bd68
65		0.007806	4f15dbe822088a4553ad1a5a323adebb12e8bdd8
73		0.005780	e514db78a2388a55c37d1a4ab26adecba8e7bd68

V. DECRYPTION PROCESS

Take the input message, for example “0.000256, 0.000239, 0.000165, 0.000130, 0.000096”. Here every element of input message will be decrypted by calling the trigonohash technique. The input(xx) is taken from input string to be decrypted, first in the input, the effect of Mod_Int3 function of encryption module is nullified by arranging the bytes of input in original form by using Mod_Int3 function of decryption module. This output is given to *Sin* function and the output of *aSin* function will go as input to *aCos* function. The output of *aCos* function is given to hash function(HA-160). The calculated hash code is matched with the received hash code, if matched then the output of *aCos function* is accepted as the message transmitted by sender else rejected. So the process of decryption can be described as four step process as illustrated below:

```

for(ii=0;ii<nx;ii++) // nx → Number of input data
{
    ss1=int(yy11[ii]);
    frac1= yy11[ii]-ss1;
    ss2=Mod_Int3(ss1);
    yy11[ii]=ss2;
    yy11[i]=yy11[ii]+frac1;
    kk2[i]=(asin(yy11[ii])*(180.0/key));
    kk3[i]=(acos(kk2[ii])*(180.0/key));
    qq1=(int)round(kk3[ii]);
    ab2[i]=hash1(qq1);
}

int Mod_Int3(int xx)

```

```

{
    int static r1;
    int AA1, BB1, AA2, BB2;
    AA1 =xx & 0xff;
    BB1 = (xx >> 8) & 0xff;
    AA2 = (xx >> 16) & 0xff;
    BB2 = (xx >> 24) & 0xff;

    r1 = (AA1 << 24) + (BB1 << 16) + (AA2 << 8) + (BB2);
    return r1;
}

```

These values(xx) after rearrangement with Mod_Int3 function will go into the *aSin* (xx), and the output of *aCos* (yy) after calculating the message digest of output and comparing it with received message digest, if the comparison is positive will form the plane text or message to be accepted by receiver. As we have changed the value of $\pi = SK_1 = 1.137201$ (for example), the result will be automatically changed as seen in the plane text.

The general methodology for calculating *plain text* is a four step process:

1) The input to decryption algorithm is manipulated using Mod_Int3 function which rearranges the bytes in the input i.e input least significant byte is placed most significant bytes place that means interchanged then (least significant -1) byte is interchanged with (most significant -1) this is repeated for all the bytes /2 in the input. The output of Mod_Int3 is given as input to next stage.

2) The output of previous stage acts as the input to *Sin* function i.e. the out will be given to (*aSin*(*aTan* output)*key/180.0), as *aSin* also takes radian value the same procedure is adopted as above for generating output.

3) The output of *aSin*function will be the input to *aCos* function i.e. the output will be given to (*aCos*(*aSin* output)*key/180.0), as *aSin*function also takes radian value the same procedure is adopted as above for generating output as above for generating output. The output of *aCos* is given to Hash function(HA-160) for generating message digest, then the generated message digest is compared with received message digest(from sender), if matched then only the output of *aCos* is accepted as received message(plain text) otherwise it is rejected. Experimental results are shown in Table IV.

TABLE IV. DEMONSTRATION OF DECRYPTION PROCESS

Ciphertext	Key(pi)	Hash value check	Plain Text
0.015350	3.0000	Hash value check for each cipher text	24
0.014351	3.0000		32
0.009919	3.0000		56
0.007806	3.0000		65
0.005780	3.0000		73

VI. MYSTERY KEY FOR SESSION UPDATE PHASE AND USER PASSWORD UPDATE PHASE

Right when the data correspondence time among sender and beneficiary methodologies update period, GN Server makes information for update secret key for next session(DUMK) and encodes it and sends it to both User and Sensor Node.

$$EIDUMKSs = X IDs - GN \oplus DUMK$$

$$EIDUMKUm = MM2 \oplus DUMK$$

At that point server sends EIDUMKSs (encrypted data for update mystery key for next(new) meeting) to the Sensor Node and EIDUMKUm (encrypted data for update mystery key for next(new) meeting) to the User. Subsequent to getting this data, Sensor Node quits sending information and hangs tight for another mystery key meeting. Subsequent to getting the solicitation both User and Sensor Node suspends their activity and User sends $\{ IDm, RPWm, \sigma m \}$ to GN Server and the cycle of enlistment and verification and meeting key synchronization start which brings about formation of new meeting key(i.e. second meeting so SKS₂) at both at mobile device(user) and Sensor Node(Ss) additionally at this junction, if client needs he too can change his password(PWm).

The aggregate of the above are the nuances of the lightweight arrangement. The arrangement consolidates three phases. In like manner approval stage, we present a grouping key $\{ xx1, bm, X_{GN}$ and $X_{IDs - GN} \}$ and pre realized a solitary bearing hash function(H1). It can comprehend shared approval just as hinder various attacks. Variable gathering key is proper for different applications and prevents attacks satisfactorily.

VII. DISCUSSION OF SECURITY FEATURES

In this part, we show our protocol achieves some excellent functional properties and can resist well-known attacks. Besides, the comparative analysis of our protocol and other relevant protocols are also given in this part.

A. Proper Mutual Affirmation

The versatile device(user) and GN Server can validate each other in light of the fact that toward the beginning of verification, mobile device(user) himself checks if he is certified by contrasting a worth he makes i.e $Am = A'm$. The client registers A'm esteems from the information which he supplies at the hour of verification. The Am is pre figured worth done at the hour of enrollment from the data(Bm) given by GN server. So whenever figured value(A'm) is equivalent to preexisted value(Am) which implies that client is certifiable and when GN server registers $MM5 = MM'5$, GN server gets a confirmation that the client is real Also the sensor hub get an affirmation about the GN server when it checks $MM7=MM'7$. The GN server gets an affirmation about the sensor hub when it figures $MM'8$ and contrasts it and the got $MM8$ esteem from the sensor hub i.e $MM'8=MM8$ gives a confirmation to the GN server about the sensor hub. The portable device(user) get an affirmation about GN server when he figures $MMM'10$ and contrasts it and the got $MMM10$ esteem from the GN server. The GN server and mobile device(user) has a pre shared one way hash work $H1(\cdot)$, just real GN Server and portable device(user) realizes the arranged one way hash work $H1(\cdot)$

and the GN worker and sensor hub has a pre shared one way hash work $H1(\cdot)$, just real GN Server and sensor node(Ss) realizes the arranged one way hash work $H1(\cdot)$.

B. Quickly Identification for Unapproved Login

Here the plan the Sensor Node information can be gotten to by versatile device(user) simply in the wake of passing shared confirmation and key synchronization stage and toward the beginning of the stage the client's login(identity ID), secret word and biometric needs to pass the check. Some unacceptable secret key checking instrument permits the cell phone to rapidly recognize and dismiss unapproved login brought about by wrong secret word, and the confirmation cycle can be found in the Step 1 of shared verification and key synchronization stage. Nonetheless, the protocol [17] had no such instrument to recognize wrong secret word.

C. Oppose Mobile Device Misfortune Attack

In the event that an aggressor A gets Um's cell phone ,A can recover the parameters $\{ Bm, Gen(\cdot), Rep(\cdot), XX, \tau m \}$ in the cell phone by utilizing power examination attack[18]. Here $Bm = Am \oplus (RPWm \oplus \sigma m)$, $RPWm = (PWm \oplus \tau m)$.

$Am = Bm \oplus (RPWm \oplus \sigma m)$, and σm is an irregular number extricated from Um's biometric information . As Bm, RPWm and Am are figured from obscure arbitrary numbers RPWm, τm , σm and $H1$ (one way hash work), the assaulter A will be unable to figure these boundaries precisely which may bring about finish of validation stage as the variable A'm at the hour of verification may not match so our protocol can dodge cell phone misfortune assault.

D. Customer Mystery and Intractability

Believe that a login demand message $\{ MM3, MM4, MM5 \}$ is snooped by aggressor A, where $MM3 = IDm \oplus MM2$, $MM4 = IDs \oplus MM2$ and $MM5 = H1(A' m)$ where $MM2 = bm * XX$ is critical and this key is the result of private key of user(bm) utilizing elliptic curve idea and XX is the public key of GN server utilizing elliptic curve idea i.e. key is made by Diffie-Hellman key exchange idea, assailant A can't recover $MM3, MM4$ until it have the data about the private key of GN Server or private key of mobile device(user). $MM5$ is created from Am which is known to just mobile device(user) and GN server and from hash code the assaulter A can't recover Am. So by utilizing private key and public key (elliptic curve idea and Diffie-Hellman key exchange), symmetric mystery key of GN server(X_{GN}) and one way hash work $H1(\cdot)$, the proposed convention accomplishes the component of client mystery. Here after GN Server gets the login data it can recover the Um recognizable proof (IDm) as it has the key additionally in this protocol first the client personality check is done carefully toward the beginning of validation and key synchronization phase($Am=A'm$). Likewise the component of login request message relies upon bm(user private key), in this way aggressor A can't follow User(as no fundamental data is passed straightforwardly) from public channel. In any case, the scheme [17] can't accomplish the capacity of intractability.

E. Sensor Hub Mystery

In this plan, Sensor Node's personality isn't sent out in the open channel as plaintext. At the point when Um needs to get

to the Sensor Node(Ss) information, Sn character is scrambled when Um send the solicitation to GN server for correspondence with Sensor Node Ss which GN Sensor can recover by decoding it with comparing secret key. Any aggressor without the information on relating secret word can't get Sensor Node's personality IDs and the plan accomplishes the component of sensor hub namelessness.

F. Appropriate for Internet of Things Applications

In our plan, GN Server is a confided in outsider for mobile device(user) and Sensor Node(Ss) and the verification and key synchronization stage is finished with the assistance of GN Server . With this the over-burden of both User and Sensor Node is diminished additionally it saves energy utilization of both portable device(user) and Sensor Node(Sn) (as the energy utilization of sensor hub is straightforwardly extent to the correspondence distance) also expands the life cycle of sensor nodes. Alongside when the common validation and key synchronization stage finishes, a meeting key SKS at both mobile device(user(Um)) and Sensor Node(Ss) is given by GN server, which can guarantee the resulting secure correspondence among Um and Ss. Consequently the portray design for our plan is reasonable for most IoT application. However, in some past related work [13], the client discuss straightforwardly with sensor hubs, it would lessen the existence pattern of sensor organizations and isn't reasonable for IoT applications.

G. Restrict Impersonation Attack

From the depiction of verification and meeting key Agreement stage, assailant A necessities the data of IDm and Am = (IDm ⊕ X_{GN}) to mimic as a real client to make a substantial login demand. However, from the description above, the proposed protocol provides the feature of user anonymity, and any gathering with the exception of GN can't get client's personality IDm. Likewise, Am must be recovered by Um from Bm utilizing PWm, rm and σm or can be determined by GN server utilizing IDm and GNs(secret key of GN server) when he/she gets IDm. Therefore, aggressor A can't get Am without knowing required data, and our protocol can maintain a strategic distance from client impersonation assault. Additionally, to mirror as the GN worker, the private keys xx1 and GNs(secret key of server) is fundamental data for assailant A to produce substantial correspondence messages. Nonetheless, xx1(private key) and mystery key(X_{GN}) of GN server and known to GN server just, and the proposed convention can stay away from entryway impersonation assault.

H. Oppose Replay Assault

In this plan, the single direction hash work system is received to oppose replay assault. In various strides of confirmation and key synchronization period of the plan, the hash code is created by Um, GN Server and Sensor Node for processing correspondence messages. Since the hash code carefully follow avalanche effect little change in message will create gigantic change in hash code so it would be hard for an assaulter A to figure the shared message on which the hash code are produced, so it is hard to perform replay assault. Consequently, our plan is secure from replay assault.

I. Simple Secret Key Change

In our plan, as the update time frame for key finishes. The GN worker starts a solicitation for change of mystery key for meeting. For this the GN server sends encoded message to User and sensor hub, so the cycle of progress of secret phrase begins. In the wake of accepting the solicitation both mobile device(user) and Sensor Node(Ss) suspends their activity and User sends { IDm , RPWm, σm} to GN Server and the cycle of enlistment and confirmation and meeting key synchronization start which brings about production of new meeting both at User and Sensor Node. At this intersection the mobile device (user) can likewise change his password (PWm).

TABLE V. PARAMETERS VALUES

Parameters Values(in bits)
IDm - 80
IDs - 160
X _{GN} , X _{IDS-GN} , PWm(All Password) - 160
bm,xx1,XX,MM1 - 320
Hash value - 160

Note: Values taken as per Xiong-Li[21]

VIII. PERFORMANCE ANALYSIS

A. Process Computation Time

We look at the time of the common attestation stage, which is executed using[19][20] in the three conventions (DES algorithm, Xiong -li scheme(2017) and our convention). Analysis, were led on a PC with windows7 32bit, 2.00 GHz with 2 GB of internal memory utilizing C/C++ language. The yield says that our plan accomplishes the best time execution of 31000 μs(microseconds) as shown in Fig. 8, DES algorithm 31000 μs(microseconds) and Xiong-Li(2017) common attestation time is 31000 μs(microseconds) , which is shown in the Fig. 9 beneath.

Our proposed plot is having less overhead, and can restrict various attacks. Researching Xiong-Li(2017)[21] basic affirmation in this arrangement it allows the sensor device to go about as server as it makes the mysterious expression and accommodates customer and does an incredible arrangement getting ready which is illicit of lightweight count rules. Also ChangLi(P1) plan like shared confirmation isn't done suitably, faces stolen smart card attack, further more encounters tracking attack, not applicable to practical application, delayed and costly detection of wrong password input, lack of identity detection mechanism and unfriendly mystery key change In our arrangement, we simply use XOR and XNOR action, one way Hash work and elliptic curve cryptography methodology for shared approval and meeting key synchronization. Moreover, session key is created by server and synchronized with IoT node and user (mobile device). Fig. 9 shows that our scheme takes minimum time for authentication. It is a nice technique to minimize the overhead of IoT hub.



Fig. 8. Result of Execution of Our Algorithm.

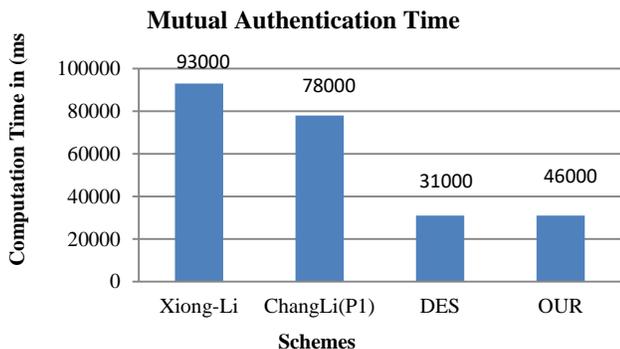


Fig. 9. Examination of Shared Confirmation Time among Xiong -Li(2017), Our Scheme and DES Algorithm.

B. Performance Analysis of Hash Algorithm(HA-160)

These three calculations HA-160, SHA1 and MD5 were tried for comparison dependent on the execution time necessities as shown in Fig. 10, Fig. 11 and Fig. 12. All the calculations have been actualized in C/C++ and run on windows 7 32-bit, CPU 2.00 GHz with 2 GB of internal memory. With the aftereffects of the analysis, it was discovered that SHA1 and MD5 requests more execution time than HA160 to create hash code.

The outcome in above Fig.11 shows time taken by three calculation (MD5, SHA1, and HA160(size of message taken by HA160 is double than SHA1)) for producing hash code for the message “ The quick brown fox jumped on the lazy dog “.



Fig. 10. Processing Time by HA-160 Algorithm“The Quick Brown Fox Jumped on the Lazy Dog”.

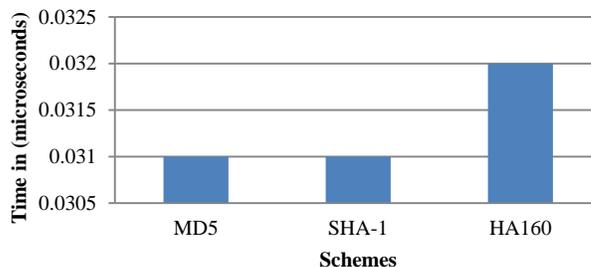


Fig. 11. Assessment of MD5, SHA1 and HA 160 Concerning Time Taken to Execute the Message “The Quick Brown Fox Jumped on the lazy Dog “ i.e. 34 Bytes.

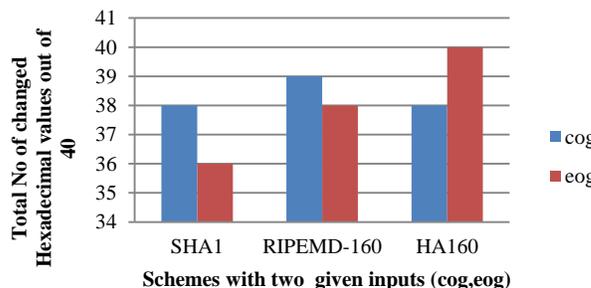


Fig. 12. Demonstrating Avalanche Effect. The Input Message used and the Two Altered Message (Dog is Replaced by Cog and Eog in Input Message) for Comparison is Shown Below.

A. Message: “The quick brown fox jumped on the lazy dog “
 Compared with
 cog :“The quick brown fox jumped on the lazy cog “
 eog:“The quick brown fox jumped on the lazy eog “

Table VI shows the message digest conveyed by three calculation (HA160, SHA-1 and RIPEMD160) for the message "The quick brown fox jumped on the lazy dog". Table VII shows the adjustment in hash code when the message in Table VI is modified for example dog is dislodged with cog, HA-160 produces a hash code with change of paying little heed to hexadecimal attributes from e and f(starting from left showed up in solid), SHA-1 yield is a hash code with change of with or without hexadecimal attributes from 9 and 1(starting from left showed up in serious) and RIPEMD 160 produces a message digest with change of paying little mind to hexadecimal attributes from f (starting from left showed up in strong) and Table VIII shows the adjustment in hash code when the message in Table VI is adapted to occurrence dog is supplanted with eog, HA-160 produces a message digest with change of each hexadecimal quality , SHA-1 yield is a hash code with change of with or without hexadecimal attributes from f, 2, 4 and 3(starting from left showed up in unprecedented) and RIPEMD 160 produces a message digest with change of with or without hexadecimal qualities from 4 and 4(starting from left showed up in strong).So on an ordinary HA-160 shows better result when any character is changed. The result is better than SHA160 and RIPEMD160 as shown in Fig. 11.

TABLE VI. RESULT

Algo/Input: "The quick brown fox jumped on the lazy dog" HA160: a5abc142 e821f1a2 48d60829 f8544618 697ea332 SHA1: 42914f6d 22976baf d4bab416 fe23b03d 4fa74963 RIPEMD160 :39349b80 fc87ec4c 887ccd87 4943f253 ee051126
--

TABLE VII. RESULT (IN INPUT DOG CHANGED TO COG)

Algo/Input: "The quick brown fox jumped on the lazydog" HA160: 58713d84 ebe7200f 24c245a4 fc197485 1c441f74 SHA1 :99e9c9dc 4698e297 efe02e13 dc3d7e97 f5355001 RIPEMD160:5fb682d2 4aa561f7 070774db 001cf3d6 dbdb7061
--

TABLE VIII. RESULT (IN INPUT DOG CHANGED TO EOG)

Algo/Input: "The quick brown fox jumped on the lazyeog" HA160: 04b36898 26446645 7659cb26 3676babb c8864a88 SHA1 :7f036e54 7b914f0f 3cadc54d 572cfa6b eb634b13 RIPEMD160:ac46cdbc 2ff1300a 1eeff31a 4f4f956c dfdb483f
--

C. Communication Overhead

To investigate the correspondence overhead, we expect that the eight messages (i.e., Message1, Message2, Message3, Message4, Message5, Message6, Message7 and Message8) are communicated during the enrollment system and the confirmation method. All the more accurately, Message1 and Message2 are sent in the enlistment strategy and Message3, Message4, Message5, Message6, Message7 and Message8 are communicated in the verification technique. Message1 incorporates the sensor's character ID_m, RPW_m(encrypted password) and uniform random number generated from Biometric information(σ_m) and Message-2 includes B_m(a value generate from GWN server from A_m) and XX(public key of GWN Server) . In addition, Message3 includes MM1, MM3, MM4 and MM5 which are calculated as follows:

$$|MM1| = |bm * PP|$$

$$|MM3| = |MM2 \oplus ID_m|$$

$$|MM4| = |MM2 \oplus ID_s|$$

$$|MM5| = |H1(A' m)|$$

Moreover, the parameters MM6 and MM7 in Message 4 , are calculated as follows:

$$|MM6| = |X'_{ID_s - GN} \oplus ID'_m|$$

$$|MM7| = |H1(X'_{ID_s - GN})|$$

then, the parameter MM8 in Message 5 , is calculated as follows:

$$MM8 = H1(X_{ID_s - GN} \oplus ID'_m)$$

also, the parameter MM9, MM10 and MMM10 in Message 6 , is calculated as follows:

$$|MM9| = |X'_{ID_s - GN} \oplus SKS|$$

$$|MM10| = |MM'2 \oplus SKS|$$

$$|MMM10| = |H1(A''m \oplus ID'_m)|$$

Finally, the parameter MM11 in Message 7 and MM12 in Message 8 is calculated as follows:

$$|MM11| = |H1(MM10)|$$

$$|MM12| = |H1(MM9)|$$

Table V contains the setting of boundaries that we have accepted for assessing the correspondence overhead of the proposed component as shown in Table IX. Subsequently, the general transmission capacity overhead of the proposed component is determined as follows:

$$bw = \sum_{i=1}^8 \text{Message } i$$

Message 1 = |ID_m| + |RPW_m| + | σ_m (random number)| = 400 bits
 Message 2 = |B_m| + |XX| = 480 bits
 Message 3 = |MM1| + |MM3| + |MM4| + |MM5| = 1120 bits
 Message 4 = |MM6| + |MM7| = 320 bits
 Message 5 = |MM8| = 160 bits
 Message 6 = |MM9| + |MM10| + |MMM10| = 480 bits
 Message 7 = |MM11| = 160 bits
 Message 8 = |MM12| = 160 bits

TABLE IX. COMMUNICATION COST OF OUR PROPOSED MECHANISM

Scheme	Communication Cost
Our Scheme	3136 bits
Xiong-Li[21]	2688 bits
Chang-Le's[21]	2400 bits

D. Computational Expense

To figure the computational expense of the proposed system, we have thought about the accompanying documentations: Th indicates the expense of single direction hash work, T_m signifies the expense of Multiplication activity, T_s means the expense of encryption and unscrambling activity. In view of the three parts (i.e. brilliant sensor, client and verification worker) which are utilized in the proposed instrument, the computational expense of every segment is introduced as follows:

- 1) Smart Sensor: The brilliant sensor performs calculations just in the verification method.
- 2) User: The client performs calculations in the enlistment and verification stage.
- 3) GN Server: The GN server performs calculation activities in enlistment and validation stage. The proposed component's computational expense is delineated in Table X. Because of the way that the proposed verification system depends just on scalar duplication activity utilizing ECC, a symmetric encryption/decoding activity and hash activity.

We use the simulation result [23] where T_m(cost based only on scalar multiplication operation using ECC) = 1.226ms, T_s(cost based only on symmetric encryption/decryption operation) = 2.049 μ s and T_h(cost based only on hash operation) = 2.580 μ s.

TABLE X. COMPUTATIONAL COST OF OUR PROPOSED MECHANISM

Scheme	User(Mobile device)	GWN Server	Smart Sensor	Total Cost
Our Scheme	3*Th + 2*Tm	6*Th + 1*Tm	3*Th	2482.96 μ s
Xiong-Li[21]	7*Th+ 2*Tm + 2Ts	8*Th + 1*Tm + 4Ts	4*Th + 2Ts	3743.412 μ s
Chang-Le's[21]	6*Th + 2*Tm	8*Th	5*Th + 2*Tm	4953.000 μ s

Note: Considering the process computation time for mutual authentication, computation cost, communication cost and other indicators, our algorithm performs well in most of the cases. The most prominent advantage is its process computation time for mutual authentication and communication cost.

E. Encryption and Decryption Algorithm Analysis

The investigation of the proposed trigonohash calculation for encryption and unscrambling has been done and exhibited in Table III and Table IV. The Encryption and Decryption Algorithm was coded in C/C++ Language[19][20]. It was compiled with MinGW-GCC 4.8.1, on the Core 2 Duo Processor, 2.00 GHz under windows 7 OS(32 bit). The investigation boundaries are plain content size (in Bytes) and time taken in encryption and unscrambling (in μ seconds).

a) Comparative PerformanceAnalysis: The proposed Trigonohash algorithm(LW-algorithm) has been compared with other existed algorithms like RC4, Hill-Cipher[19][20], RSA, Present(Block Cipher with key 80 bits and plain content 64 bits or 8 bytes) [27] and ELSCA(LW-Stream Cipher) - (Varient-HassanNoura(2019(key consist of 4 location)) stream cipher)[22]. Likewise, the results are shown below. From the going with Fig. 13 and Table XI we can say that the run time unpredictability of trigonometric Algorithm is underneath than the other existed Algorithm plot.

TABLE XI. SAMPLE RESULT FROM DIFFERENT ALGORITHM.(RESULT IN MICROSECOND)

No of Bytes	RC4 (stream cipher)	Hill-Cipher (Symmetric encryption)	RSA (public key encryption)	Present(LW-Block Cipher(64 bits))	ELSCA(LWStreamCipher-2017)	Trigonohash(LW-algorithm)
12	31000	16000	31000	47000	15000	00000
20	46000	46000	31000	78000	46000	16000
28	62000	63000	47000	125000	47000	31000
36	93000	78000	78000	156000	93000	62000

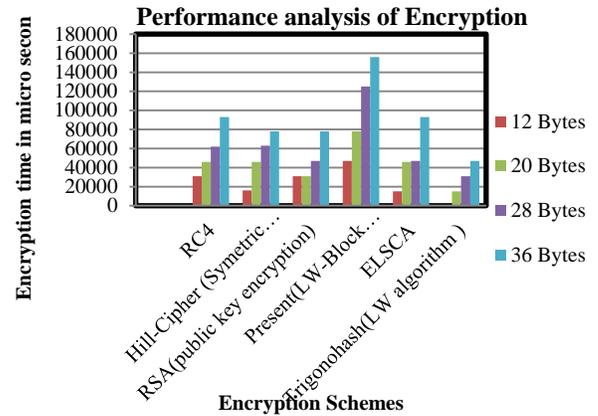


Fig. 13. Execution Examination of Encryption Algorithm of different Protocol.

IX. CONCLUSION

In this paper, we propose a strong and weighty cryptographic arrangement for Industrial Internet of Thing using one way hash work, elliptic curve cryptography and trigonohash(trigonometry and hash work) idea and endeavored to improved security and execution of a cryptographic arrangement for lightweight contraption using key generation. It gives keen and feasible instruments for basic affirmation, meeting key synchronization and meeting key update. They are proper for IoT hub with limited handling resources and force. First thing, we present the nuances of shared affirmation using elliptic curve cryptography and one way hash work and a way to deal with hinder replay attack without timestamp. Besides, again elliptic curve cryptography and one way hash idea is progressed to simply recognize meeting key synchronization lastly by using mathematical thoughts data encryption and unscrambling are done to diminish the overhead of IoT center points. In addition, we have taken a gander at the security and execution of our arrangement with some lightweight approval and cryptographic plans. The arrangement is lightweight anyway can thwart attacks satisfactorily, which is useful for guaranteeing the security of the correspondence between IoT center points, users(mobile gadget) and server.

The proposed figuring for encryption isn't simply giving the quick data encryption anyway it gives an unrivaled security diverged from other count through a most grounded key as well. The assessment of π is changing each time which is created from the common key that makes the count adequate. The estimation moreover makes the cryptanalysis cycle complex. Since there are dark variables will be more and the amount of conditions will be less when stood out from other standard encryption computations. This calculation could be well valuable for online applications like secure web based talking, data transmission between resource constraint devices and health sector.

REFERENCES

- [1] L. Atzori, A. Iera, and G. Orabito, "The Internet of Things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [2] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on industrial informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.
- [3] M. S. Hossain and G. Muhammad, "Cloud-assisted industrial internet of things (iiot)-enabled framework for health monitoring," *Computer Networks*, vol. 101, pp. 192–202, 2016.
- [4] R. Amin, S. H. Islam, G. Biswas, M. K. Khan, L. Leng, and N. Kumar, "Design of an anonymity-preserving three-factor authenticated key exchange protocol for wireless sensor networks," *Computer Networks*, vol. 101, pp. 42–62, 2016.
- [5] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in internet of things: The road ahead," *Computer Networks*, vol. 76, pp. 146–164, 2015.
- [6] M. L. Das, "Two-factor user authentication in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1086–1090, 2009.
- [7] A. K. Das, P. Sharma, S. Chatterjee, and J. K. Singh, "A dynamic password-based user authentication scheme for hierarchical wireless sensor networks," *Journal of Network and Computer Applications*, vol. 35, no. 5, pp. 1646–1656, 2012.
- [8] D. Wang and P. Wang, "Understanding security failures of two-factor authentication schemes for real-time applications in hierarchical wireless sensor networks," *Ad Hoc Networks*, vol. 20, pp. 1–15, 2014.
- [9] K. Xue, C. Ma, P. Hong, and R. Ding, "A temporal-credential-based mutual authentication and key agreement scheme for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 316–323, 2013.
- [10] Q. Jiang, J. Ma, X. Lu, and Y. Tian, "An efficient two-factor user authentication scheme with unlinkability for wireless sensor networks," *Peer-to-peer Networking and Applications*, vol. 8, no. 6, pp. 1070–1081, 2015.
- [11] H.-L. Yeh, T.-H. Chen, P.-C. Liu, T.-H. Kim, and H.-W. Wei, "A secured authentication protocol for wireless sensor networks using elliptic curves cryptography," *Sensors*, vol. 11, no. 5, pp. 4767–4779, 2011.
- [12] W. Shi and P. Gong, "A new user authentication protocol for wireless sensor networks using elliptic curves cryptography," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 730831, 7 pages, <http://dx.doi.org/10.1155/2013/7308312013>.
- [13] Q. Jiang, J. Ma, F. Wei, Y. Tian, J. Shen, and Y. Yang, "An untraceable temporal-credential-based two-factor authentication scheme using ecc for wireless sensor networks," *Journal of Network and Computer Applications*, vol. 76, pp. 37–48, 2016.
- [14] X. Li, J. Niu, S. Kumari, F. Wu, A. K. Sangaiah, and K.-K. R. Choo, "A three-factor anonymous authentication scheme for wireless sensor networks in Internet of things environments," *Journal of Network and Computer Applications*, 2017, doi:10.1016/j.jnca.2017.07.001.
- [15] X. Li, J. Niu, M. Z. A. Bhuiyan, F. Wu, M. Karuppiah, and S. Kumari, "A robust ecc based provable secure authentication protocol with privacy protection for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–1, 2017.
- [16] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2004, pp. 523–540.
- [17] C.-C. Chang and H.-D. Le, "A provably secure, efficient, and flexible authentication scheme for ad hoc wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 357–366, 2016.
- [18] T. S. Messerges, E. A. Dabbish, and R. H. Sloan, "Examining smart-card security under the threat of power analysis attacks," *IEEE Transactions on Computers*, vol. 51, no. 5, pp. 541–552, 2002.
- [19] 3GPP TS 35.201 V14.0.0, Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 1: f8 and f9 specifications, 2017.
- [20] 3GPP TS 35.202 V14.0.0, Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 2: Kasumi algorithm specifications, 2017.
- [21] Xiong Li, J. Peng, J. Niu, F. Wu, J. Liao, Kim Kwang, R. Choo, "A Robust and Energy Efficient Authentication Protocol for Industrial Internet of Things" DOI 10.1109/JIOT.2017.2327-4662 (c) 2017 IEEE.
- [22] Hassan Noura and Ali Chehab "An Efficient and Secure Variant of RC4 Stream Cipher Scheme for Emerging Networks" 978-1-5386-5657-0/18/\$31.00_c 2019 IEEE.
- [23] D. Wang, D. He, P. Wang, and C.-H. Chu, "Anonymous two-factor authentication in distributed systems: certain goals are beyond attainment," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 4, pp. 428–442, 2015.
- [24] Venkateswara Rao Pallipamu, K. Thammi Reddy, P. Suresh Varma "ASH-160: A novel algorithm for secure hashing using geometric concepts" <http://dx.doi.org/10.1016/j.jisa.2014.05.0012214-2126> © 2014 Elsevier.
- [25] Venkateswara Rao Pallipamu, K. Thammi Reddy, P. Suresh Varma "ASH-512: Design and implementation of cryptographic hash algorithm using co-ordinate geometry concepts" <http://dx.doi.org/10.1016/j.jisa.2014.10.006.2214-2126> © 2014 Elsevier.
- [26] Venkateswara Rao Pallipamu, K. Thammi Reddy, P. Suresh Varma "Design and implementation of geometric based cryptographic hash algorithm: ASH-256" pp 275-291, [Rao* et al., 5(7): July, 2016] IC™ Value: 3.00 2016 IJESRT.
- [27] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. Robshaw, Y. Seurin, and C. Vikkelsoe, "Present: An ultra-lightweight block cipher," in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 2007, pp. 450–466.
- [28] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wirel. Commun.* 1 (4) (2002) 660–670.

Detecting Malware Infection on Infrastructure Hosted in IaaS Cloud using Cloud Visibility and Forensics

Lama Almadhoor¹

Department of Computer Science
Jouf University, Sakaka, KSA

A. A. bd El-Aziz²

Department of Information Systems
Jouf University, Sakaka, KSA
Department of IST, Cairo University
FGSSR, Egypt

Hedi Hamdi³

Department of Computer Science
Jouf University
Sakaka, KSA
University of Manouba, Tunisia

Abstract—Cloud computing has been adopted very rapidly by organizations with different businesses and sizes, the use of cloud services is rising at an unparalleled rate these days especially IaaS services as cloud providers offer more powerful resources with flexible offerings and models. This rapid adoption opens new surface attacks to the organizations that attackers abuse with their malware to take advantage of these powerful resources and the valuable data that exist on them. Therefore for organizations to well defend against malware attacks they need to have full visibility not only on their data centers but also on their resources hosted on the cloud and don't take their security for granted. This paper discusses and aims to provide the best approaches to achieve continuous monitoring of malware attacks on the cloud along with their phases (before, during, and after) and the limitations of today's available techniques suggesting needed developments. Logging and forensics techniques have always been the cornerstone of achieving continuous monitoring and detection of malware attacks on-premises, this paper defines the best methods to bring loggings and forensics to the cloud and integrate them with on-premises visibility, thus achieving the full monitoring over the whole security posture of the organization assets whether they are on-premises or on the cloud.

Keywords—Malware attacks; infrastructure as a service (IaaS); amazon web services (AWS); malware detection; cloud forensics; visibility

I. INTRODUCTION

The cloud is a technology that's not new anymore. Nowadays, using cloud services is increasing at an unprecedented pace [1], it has become more popular after the advent of the Fourth Industrial Revolution (IR 4.0) [2] In 2020, about 83% of business workloads operate in the cloud, and a whopping 94% of companies now use a cloud service in one form or shape [3]. There are three most utilized cloud services include Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).

Infrastructure as a Service (IaaS) is one of the most critical and fastest-growing services in Cloud Computing. As shown in Fig. 1, according to BMC the growth rate from 2018 to 2022 of IaaS is projected to be 18% higher than that of other cloud services.

In this service model, cloud providers provide resources to users/machines such as virtual machines (VMs), raw (block) storage, firewalls, load balancers, and network devices. Resource management for IaaS gives the following advantages Quality of service, scalability, reduced overheads, optimum utility, increased throughput, specialized setting, reduced latency, a streamlined interface, and cost-effectiveness. Virtualization technology is used to offer Infrastructure which enables multiple consumers or tenants to share the same hardware. Virtual machines (VMs) play an important role in Cloud Computing because they allow powerful and systematic use of the hardware available [4].

All these outstanding features make it simple and convenient to access the IaaS service. However, many individuals use this technology in an effective way and few challenges to use it [5]. Infrastructure hosted on the IaaS cloud is becoming targets to many attacks like malware for the following reasons:

1) Cloud service providers steadily offer higher performance with high computation power for their customers. These VMs are big targets for crypto currency mining malware, which are becoming more sophisticated to take the resources of the server without getting noticed.

2) The increase of remote working and globally dispersed workforce and application accessibility especially after the COVID 19 give the attackers more chances to hide their malicious traffic to compromise the cloud-hosted VMs, and use them for their malicious campaigns (phishing campaigns, botnet command, and control, so on).

3) The increase in IoT applications that use cloud-hosted infrastructure to analyze the enormous amounts of data generated by these applications to create business value and insights. Most of these IoT appliances are built with no or weak information security measures thus attackers can easily get their way to the backend cloud-hosted VMs through these IoT devices and applications.

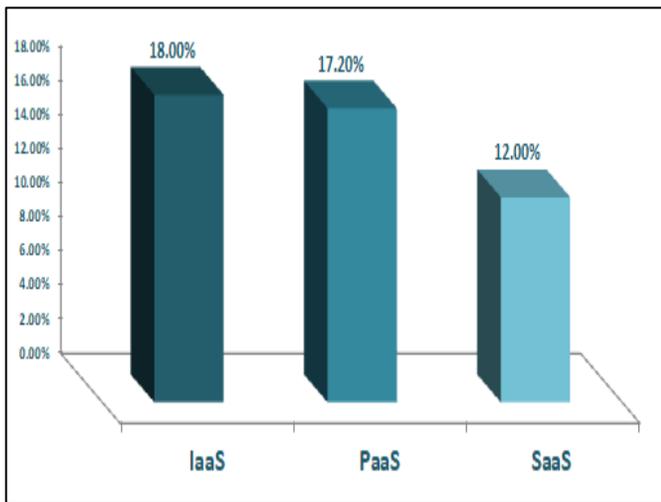


Fig. 1. The Growth Rate (from 2018 to 2022) of cloud Services.

Malware protection is becoming increasingly essential for cloud network security, as it presents a significant threat to network security. Since malware security is minimal, only PC-based protection is up to date [6]. According to Noëlle and others [7] present attacks in the IaaS cloud can be investigated using VMI-based mechanisms. [8] Gives a summary of the different malware detection methods used for intrusion detection. Various machine learning methods can be used to provide a Cloud Computing detection mechanism. In that way, an improved technique is needed to ensure effective intrusion detection for such techniques. The researchers have used a malware detection technique to detect malware in [9], which is SAIDS, a security monitoring system tailored for IaaS clouds, but it only handles one form of IDS. G. Murali and N. Moses [6] in this work propose a framework that deals with anomaly detection malware at the network level, stating that malware distribution in terms of networks varies. Many studies in [10] address cloud forensics in general terms without offering details.

Therefore, this research combines technologies and forensics to monitor and detect malware attacks that targeting VMs in the IaaS cloud. Begin to classify malware attacks in the infrastructure hosted on the IaaS cloud, shed the light on the importance of cloud services visibility to thoroughly track each malicious activity on the cloud-hosted infrastructure and quickly react to malware attack, then list the different approaches and techniques used to gain this visibility and how to perform analytics on this collected data to get insights helping improve the reactive and proactive defenses. Then testing the existing technologies and methodologies for monitoring cloud-hosted infrastructure and performing digital forensics on the cloud-hosted assets and how both can be used to speed up the detection of malware attacks then quickly deploy countermeasures to better stop them from reaching their goals.

This research which addressing the security of infrastructure in IaaS cloud is noble, because it focuses on an investigation that will render the safety of this service, which is conducted on the IaaS cloud. This cloud service has dominated not only small businesses but also global

enterprises including big multinationals. As such this topic is very informative since it will detect malware in its early stages. [11] All this provides organizations a feeling of reassurance that their assets are safe and secure.

II. BACKGROUND

A. Infrastructure as a Service (IaaS) Cloud

The cloud service model's bottom tier is IaaS [12]. It is the most fundamental and critical service, offering basic computing services such as servers, networking, and storage. These resources make use of virtualization technologies to execute services. IaaS also provides users with data security, backup, and maintenance [13, 14]. Consumers have full control over these tools, which are aggregated and controlled [15]. Some companies cannot afford to purchase a computer, so instead of buying the infrastructure, it can be leased or rented according to the needs of the users. This service enhances system availability while also lowering costs and offering a more flexible system.

Common examples of this service: Amazon Web Services (AWS), Google Compute Engine (GCE), Microsoft Azure, and Cisco Metapod.

Despite all of the technical developments in IT security over the last three decades, the Latest statistics also show an increase in malware activity, Deep Instinct conducted analysis and have published a study on the hundreds of millions of attempted cyber-attacks that occurred every day in 2020, Revealing that malware increased by 358 percent overall and Ransom ware increased by 435 percent compared to 2019.

B. Malware Attacks

Malware is a term that combines the word malicious and malware, thus malware is described as software that has a malicious and harmful effect on networks, software, operating systems, or other components [16].

Malware, according to [17], is a software program that is intended to help malicious attackers accomplish their goals. It was created to help attackers accomplish their objectives. Disturbing device processes, altering or hijacking core computing functions and network resources, tracking users' behavior, and stealing, encrypting, or deleting confidential data without the user's permission are just a few of these objectives.

One of the biggest challenges in the IaaS cloud world is malware attacks; malware has long been a major concern to home and business devices, as well as cloud virtual machines [18]. Virtualization, as one of the most important Cloud Computing techniques, blurs the lines between time and space. Virtualization would undoubtedly increase resource efficiency and reduce system management costs [19]. But unfortunately, virtualization creates new vulnerabilities, which are being exploited by malware.

According to Malwarebytes' 2021 Malware Study, attackers exploited the COVID-19 public health crisis in unthinkable ways previously, not only preying on uncertainty and fear during the early months of the global pandemic, but also enhancing malware, retooling assault tactics, and

extorting targets to the tune of \$100 million. Malware is multiplying at an unprecedented pace, every day AV-TEST registers more than 350,000 new malware in 2021, and there is an increase in the total of malware compare with the last five years as shown in Fig. 2.

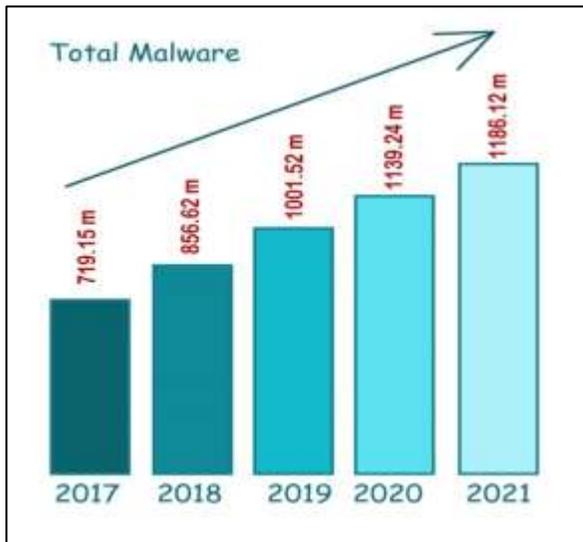


Fig. 2. Total Malware in Last Five Years.

The malware must be discovered before it affects the most resources and networks. In order to protect infrastructure hosted in an IaaS cloud, malware detection must be successful.

C. Malware Detection Methods

In the IaaS cloud world, There are several unique and has different on-demand services available for consumers. These services provide easy and straightforward access to a range of web apps. The virtualized nature of IaaS environments may be a weakness when it comes to malware [20], As a result, malicious attacks and breaches may strike at any time, destroying important apps and files [5]. One of the most difficult challenges in the creation of resilient and stable cloud-based mechanisms is the accurate detection and identification of malware. This is attributed to that malware is often the starting point for phishing, massive Distributed Denial of Service (DDoS) attacks [21], and email spamming by the use of a botnet. For that detecting malware as soon as possible is very important. To resolve the many different attacks and threats that exist in IaaS clouds, continuous and automatic security monitoring of Infrastructure has become a primary requirement [22].

There has been a rapid rise in the number of research studies on malware detection in the IaaS cloud in recent years. A variety of techniques have been proposed to detect malware. These techniques can be divided into four categories: Signature-Based Techniques, Behavior-Based Techniques, cloud-based Techniques, and Machine Learning-Based Techniques.

The procedure for extracting features differs from one Technique to the next. It is difficult to prove that one detection approach is superior to another since each approach has its

own set of benefits and drawbacks [23]. Since malware comes in a variety of shapes, sizes, and behaviors, as well as various levels of danger, the same detection methods and mechanisms cannot be used in every situation. It is impossible to detect malware with only one approach to security. [24] As a consequence, having different detection techniques for various conditions is inevitable.

The following describes in detail the approaches to monitor and detect malware attacks:

1) *Signature-based techniques for malware detection:* A signature is often a short sequence of bytes that is exclusive to each known malware and enables new files to be correctly detected with a low error average [25]. They are commonly used in commercial antivirus software to detect known malware. Traditional signature-based techniques preserve a listing of signatures that are stored in the database, if a match is found, an alert is triggered, and the database is updated [26]. These signatures are formed by disassembling and analyzing the code, there are many disassemblers and debuggers tools are available to aid in the disassembly of portable executable (PE). In this way, code is analyzed then features are extracted. As a consequence, these features play a primary and important role in creating a malware family's signature [27].

Signature-Based techniques are very quick and effective to create a signature. The major feature of this method is the precision with which they detect malware. It is capable of detecting known malware instances quickly and with a smaller amount of malware. A downside of this technique is that of being unable to detect unknown malware. When new malware is released into the market, it must wait until it has infected many systems before its signature can be created and added to the databases. These techniques also have the drawback of being unable to identify encrypted or polymorphic malware [26] that indicates they are not resistant to malware obfuscating techniques.

2) *Behavior-based techniques for malware detection:* It's also known as Heuristic-Based Detection or anomaly Detection. A behavior-based detection technique uses monitoring tools to observe the programs' behavior and identified whether is it malware or benign? The main goal of this technique is to examine the behavior of malware, known and unknown [23]. It is divided into two phases [28, 29]:

a) *The training (learning) phase:* Involves monitoring the system's behavior in the absence of an attack or malware, also a learning task is carried out in order to teach the classifier to behave normally.

b) *The testing (monitoring) phase:* The normal stage is compared with the current behavior then detects suspicious behavior, identifying anomaly activities, looking up the protocols and ports used, and indicate any malicious activities in order to inform the responsible of deviations or major variations from the baseline [30].

The main feature of the behavior-based approach is the capability to detect both unknown malware. However, there are major drawbacks of this technique which are a high False-

Positive Rate (FPR) and a long monitoring duration [31]. Further, the ability to detect malware like zero-day attacks is directly impacted by the decrease of thousands of extracted features, evaluating similarities between them, and surveillance of malware activities [32].

3) *Cloud-based techniques for malware detection*: Cloud Computing has grown in popularity as a result of its numerous benefits, including easy access, on-demand storage, and lower costs. Because the cloud is so widely used, it's also been used to detect malware. Cloud-based malware detection improves the detection performance for devices and VMs with much larger malware databases and strong and many computational resources. This detection approach employs a variety of detection agents distributed across cloud servers and provides security as a service. The users can upload any type of file and get a record of whether or not the file is malware [23]. In [33] Sun et al created a cloud-based detection malware system, called Cloud Eyes, which has provided resource-constrained devices with effective and trusted security services. Cloud Eyes detected malicious buckets through cross-filtering on the

cloud server. In [34] the researchers implemented a malware detection infrastructure realized by an intrusion detection system (IDS) with cloud and mist computing to overcome the IDS sending problem in brilliant objects due to their constrained resources and heterogeneous sub networks.

IaaS cloud is the most flexible model. Users have more choices when it comes to performing IDS over this infrastructure. Intrusion Detection Systems (IDS) can be used in a variety of ways over the IaaS cloud layer [22], including:

4) *Network-based Intrusion Detection Systems (NIDS)*: All network packets are collected and analyzed in the cloud environment using signature or behavior-based detection approaches to identify malicious events and activities like port scanning, DoS attacks, user to root attacks, etc [35-37]. It's used to keep track of network traffic between VMs and host machines, as well as between VMs [22].

There are several models for adapting NIDS to the cloud shown in Table I:

TABLE I. VARIATION BETWEEN CLOUD NIDS/NIPS MODELS

Cloud NIDS model	Ease of management	Cost of operation	Cost of implementation	Customization and Integration capabilities	Level of visibility
IDS on-premises and usage of VPC endpoints	Easy to be manageable as the on-premises IDS scope will just be extended to the cloud environment	High cost due to data transfer expenses as all data will be sent from the cloud environment to the on-premises IDS for inspection. plus, the cost of operating high-performance VPN tunnel	low cost as it doesn't involve purchasing new appliances or special cloud deployments or modifying network architectures	The IDS rules and policies will be consistent over both the on-premises environment and the cloud. The cloud will just look like an extension to the on-premises environment	The IDS will be able to monitor the traffic going in and out the cloud environment but there is no visibility on inter traffic within the cloud environment itself
IDS on VPC NAT instances or dual-homed systems	Easy to deploy model but hard to be manageable because the organization is responsible for the configuration and customization for each aspect of the IDS and keeping it in pace with new threats and attacks	High cost of operation One single point of failure that can lead to services unavailability when it's malfunctioned or misconfigured Needs high skilled engineers which can be expensive	The cost of implementing a VPC NAT instance will depend on the size of the VPC and its instances which can be Medium to high cost	Highly Customizable Security teams can add up new features at no cost Highly customizable dashboards and deferent detection mechanisms Due to the high customization, there won't be obstacles bringing the rules and policies from on-premises IDS, thus providing consistency along both environments	Has visibility on the cloud traffic that will go only through the cloud gateway Doesn't extend the capabilities of the on-premises IDS
Usage of 3rd party AMI as the NIDS	Custom route and traffic control is required Requires distinct network zone for the appliance for centralized monitoring Management of the appliance is the responsibility of the security team for the organization which adds extra tasks and extra appliances to be managed	Cost will depend the amount of data fed to the appliance As it's another appliance to be managed, there will be a cost of management will be added in terms of providing the adequate numbers to manage organizations' appliances	Cost will depend on the Vendor, the features purchased and the amount of data fed to the appliance which will range from Medium to High Cost Due to the high cost If the appliance technology differs from the IDS used on premises, then there will be additional costs for team training or hiring new skills	Customization is limited to the capabilities and features available of the purchased appliance If the technology of the cloud-based IDS differs from the on-premises IDS then there might be rules and policies inconsistency between both environments	Has visibility on the cloud traffic that will go only through the appliance which is to or from the VNets Due to the pay per the data ingested approach organizations usually don't integrate it with the test and pre-production environments on the cloud

5) *Host-based Intrusion Detection Systems (HIDS)*: At the host level, data is processed, tracked, and analyzed. It monitors and detects modifications in the host kernel, program behavior, and file system [38-40]. These IDS may be installed on a host, virtual machine (VM), or hypervisor (VMM) to detect intrusion events by analyzing device logs against user credentials and access control (AC) policies [41]. In this way, Customers then notify managers if they notice any abnormal activity [22].

The cloud user is in charge of monitoring HIDS deployed on a VM, while the cloud provider is in charge of HIDS deployment on the cloud [42].

6) *Distributed Intrusion Detection Systems (IDS)*: Multiple intrusion detection systems (IDSs) (such as NIDS and HIDS) are deployed over a wide network to track and analyze traffic patterns for intrusion detection, and they can function independently or collaboratively [43,44].

7) *VMM/Hypervisor-based Intrusion Detection Systems (HypIDS)*: Hypervisors that host VMs can easily access performance data; these data offer insight into the activities taking place inside a virtual machine without requiring direct knowledge of the virtual machine's operating system, software, or private data [45].

This type of IDS is installed at the hypervisor layer and monitors and analyzes information transmitted in communications between VMs (i.e., VM-VM), between the VM and the hypervisor, and between the cloud environment and the outside world [46, 47].

8) *Machine Learning-Based Techniques Malware Detection*: Since Machine learning (ML) can be generalized to never-before-seen malware families and polymorphic strains, machine learning is a common approach to signatureless malware detection [48]. There are a lot of Well-known ML algorithms that particularly useful in behavior-based detection and other detection methods like Artificial Neural Network (ANN), Decision Tree(DT), XGBoost, naive Bayes (NB), Associative Classifier (AC), C4.5 decision tree variant (J48), random forest tree (RF), k-nearest neighbor (KNN), support vector machine (SVM), logistic model trees (LMT), Shared nearest neighbor (SNN), multilayer perceptron (MLP), Bayesian network (BN), simple logistic regression (SLR), RIPPER, Deep Learning (DL), and sequential minimal optimization (SMO) [49-54].

D. Cloud Forensics

Cloud Digital Forensic techniques are typically used to gathering and preserving evidence, reconstructing incidents, deciding how, where, and where an incident happening, and producing threat information. Threat information includes Indicators of compromise that can be used to help an organization defend itself.

III. METHODOLOGY

The methodology has been divided into two practical parts:

The First: when the malware attack happened, make cloud analysis for malware detection.

The Second: is Forensics Analysis in the IaaS Cloud after the malware attack happens.

A. Cloud Analysis to Malware Detection

In this practical part, flow many steps as shown in Fig. 3:

1) *Choosing test environment*: The tests were performed on Amazon Web services (AWS) hosted infrastructure. choosing the Amazon Web services (AWS) for this research because it the market leader for public cloud services offering and has a wide service catalog making it a suitable choice for most organizations, Named as a Leader in Gartner's Infrastructure as a Service (IaaS) Magic Quadrant for the 7th Consecutive Year. (AWS) innovated many tools and techniques for data collection, monitoring, analysis for their customers which most of the other cloud service providers follow.

2) *Data set*: Fortunately, there are community initiatives that define and classify each cloud attack technique publicly witnessed; such as the NIST Cybersecurity Framework [55] and MITRE ATT&CK cloud framework.



Fig. 3. Flow Chart for Cloud Analysis to Malware Detection.

To define the data to be collected by using the MITRE ATT&CK framework, the combined techniques used to compromise cloud-hosted infrastructure which can be leveraged by the attackers to gain initial access and control over the environment- along with the malware techniques used in enterprise infrastructure. With that combination, the organizations implement continuous monitoring for their infrastructures whether it's on-premises infrastructure or cloud-hosted one. For the scope of research, Continuous monitoring on IaaS can be accomplished by gathering and processing the following [56]:

- API calls Monitoring (In AWS it can be achieved through CloudTrail's logs).
- Host logs and logs of deployed Host Intrusion detection System (HIDS).
- VPC flows.
- Logs of the cloud resources (in AWS it's the CloudWatch Logs) [57].
- Image and instance integrity validation.
- Automation through tools like AWS Lambda and AWS Config.

3) *Testing and analysis*: Performed the malicious activities performed by the malware without using real malware in this environment.

Use many tools like:

- Amazon CloudWatch: Collect and track metrics, collect and monitor log files, set alarms, and automatically react to changes.
- AWS CloudTrail: A web service that logs your account's AWS API calls and provides you log files.
- AWS Config: provides a comprehensive view of the AWS resource configuration in your AWS account. This involves how the services are connected to one another and how they were previously configured, allowing you to track how the settings and relationships change over time.
- SIEM software: Providing data analysis, event correlation, aggregation, reporting, and log management.
- SIFT is open-source: Includes most tools required for digital forensics analysis and incident response examinations.

Implementation of many testing labs:

a) *Create a billing alarm for AWS account*: According to the MITRE ATT&CK framework for cloud attacks, one of the most used attack vectors for Cloud attacks and malware attacks targeting cloud-hosted environments is cloud account takeover. There are many ways to detect cloud account takeover, one of the best ways is detecting changes in the usual billing as most cloud malware attacks aim to abuse the environment's resources or deploy new resources. Most public

cloud providers provide features to enable their customers to create billing and send them emails when these alarms are triggered

To create a billing alarm for the AWS account that can be used later in detecting any suspicious abuse of IaaS resources, first enable receive billing alerts for the account and then use AWS CloudWatch, to create a metric that will trigger an alarm whenever the billing exceeds a specific threshold.

Also, use the AWS Simple Notification Service (SNS) for sending an email once the alarm is triggered.

Results

When malware misuses the resources of the cloud or publishes expensive new resources, AWS calculates the charges and the estimates charges as per your approach and now have a threshold of 5\$ that whenever the charges are exceeded the alarm will be triggered as shown in Fig. 4 and receive an email as a notification.

b) *Perform continuous monitoring in the AWS environment*: AWS offers a service called AWS Config, this service allows monitoring AWS resource configurations and track resource inventory and changes, which can be used to detect any malicious configuration changes the attacker tries to make to gain control or persistence over the compromised account's resources. This monitoring feeds then can be consumed using AWS CloudWatch and SNS Notifications can be created based on them.

Malware attacks target and modify the data stored and any misconfigured cloud storage leading to leaked data. By using AWS Config to make many rules like sure storage versioning is enabled for AWS storage (S3). By enabling the s3-bucket-versioning-enabled rule, another action performed by attackers is to try to hide their malicious API calls by disabling API calls monitoring, configured a rule to detect if cloudTrail enabled or not and another rule to detect whether the volumes used are encrypted or not. Also to prevent the misuse of the root account, enable another rule to detect whether multi-factor authentication (MFA) is enabled for the root account or not.

In Fig. 5 the Dashboard of AWS Config detected the security issues in resources, which whether indicate a misconfiguration or malicious change.

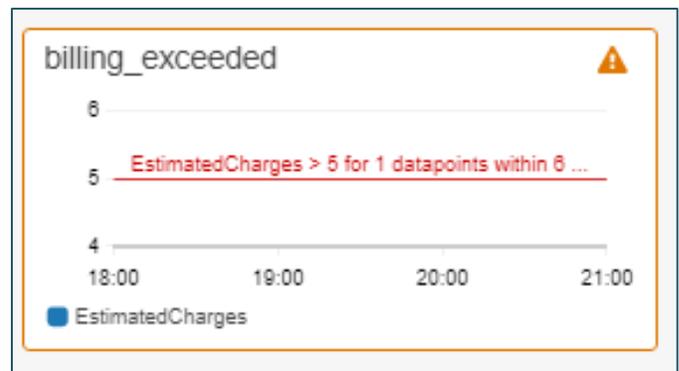


Fig. 4. Alarm for Exceeding Bill for AWS Account.

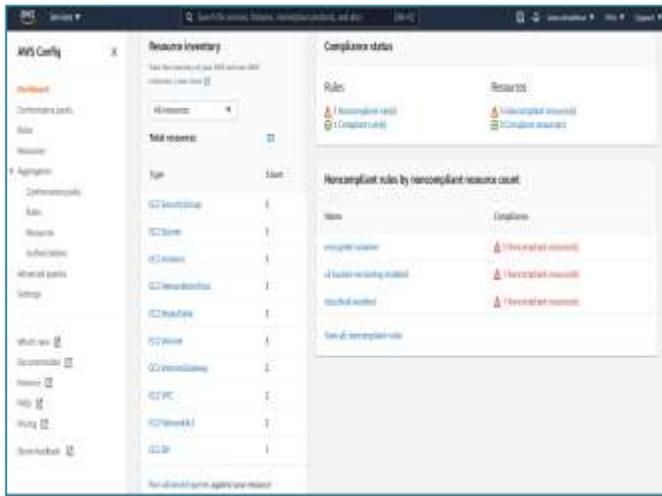


Fig. 5. AWS Config Dashboard.

Using the AWS Config resource inventory to view existing or deleted resources recorded. For a specific resource, view the resource details, configuration timeline, or compliance timeline. The resource configuration timeline allows you to view all the configuration items captured over time for a specific resource. The resource compliance timeline allows you to view compliance status changes. For example, used AWS Config to track the timeline of changes for an s3 bucket resource.

Results

- ✓ By using AWS Config detected Malware attacks that modify the data stored, and target any misconfigured cloud storage.
- ✓ By using rules in AWS Config can see if malware attacks try to hide their malicious API calls by disabling API calls monitoring.

Limitations

- ✗ Not all monitoring options are available for each region, cloud providers may offer their services to a specific region but monitoring these services may not be available for this region or came late.
- ✗ The AWS config service takes a few minutes to rebuild asset inventory, thus not providing real-time monitoring.

c) *Monitoring cloud API Calls:* In AWS CloudTrail is used to monitor account activity and API calls, this is a very important feature as cloud providers offer their services via APIs. CloudTrail feeds can be integrated with CloudWatch to create metrics generating alarms for any suspicious account's behavior or any account misuse.

Create a trail; I enable ongoing delivery of events as log files to an Amazon S3 bucket. Then get the logs and API Calls from CloudTrail Event history.

Fig. 6 detects console logins (recording the account used and the source IP that the user has when he logged in) which reveals any attempts to login from suspected places or countries.

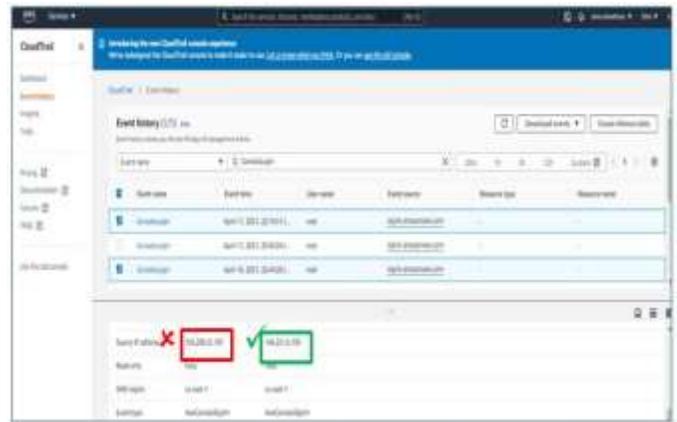


Fig. 6. Detecting Console Logins in AWS CloudTrail.

Also can find that AWS generated logs are formatted in JSON.

CloudTrail has a lot of great, very detailed log data. also, By setting up a CloudTrail trail I deliver CloudTrail events to Amazon S3 then review this in the console or view/download from S3.

Results

- ✓ Detect console logins from suspected places or countries.
- ✓ Because most of the time, organizations transfer cloud logs to their own data center for long term storage and correlation with other on-premises tools, it's very important to generate the logs in a text-like format such as JSON, thus enabling serialization of complex, high-quality log data and decouple the interpretation of logs from specific solution or vendor. From a test, notice AWS uses this concept in their generated logs and flows.
- ✓ Leverage the storage API (s3 API) to import cloud trails to a search and indexing platform or security management systems like (SIEM solution) for building more unified and robust use cases monitoring the security posture over the entire environment.

Limitations

- ✗ By default, there are no trails configured to log any of AW'S activities. Must enable what wants to monitor.
- ✗ CloudTrail Logs directly in the console is not as efficient as pulling the logs into some tool for parsing and normalization with better searching and filtering features.

d) *Generating VPC network flows (VPC Flows):* Producing flow logs of communications that occur in the VPC (north-west communications and east-west communications) is very crucial to detect activities related to malware attacks such as communications with malicious IPs and command and control servers, attempts of pivoting and lateral spreading, and suspicious communications behaviors and data exfiltration.

CloudWatch, created a flow log group to watch for flows within VPC [58] while one instance is performing lateral scanning, to detect the scan.

By generating these flows and sending those to CloudWatch then create metrics watching for specific suspicious behaviors or policy violation communications. For example, made a metric to watch for Telnet communications that can due to policy violation or scanning attempt and received an alarm as shown in Fig. 7.

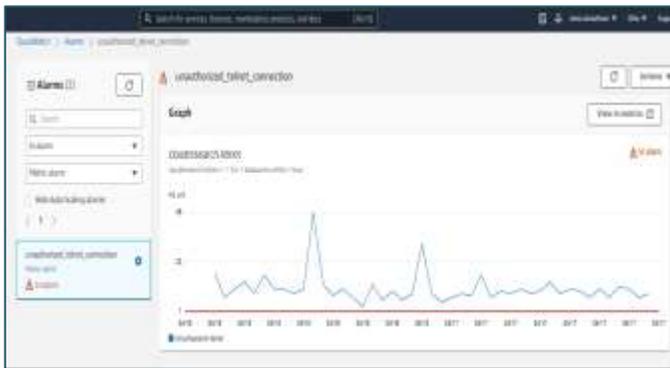


Fig. 7. Alarm for unauthorized Telnet Communication.

Also, receive this notification by email.

Results

- ✓ Seeing all flow logs of communications that occur in the VPC.
- ✓ By sending flow logs to CloudWatch then create metrics watching for specific suspicious behaviors or policy violation communications.
- ✓ Instead of sending flow logs to CloudWatch, can send them to s3 bucket storage to be aggregated with other log sources and offloaded to on-premises with Storage API calls and leveraging them using SIEM Solution.

Limitations

- * Flow logs may take up to 5 minutes to appear, thus not providing real-time visibility.
- * Reviewing VPC Flow Logs directly in the console is not efficient while using CloudWatch Metrics to make use of these flows, leveraging the high potential of captured flows is achieved through sending them to some security event management tool like SIEM Solution for more flexible correlation and enrichment.

e) *Transferring logs to SIEM:* Must bringing cloud logs into a single point where they can be aggregated with on-premises security events and other security and intelligence feeds, thus enabling the threat management team to have a single pane of glass from which they can monitor the whole security posture of their organization.

To achieve this purpose, install IBM Qradar Community edition (SIEM Solution) locally and configure it to receive the AWS CloudTrail logs which configured earlier to monitor different services of the AWS account and its resources [56].

1) The first thing to notice that there are many protocol options to bring the CloudTrails Logs to SIEM solution.

2) Fig. 8 contains the different types of event formats that can be stored in s3 bucket (the important ones are AWS CloudTrail and AWS VPC Flow logs).

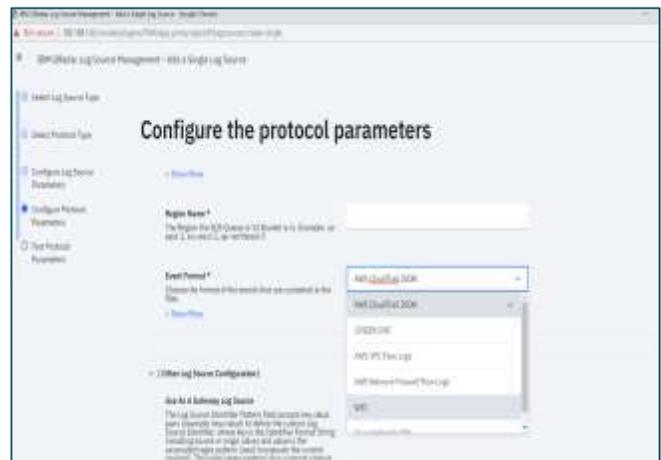


Fig. 8. Setting Options to Select Event Formats that Stored in S3 Bucket.

3) Specify the S3 bucket name and the directory where AWS CloudTrails resides.

4) The maximum frequency to poll the logs from the AWS s3 bucket is 1 minute, providing an adequate timeframe to detect attacks but still not real-time visibility.

Results

The different ways to transfer Flow logs from the cloud to threat management like SIEM is shown in Table II.

- ✓ For most IaaS use cases, the best way to transfer monitoring logs to on-premises threat management tools like SIEM is through the usage of Storage APIs, yet there are other methods suitable for the less common of today's use cases.
- ✓ Also, a statistic in Fig. 9 determines the number of logs that are generated per second from resources on the cloud, it helps us to determine the amount of visibility that gets from the logs, for example, NGFW and HDIS have many activities, so they give me the greatest amount of visibility and helps us see all the events in real-time and discover malware quickly.

B. Forensics Analysis in The IaaS Cloud

Using cloud forensics to assist companies in enhancing their incident response and threat detection capabilities [59] organizations must have sufficient forensics investigation expertise to apply to their cloud infrastructure to ascertain the root cause of an attack, detect signs of vulnerability, and better protect against IaaS malware attacks, as well as quickly locate malware and its objectives before they have an impact on the companies' operations.

In the event of a hacked virtual machine, several cloud users automatically terminate and destroy the virtual machine (VM), erasing all proof in the process.

TABLE II. THE BEST WAYS TO TRANSFER FLOW LOGS FROM THE CLOUD TO THREAT MANAGEMENT

	SysLog or SysLog Server	Storage API	Data Stream
Data transfer Speed	It's not considered a real-time data communication. Data is ready to be sent as soon It's generated or collected by the aggregation server. Because in most cases it's not compressed, it will take few minutes to arrive at the premises,	Faster than sending Syslog messages, but still will take time for the data to be written in the storage bucket. Fast but not real-time data communication.	Data processed as soon it's generated. Because it makes use of data analytics features the data transferred is small and fast. it provides real-time visibility. Example AWS Kinesis streams
Cost	High cost around 9 cents per Gigabyte	-Low cost around .007 cent per Gigabyte. -There will be additional cost for the storage bucket	Expensive service
Size of logs transferred	If there are big number of resources and no compression used, the size of logs will be big leading to more costs	High quality compressed data generated small size data transfers	Rich and Small size of data streams
Data enrichment and preprocessing	No analysis or processing is made before sending	Better quality data and fully inflated data to flows and logs	performs data analytics, correlation and enrichment with other feeds, sending only good quality actionable insights
Scalability	Not scalable as the volume of data increases	Scalable as it uses API communication methods	Scalable
Log Data protection	By default, data is sent in plain text unless sent over VPN connection or syslog collector used	the connection is encrypted	Communication is encrypted
Need for additional tools or management	For optimum experience VPN tunnel and syslog collector are used	configure additional storage to store the logs	Purchasing the data analytics stream service

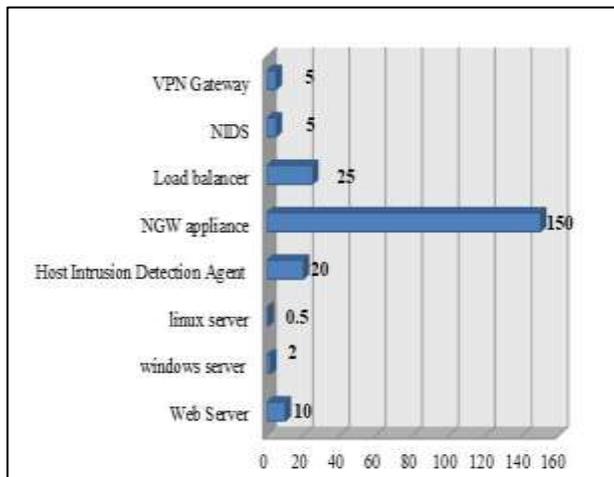


Fig. 9. Logs Generated from Cloud Resources Per Second.

It can be difficult to plan for forensics in the cloud. Until recently, there have been few tools to assist analysts in inspecting applications and collecting data [60]. When it comes to gathering and analyzing evidence, must look for the following:

- Network packet captures (PCAPs) for network forensics.
- Memory for instance.

- A disk for instance.
- Event data and logs.

Recently, more vendors and community members have been concentrating their efforts on resolving forensics issues in the cloud. How do we best gather evidence? How do we store it properly? What tools work in the cloud well?

1) Evidence Capture

- Capture a disk is getting easier in a running instance. You take a snapshot of EBS (Elastic Block Store) in Amazon EC2 (Amazon Elastic Compute Cloud) [61], after that, attach it to a forensic workstation (covered in a moment). The Azure, You can grab IaaS OS and Data drives directly from the portal.
- Capturing memory in a shared environment would necessitate some form of per-instance basis capture. To put it another way, instances' running memory would have to be acquired using separate tools (local or remote).

2) Test environment preparation

3) Provisioning the Forensic machine and installing the required forensics investigation tools. Fortunately, a package that provides access to most of the forensics tools from one executable package is called SIFT. Prepare forensics investigation machine as follows:

a) create an instance that will be used to perform the investigation. In a test, named it cloudresearch-forensics-instance.

b) SSH into the cloudresearch-forensics-instance to download the SIFT tools as shown in Fig. 10 [in research time the latest SIFT version was 1.10.0-rc5].

```
ubuntu@ip-10-0-0-11:~$ sudo wget https://github.com/traefik/sift-cli/releases/download/v1.10.0-rc5/sift-cli-linux-3821-45-45-22-45-50 -- https://github.com/traefik/sift-cli/releases/download/v1.10.0-rc5/sift-cli-linux
Resolving github.com (github.com)... 149.12.124.4
Connecting to github.com (github.com)|149.12.124.4|:443... connected.
HTTP request sent, awaiting response... 302 Found
Location: https://github-releases.githubusercontent.com/33119626/66046f90-6219-11eb-79cc-6741e264ab70?mac=4
ast=3A2F934Fae4_request=AM-Get=332158222450844&mac-Expires=300&mac-Signature=62c242277b11e1e1255
key_id=4e9e0_id=92619626&response-content-disposition=attachment;filename=sift-cli-linux&response-co
-2921-45-45-22-45-50 -- https://github-releases.githubusercontent.com/33119626/66046f90-6219-11eb-79cc-6741e
264ab70?mac=4e9e0_request=AM-Get=332158222450844&mac-Expires=300&mac-Signature=62c242277b11e1e1255
&ast=3A2F934Fae4_request=AM-Get=332158222450844&mac-Expires=300&mac-Signature=62c242277b11e1e1255
Connecting to github-releases.githubusercontent.com (github-releases.githubusercontent.com)|149.12.124.4|:
443... connected.
HTTP request sent, awaiting response... 200 OK
Content-Length: 8766209 (8.4M) [application/octet-stream]
Saving to: 'sift-cli-linux'

sift-cli-linux          88%====
3821-45-45-22-45-50 (64.8 MB/s) - 'sift-cli-linux' saved [8766209/8766209]

ubuntu@ip-10-0-0-11:~$ ls
sift-cli-linux
ubuntu@ip-10-0-0-11:~$ cd sift-cli-linux
ubuntu@ip-10-0-0-11:~$
```

Fig. 10. Download the SIFT Tools.

c) For ease of access move the executable file to the account binaries directory

```
mv ./sift-cli-linux /usr/local/bin/sift
chmod +x /usr/local/bin/sift
sudo sift install
```

- 1) Create a snapshot from the instance to perform forensics analysis on it.
- 2) Create a volume from it with the snapshot in the same available zone as cloudresearch-forensics-instance.
- 3) The Evidence Volume will be available to be attached to cloudresearch-forensics-instance to perform investigation.
- 4) Then attaché the evidence cloudresearch-forensics-instance by follow:

From volume actions we selected Attach Volume and choose to attach it to cloudresearch-forensics-instance as shown in Fig. 11 [62].

We confirmed the successful attachment from the console as shown in Fig. 12.



Fig. 11. Attaching the Evidence to the SIFT Workstation.



Fig. 12. Successful Attachment.

We verified that using lsblk command from cloudresearch-forensics-instance

```
$ sudo lsblk
NAME MAJ:MIN RM SIZE RO TYPE
MOUNTPOINT
Xvda 202:0 0 8G 0 disk
└─Xvda1 202:1 0 8G 0 part /
Xvdf 202:80 0 8G 0 disk
└─Xvdf1 202:81 0 8G 0 part
```

The *xvdf* disk is attached volume with a single partition not mounted yet.

We determined the format of partition using *file* command

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo file -s /dev/xvdf1
/dev/xvdf1: SGI XFS filesystem data (blksz 4096, inosz 512, v2 dirs)
```

As we are analyzing a Linux machine, we can see that the format is XFS.

Mount the evidence Linux file system.

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo mount -o ro /dev/xvdf1 /mnt/evidence /
```

Verify a successful mount of the evidence volume using *ls* command.

```
$ sudo ls -als /mnt/evidence/
total 20
dr-xr-xr-x 18 root root 257 Apr 14 17:37 .
drwxr-xr-x 18 root root 4096 May 4 05:45 ..
-rw-r--r-- 1 root root 0 Apr 14 17:37 .autorelabel
lrwxrwxrwx 1 root root 7 Mar 26 17:35 bin ->
usr/bin
dr-xr-xr-x 4 root root 4096 Mar 26 17:36 boot
drwxr-xr-x 3 root root 136 Mar 26 17:36 dev
drwxr-xr-x 81 root root 8192 Apr 17 04:05 etc
drwxr-xr-x 3 root root 22 Apr 14 17:37 home
lrwxrwxrwx 1 root root 7 Mar 26 17:35 lib ->
usr/lib
lrwxrwxrwx 1 root root 9 Mar 26 17:35 lib64 ->
usr/lib64
drwxr-xr-x 2 root root 6 Mar 26 17:35 local
drwxr-xr-x 2 root root 6 Apr 9 2019 media
drwxr-xr-x 2 root root 6 Apr 9 2019 mnt
drwxr-xr-x 4 root root 27 Mar 26 17:36 opt
drwxr-xr-x 2 root root 6 Mar 26 17:35 proc
dr-xr-x--- 3 root root 103 Apr 14 17:37 root
drwxr-xr-x 3 root root 18 Mar 26 17:36 run
```

```
lrwxrwxrwx 1 root root 8 Mar 26 17:35 sbin ->
usr/sbin
drwxr-xr-x 2 root root 6 Apr 9 2019 srv
drwxr-xr-x 2 root root 6 Mar 26 17:35 sys
drwxrwxrwx 8 root root 172 May 4 03:54 tmp
drwxr-xr-x 13 root root 155 Mar 26 17:35 usr
drwxr-xr-x 19 root root 269 Apr 14 17:37 var
Ubuntu@cloudresearch-forensics-instance:~
```

5) *Forensic analysis on linux EC2 instance:* Fig. 13 contains perform common disk image forensic investigation domain exercises on the snapshot of EC2 instance's volume and define any limitations and special configurations needed to facilitate the investigation [63].

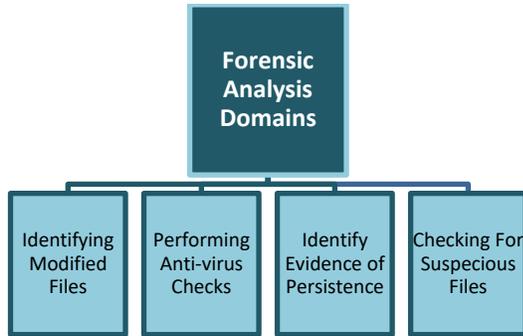


Fig. 13. Forensic Analysis Domains Test.

d) Identifying modified and added files in the invested file system:

1) Create a new instance from the same type of instance to perform forensics analysis on. This new instance will be used as a reference to a known good state. Achieve that by taking a snapshot of the newly created instance, mounting it to the cloudresearch-forensics-instance in read-only then creating hashes of all files on the reference volume to help identify the differences between the Evidence and the Baseline.

```
Ubuntu@cloudresearch-forensics-instance:~
$ lsblk
NAME        MAJ:MIN RM  SIZE RO  TYPE MOUNTPOINT
Loop0       7:0      0  55.5M 1  loop /snap/core18/1988
Loop1       7:1      0  33.3M 1  loop /snap/amazon-ssm-
Loop2       7:2      0  55.5M 1  loop /snap/core18/1997
Loop3       7:3      0  70.4M 1  loop /snap/lxd/19647
Loop4       7:4      0  32.3M 1  loop /snap/snapd/11588
Loop5       7:5      0  31.1M 1  loop /snap/snapd/11036
Loop6       7:6      0  69.9M 1  loop /snap/lxd/19188
Xvda        202:0    0    8G    0    disk
├─ Xvda1    202:1    0    8G    0    part /
├─ xvdf    202:80   0    8G    0    disk
├─ xvdf1   202:81   0    8G    0    part /mnt/evidence
├─ xvdg    202:96   0    8G    0    disk
└─ xvdg1   202:97   0    8G    0    part /mnt/baseline
```

2) creating a hash database of all Baseline Volume files as follows:

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo find /mnt/baseline/ -type f -exec /usr/bin/md5sum {} \; >
baseline_files.md5
```

3) Then making a hash of all files on the volume under investigation. And compare them with the reference file hashes, storing them in changed_files.txt

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo find /mnt/evidence/ -type f -exec /usr/bin/md5sum {} \; >
investigate_files.md5
```

```
Ubuntu@cloudresearch-forensics-instance:~
$ ls -al *.md5
-rw-rw-r-- 1 ubuntu ubuntu 4304670 May 6 15:18
baseline_files.md5
-rw-rw-r-- 1 ubuntu ubuntu 3865228 May 6 15:15
investigate_files.md5
```

4) After making a hash index of the both files using *hfind* command, the 2 hash tables to find the difference. then make the output in a text file.

```
Ubuntu@cloudresearch-forensics-instance:~
$ awk '{print $1}' investigate_files.md5 | hfind baseline_files.md5 |
grep "Hash Not Found" | awk '{print $1}' > difference.md5
```

```
Ubuntu@cloudresearch-forensics-instance:~
$ hfind -f difference.md5 investigate_files.md5 > different_files.txt
```

Fig. 14 shows a sample of Files that has been changed.

```
Ubuntu@cloudresearch-forensics-instance:~
$ head different_files.txt
0064706204FE32E475F0C354946F4 /mnt/evidence/etc/openssh/ssh-agent-script/ssh-agent-functions
7E457E70481911111111111111111111 /mnt/evidence/etc/openssh/ssh-agent
00F978C0541894F611FF64077173 /mnt/evidence/etc/statd
00E6A271C94F77041125F1170A890C /mnt/evidence/etc/ntp/cpe-ntp/ntp-32-3201-7
0F03C5C840D7F01127C4941531161 /mnt/evidence/etc/ntp/cpe-trust/extracted/ntp/cacerts
03082E3448F78EE13F01101219F90C /mnt/evidence/etc/yum/repos/df/ana2-ectras.repo
467740007F6142E21844111111111111 /mnt/evidence/etc/yum/repos/df/legal-test/ing.repo
E05648847E241C9629414E70A0F7E /mnt/evidence/etc/yum/repos/df/legal.repo
08342C35E74517711421111111111111 /mnt/evidence/etc/yum/repos/amazon1.repo
007547000245702138A8F510358E0 /mnt/evidence/etc/yum/repos/awsgoto
```

Fig. 14. Sample of Output.

5) *Finding keys on the compromised system:* Private keys are usually found in hidden directories and the SSH by analysts. Looking for private keys in SSH and AWS hidden directories (.ssh/ and .aws/). In an EC2 case, unprotected private keys are a bad security practice that could be violating the company's security policy. If any private SSH keys or AWS keys are present, they must be supposed to be compromised.

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo ls /mnt/evidence/home/ec2-user/.ssh/authorized_keys
```

Sample of the command searching for the pattern AKIA[A-Z0-9] shown in Fig. 15.



Fig. 15. Sample of Output from Command AKIA [A-Z0-9] Pattern.

Another common practice is to look at the list of public keys used to SSH into the instance (home/*/.ssh/authorized keys), which can be used in conjunction with syslog messages to figure out who has accessed the system as shown in Fig. 16.



Fig. 16. List of Public Keys those are used to SSH to the Instance.

e) Performing Anti-Virus checks

1) Find any security software installed: Most cloud providers provide a mechanism for central management of VMs' systems deployed on the IaaS. These systems are used to patch, configure and audit the VMs to a baseline. Then information available through these systems may support the forensics investigation. For AWS this system is called AWS System Manager. If the AWS system Manager is present its default location will be /usr/bin/amazon-ssm-agent and its log path will be /var/log/amazon/ssm/amazon-ssm-agent.log. Another tool that can be offered by the Cloud provider to the deployed VMs is vulnerability scanner. AWS offers AWS Inspector. If this tool was deployed on the compromised machine indicates that there may be vulnerability data available via the console which can help focus the investigation.

2) Scan the image with antivirus: To validate if malware files can be detected by using AntiVirus Scan tools, download EICAR file to an investigated instance before taking a snapshot of its volume. EICAR is ANTI MALWARE Test file. The scanning of a snapshot volume has no different than

scanning conventional forensics images and able to detect the EICAR file. For this purpose, use ClamAV that comes with the SIFT package. Then scan the mounted evidence with ClamAV.

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo clamscan -i -r --log=/cases/clam.log /mnt/evidence/
```

The previous command results are shown in Fig. 17.

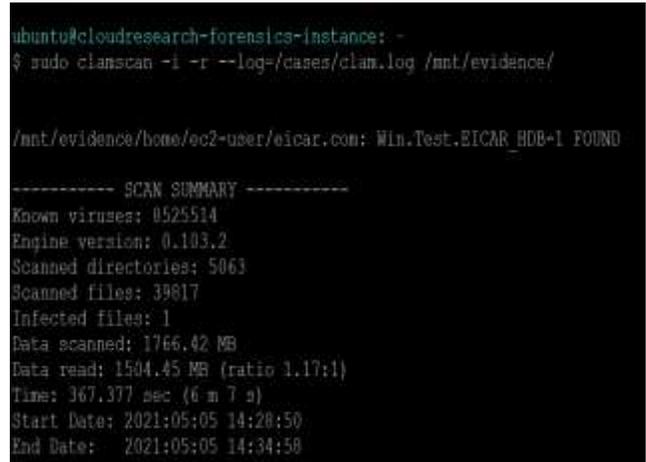


Fig. 17. Result of Scan the Image with Antivirus.

f) Identify Evidence of Persistence

The malware must use a persistence mechanism to survive a reboot. The two most popular methods are start-up scripts and cron jobs.

1) View the cron task by viewing crontab file and listing cron files as shown in Fig. 18:

```
Ubuntu@cloudresearch-forensics-instance:~
cat /mnt/evidence/etc/crontab
```

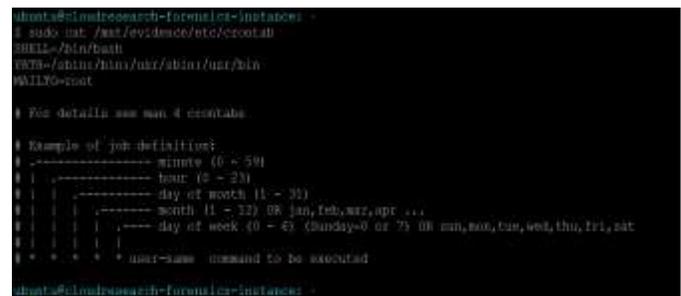


Fig. 18. Snapshot of Corntab File and Corn Files.

Also check the corn jobs for all users.

```
Ubuntu@cloudresearch-forensics-instance:~
$ sudo ls /mnt/evidence/etc/cron.*
```

2) Looking for unusual start-up scripts: While entering a specific run stage, some malwares can use the start-up scripts that Linux runs at boot time. These scripts can be found in

/etc/init.d on some Linux distributions, but they will be in /etc/rc*.d on Amazon Linux as shown in Fig. 19:

```

ubuntu@cloudresearch-forensics-instance: ~
$ sudo ls -al /mnt/evidence/etc/rc*.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc0.d -> rc.d/rc0.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc1.d -> rc.d/rc1.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc2.d -> rc.d/rc2.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc3.d -> rc.d/rc3.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc4.d -> rc.d/rc4.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc5.d -> rc.d/rc5.d
lrwxrwxrwx 1 root root 10 Mar 26 17:35 /mnt/evidence/etc/rc6.d -> rc.d/rc6.d

/mnt/evidence/etc/rc.d:
total 16
drwxr-xr-x 10 root root 127 Mar 26 17:36 .
drwxr-xr-x 81 root root 5192 Apr 17 04:05 ..
drwxr-xr-x 2 root root 85 Mar 26 17:36 init.d
-rw-r--r-- 1 root root 471 Feb 3 2020 rc.local
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc0.d
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc1.d
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc2.d
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc3.d
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc4.d
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc5.d
drwxr-xr-x 2 root root 45 Mar 26 17:36 rc6.d
ubuntu@cloudresearch-forensics-instance: ~
$ sudo ls -al /mnt/evidence/etc/init.d
lrwxrwxrwx 1 root root 11 Mar 26 17:35 /mnt/evidence/etc/init.d -> rc.d/init.d
ubuntu@cloudresearch-forensics-instance: ~

```

Fig. 19. Snapshot of Start-up Scripts on Amazon Linux.

g) *Checking for Suspicious Files:* Perform all techniques related to looking and searching for suspicious files as perform them in the conventional forensic analysis. As follows:

- Look for the contents of the /tmp directory and unusual SUID files. To search for unusual SUID files, make a base line using baseline volume and compared it with the volume under investigation. Purposefully add a malicious _file that has suid permission enabled to the home directory of the ec2-user account on the compromised machine and successfully to detect it as shown in Fig. 20:

```

ubuntu@cloudresearch-forensics-instance: ~
$ sudo find /mnt/evidence/ -perm -4000 -print > suid_evidence
ubuntu@cloudresearch-forensics-instance: ~
$ sudo find /mnt/baseline/ -perm -4000 -print > suid_baseline
ubuntu@cloudresearch-forensics-instance: ~
$ cut suid_baseline -d"/" -f4 -> suid_baseline_relative
ubuntu@cloudresearch-forensics-instance: ~
$ cut suid_evidence -d"/" -f4 -> suid_evidence_relative
ubuntu@cloudresearch-forensics-instance: ~
$ diff suid_baseline_relative suid_evidence_relative
2d1
< usr/bin/mount
fd2
< usr/bin/mount
9a8,9
> usr/bin/mount
> usr/bin/mount
17a18
> home/ec2-user/malicious_file
ubuntu@cloudresearch-forensics-instance: ~

```

Fig. 20. Detect Malicious _file in Home Directory of the ec2-user Account.

- Use the same method to look for large files above a certain size limit and compare them to the base volume.
- Use the *strings* command to get a fast indication of the file's existence and to spot possible signs of compromise. The strings command's output can include IP addresses, file names, and configuration information that expose the malware's intent.

```

Ubuntu@cloudresearch-forensics-instance:~
$ string /mnt/evidence/home/ec2-user/eicar.com

X50!P% @AP[4\PZX45(P^)7CC)7]$EICAR-STANDARD-
ANTIVIRUS-TEST-FILE!$H+H*

```

- Use 3rd party tools like **VirusTotal** to investigate the suspected files detected by this exercise and validate if the samples have been seen in the wild as shown in Fig. 21.

For this purpose, calculate the SHA256 hash of detected malware (EICAR) and then look for it on VirusTotal.

```

Ubuntu@cloudresearch-forensics-instance:~
$ sudo sha256sum /mnt/evidence/home/ec2-user/eicar.com

275a021bbfb6489e54d471899f7db9d1663fc695ec2fe2a2c4538aabf6
51fd0f /mnt/evidence/home/ec2-user/eicar.com

```

C. Discussion and Area of Improvements

A common mistake is to let snapshots be shared publicly (across different AWS accounts), without great care is taken, this can lead to leakage of sensitive data and keys.

According to Amazon, the only way to capture the memory of an Amazon Linux EC2 instance is via an SSH session, passing the SSH keys that have the root privilege to a remote memory imaging tool.

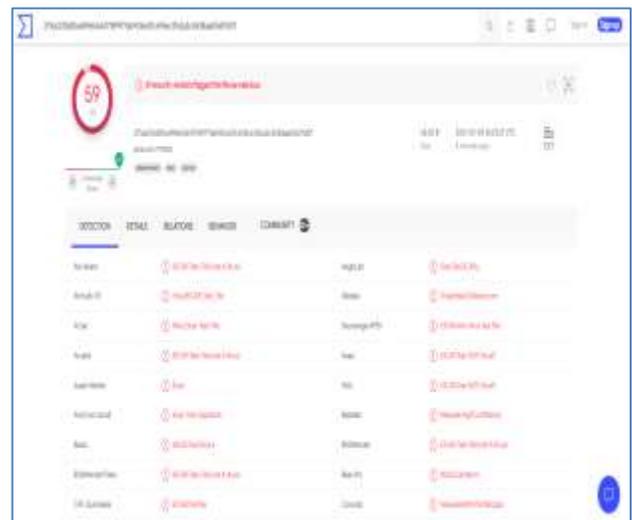


Fig. 21. Virustotal Result for the Malware we Detected on Machine.

As a best practice incident handler should provide a new forensics analysis machine for each case. Provisioning the forensics analysis machine on AWS takes some time, unlike the conventional on-prem machines. But one way to speed this up is to make an AMI (Amazon Machine Image) of the SIFT.

The volumes created from snapshots must be in the same availability zone of the forensics analysis workstation.

There is a quiet difference between conventional on-premises evidence gathering and evidence gathering from IaaS virtual machines and due to the need for rapid containment of the compromised cloud virtual machine, these evidence gathering processes must be integrated with infrastructure automated deployments tools for automating the evidence gathering.

Memory analysis for IaaS virtual machines is very challenging as most leading solutions for image analysis require a profile that is specific to the kernel of the system that was imaged, so they have profiles for operating systems used on-premises but don't have any profiles for the Cloud provided Linux images. Custom profiles can only be made for these tools if only the analyst has the source code for the kernel headers of these cloud-provided Linux images.

Amazon Linux differs from traditional on-prem Linux especially in the path for some services and files such as the path for the startup scripts and the path of all cron jobs, thus creating an inconsistency between different cloud providers Linux distributions and the forensic analyst have to learn the customizations before beginning his investigation.

The audit logs generated by the Amazon Linux process which is by default enabled will take some practice to read and learn how to read them.

For instances running web servers, the logs of the web service exist in unusual places on the volume, thus the best way to analyze these logs is to pull them into a log server or SIEM solution.

Using cloud web load balancers such as Amazon's Elastic Load Balancer may hide the address of the attacker if no special customization to the load balancer configuration.

EC2 instances are set to UTC by default, thus if anyone trying to make a timeline for malware attacks that spread from the cloud-hosted infrastructure to the on-premises infrastructure, to make a TimeZone correction before correlating cloud instances logs with on-premises logs.

Cloud instances use a default setting for Syslog that sometimes is inconsistent with the common or Syslog settings used on-premises. For example, Amazon Linux images in the marketplace use a default setting for Syslog that does not include the year in the date stamp. Ideally, a cloud administrator uses a configuration template for configuring new instances on the cloud which properly configures the Syslog timestamp to include the year, but this may not be the case with each organization or cloud administrator.

The challenge for forensics in the cloud is usually "Will it meet the chain of custody requirements?" and "Will it hold up in court?" You need to enable write-once storage that is

owned solely by the forensics and IR teams and carefully document your IAM policy for the IR/Forensics role. In addition, ALL activity around evidence acquisition and evidence storage location should be logged extensively.

Despite standardized, well-defined, and matured tools of Incident Response and evidence acquisition for suspected on-premises assets, there are no standardized processes for Digital forensics automation on the cloud and very few tools have been built to do this for the cloud today. Almost nothing is available commercially and each cloud service provider tries to provide its own solution for this purpose for example Microsoft has centered most of its security capabilities around Security Center and Sentinel within Azure.

IV. CONCLUSION

Recently, the number, severity, sophistication of malware attacks, and cost of malware infect on the world economy have been increasing exponentially. Malware should be detected before damaging the important assets in the company. In this research, we applied cybersecurity and security in-depth principles to the cloud IaaS environment. These principles embrace that defense controls will fail at some point and an attack will succeed so organizations must have response mechanisms to put off these attacks as soon as possible, to be able to detect and monitor the attack while and post its occurrence is very crucial. Log monitoring and digital artifacts gathering are the cornerstone enablers for monitoring and detection of active malware attacks. In this research, we defined a practical baseline of the security telemetry that needs to be configured on cloud IaaS environments to maintain continuous visibility and monitoring. We achieved the baseline of the practical part through using of MITRE ATT&CK framework for Cloud which classifies cloud attacks and cloud attack vectors that have been abused so far. After that, we validated the applicability and limitation of deploying this baseline using the AWS environment. Collecting logs without performing investigation and analysis of these logs will provide no help regarding the attack detection. This can be achieved by using security event management tools such as SIEM solutions that perform data correlation, enrichment, integration with other security events, and long-term storage. In the second part of the research, review the different ways to transfer logging data from IaaS environment to the SIEM solution located on-premises. The third part of this research investigated the applicability of performing Digital Forensic Investigations on the compromised IaaS VMs. In future work, we suggest that cloud providers should provide the maintenance tools for performing volatile memory analysis for their VMs. Also, develop a new automated tool for incident response and forensics investigation on the IaaS.

REFERENCES

- [1] B. Varghese and R. Buyya, "Next Generation Cloud Computing: New Trends and Research Directions," *Elsevier :Future Generation Computer Systems*, Vol. 79, pp. 1-22, September 2017.
- [2] W. Yassin, M. F. Abdollah, R. Ahmad, Z. Yunos and A. Ariffin, "Cloud Forensic Challenges and Recommendations: A Review," *OIC-CERT Journal of Cyber Security*, vol.2, issue.1, pp. 19 – 29, February 2020.
- [3] W. Dawoud , I. Takouna and C. Meinel, "Infrastructure as a Service Security: Challenges and Solutions," *Conference: Informatics and*

- Systems (INFOS), 2010 The 7th International Conference, IEEE Explore.*, April 2010.
- [4] R. V. Deshmukh and K. K. Devadkar, "Understanding DDoS Attack & Its Effect In Cloud Environment," *Procedia Computer Science, Elsevier*, Vol.49, pp. 202 – 210, 2015.
- [5] A. Bedi, N. Pandey and S. K. Khatri, "Analysis of Detection and Prevention of Malware in Cloud Computing Environment," *2019 Amity International Conference on Artificial Intelligence (AICAI), IEEE*, pp. 918-921, 2019.
- [6] N. Babu and G. Murali, "Malware Detection for Multi Cloud Servers using Intermediate Monitoring Server," *International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), IEEE*, pp. 3609-3612, Aug 2017.
- [7] N. Rakotondravony, B. Taubmann, W. Mandarawi, E. Weishäupl, P. Xu, B. Kolosnjaji, M. Protsenko, H. de Meer and H. P. Reiser, "Classifying malware attacks in IaaS cloud environments," *Journal of Cloud Computing: Advances, Systems and Applications*, Vol.6, no.26, December 2017.
- [8] H. Rathore, S. Agarwal, S. K. Sahay and M. Sewak, "Malware Detection using Machine Learning and Deep Learning," *International Conference on Big Data Analytics*, Springer; LNCS, Vol. 11297, pp. 402-411, 4 Apr 2019.
- [9] A. Giannakou, L. Rilling, C. Morin and J.-L. Pazat, "SAIDS: A Self-Adaptable Intrusion Detection System for IaaS Clouds," *IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 354-355, May 2018.
- [10] A. Alenezi, R. K. Hussein, R. J. Walt and G. B. Wills, "A Framework for Cloud Forensic Readiness in Organizations," *IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*, IEEE Xplore, 12 June 2017.
- [11] I. Ahmad and H. Bakht, "A Novel Detection and Prevention (DAP) Framework for Abuse of cloud service threat," *International Journal of Computing and Network Technology*, Vol.6, Issue.3, no.3, Sept 2018.
- [12] S. Fatima and S. Ahmad, "An Exhaustive Review on Security Issues in Cloud Computing," *Ksii Transactions On Internet And Information Systems*, Vol. 13, Issue. 6, pp. 3219-3237, Jun 2019.
- [13] K. Lokuge, "Security Concerns in Cloud Computing: A Review," *Secure Software Development Assignment, Researchgate*, December 2020.
- [14] B. Mohammed, B. Modu, K. M. Maiyama, H. Ugail and I. Awan, "Failure Analysis Modelling in an Infrastructure as a Service (IaaS) Environment," *Electronic Notes in Theoretical Computer Science*, Vol.340, p. 41–54, October 2018.
- [15] E. B. Chawkia, A. Ahmeda and T. Zakariae, "IaaS Cloud Model Security Issues on Behalf Cloud Provider and User Security Behaviors," *Procedia Computer Science*, Vol.134, pp.328–333, January 2018.
- [16] Nancy, S. Silakari and U. Chourasia, "A Survey Over the Various Malware Detection Techniques used in Cloud Computing," *International Journal of Engineering Research & Technology (IJERT)*, Vol. 5 Issue.02, pp. 398-402, February 2016.
- [17] Y. YE, T. LI, D. ADJEROH and S. S. IYENGAR, "A Survey on Malware Detection Using Data Mining Techniques," *ACM Computing Surveys*, Vol. 50, No. 3, pp. 1-41, June 2017.
- [18] X. Gao, C. Hu, C. Shan, B. Liu, Z. Niu and H. Xie, "Malware classification for the cloud via semi-supervised transfer learning," *Journal of Information Security and Applications, Elsevier*, Vol.55, 20 October 2020.
- [19] X. Zhu, J. Wang, H. Guo, D. Zhu, L. T. Yang and L. Liu, "Fault tolerant scheduling for real-time scientific workloads with elastic resource provisioning in virtualized clouds," *IEEE Transactions on Parallel and Distributed Systems*, Vol.27, Issue.12 pp. 3501 - 3517, , Dec. 2016.
- [20] A. K. Mamerides, M. R. Watson, N. Shirazi, A. Mauthe and D. Hutchison, "Malware analysis in Cloud Computing: Network and system characteristics," *IEEE Globecom Workshops (GC Wkshps)*, December 2013.
- [21] M. R. Watson, N.-u.-h. Shirazi, A. K. Mamerides, A. Mauthe and D. Hutchison, "Malware Detection in Cloud Computing Infrastructures," *Transactions on Dependable and Secure Computing, IEEE*, Vol.13, Issue.2, pp.192-205, January 2015.
- [22] M. Abdelsalam, R. Krishnan and R. Sandhu, "Clustering-Based IaaS Cloud Monitoring," *2017 IEEE 10th International Conference on Cloud Computing (CLOUD)*, June 2017.
- [23] Ö. ASLAN and R. SAMET, "A Comprehensive Review on Malware Detection Approaches," *IEEE Access*, Vol.8, pp. 6249 - 6271, 03 January 2020.
- [24] N. Praveena, S. Sofia and D. Srinivasulu, "Anomaly Detection in Infrastructure Service of Cloud Computing," *International Journal of Computer Science Trends and Technology (IJCTST)*, Vol.4, Issue.6, pp. 104-108, Nov - Dec 2016.
- [25] Y. Ye, T. Li, S. Zhu, W. Zhuang, E. Tas, U. Gupta and M. Abdulhayoglu, "Combining file content and file relations for cloud based malware detection," *the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 222–230, August 2011.
- [26] P. Srivastava and M. Raj, "Feature extraction for enhanced malware detection using genetic algorithm," *International Journal of Engineering & Technology*, Vol.7, pp. 444-449, March 2018.
- [27] M. Jain and P. Bajaj, "Techniques in Detection and Analyzing Malware Executables: A Review," *International Journal of Computer Science and Mobile Computing*, Vol. 3, Issue. 5, pp. 930 – 935, May 2014.
- [28] R. Sihwail, K. Omar and K. A. Ariffin, "A Survey on Malware Analysis Techniques: Static, Dynamic, Hybrid and Memory Analysis," *International Journal on Advanced Science Engineering and Information Technology*, Vol.8, No.4-2, pp. 1662-1671, September 2018.
- [29] A. Damodaran, F. D. Troia, C. A. Visaggio, T. H. Austin and M. Stamp, "A comparison of static, dynamic, and hybrid analysis for malware detection," *Journal of Computer Virology and Hacking Techniques*, Vol.13, pp.1-12, 29 December 2015.
- [30] J. D. Araújo and Z. Abdelouahab, "Virtualization in Intrusion Detection Systems: A Study on Different Approaches for Cloud Computing Environments," *IJCSNS International Journal of Computer Science and Network Security*, Vol.12 No.11, pp. 9-16, November 2012.
- [31] Z. Bazrafshan, H. Hashemi, S. M. Hazrati Fard and A. Hamzeh, "A survey on heuristic malware detection techniques," *5th Conference on Information and Knowledge Technology (IKT), IEEE*, pp. 113-120, May 2013.
- [32] M. H. M. Yusof and M. R. Mokhtar, "A Review of Predictive Analytic Applications of Bayesian Network," *International Journal on Advanced Science Engineering and Information Technology*, Vol.6, No.6, pp. 857-867, August 2016.
- [33] H. Sun, X. Wang, R. Buyya and J. Su, "CloudEyes: cloud-based malware detection with reversible sketch for resource-constrained internet of things (IoT) devices," *Software Practice and Experience*, Vol. 47, June 2016.
- [34] S. Shen, L. Huang, H. Zhou, S. Yu, E. Fan and Q. Cao, "Multistage Signaling Game-Based Optimal Detection Strategies for Suppressing Malware Diffusion in Fog-Cloud-Based IoT Networks," *IEEE Internet of Things Journal*, Vol.5, Issue.2, pp. 1043 - 1054, April 2018.
- [35] H. Hamad, M. Hoby, "Managing Intrusion Detection as a Service in Cloud Networks," *International Journal of Computer Applications*, Vol.41, No.1 pp. 35-40, March 2012.
- [36] D. Singh, D. Patel, B. Borisaniya and C. Modi, "Collaborative IDS Framework for Cloud," *International Journal of Network Security*, Vol.18, No.4, pp. 699-709, July 2016.
- [37] C. N. Modi and D. Patel, "A novel hybrid-network intrusion detection system (H-NIDS) in Cloud Computing," *2013 IEEE Symposium on Computational Intelligence in Cyber Security (CICS)*, April 2013.
- [38] K. Vieira, A. Schultze, C. B. Westphall and C. Merkle, "Intrusion Detection for Grid and Cloud Computing," *IEEE IT Professional*, Vol.12, Issue.4, pp. 38 - 43, July-Aug. 2010.
- [39] J.-H. Lee, M.-W. Park, J.-H. Eom and T.-M. Chung, "Multi-level Intrusion Detection System and log management in Cloud Computing," *13th International Conference on Advanced Communication Technology (ICACT2011), IEEE*, Feb 2011.
- [40] J. Arshad, P. Townend and J. xu, "An Abstract Model for Integrated Intrusion Detection and Severity Analysis for Clouds," *International*

- Journal of Cloud Applications and Computing*, Vol.1, pp. 1-16, January 2011.
- [41] R. Kumar and R. Goyal, "On cloud security requirements, threats, vulnerabilities and countermeasures: A survey," *Elsevier journal-Computer Science Review*, Vol.33, pp. 1–48, 2019.
- [42] Y. Mehmood, U. Habiba, M. A. Shibli and R. Masood, "Intrusion Detection System in Cloud Computing: Challenges and opportunities," *IEEE 2nd National Conference on Information Assurance (NCIA)*, December 2013.
- [43] C.-C. Lo, C.-C. Huang and J. Ku, "A Cooperative Intrusion Detection System Framework for Cloud Computing Networks," *IEEE 39th International Conference on Parallel Processing Workshops*, Sept 2010.
- [44] A. V. Dastjerdi, K. Abu Bakar and S. G. H. Tabatabaei, "Distributed Intrusion Detection in Clouds Using Mobile Agents," *Third International Conference on Advanced Engineering Computing and Applications in Sciences*, Oct 2009.
- [45] J. Nikolai and Y. Wang, "Hypervisor-based cloud intrusion detection system," *International Conference on Computing, Networking and Communications (ICNC)*, IEEE, February 2014.
- [46] N. Pandeewari and G. Kumar, "Anomaly Detection System in Cloud Environment Using Fuzzy Clustering Based ANN," *Mobile Networks and Applications*, Vol.21, pp. 494–505, 16 August 2015.
- [47] A. Aldribi, I. Traoré, B. Moa and O. Nwamuo, "Hypervisor-based cloud intrusion detection through online multivariate statistical change tracking," *Elsevier: Computers & Security*, Vol.88, January 2020.
- [48] H. S. Anderson, A. Kharkar and B. Filar, "Evading Machine Learning Malware Detection," *Black Hat USA*, pp. 22-27, July 2017.
- [49] H. El Merabet and A. Hajraoui, "A Survey of Malware Detection Techniques based on Machine Learning," (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, Vol. 10, No. 1, pp. 366-373, 2019.
- [50] P. HarshaLatha and R. Mohanasundaram, "Classification Of Malware Detection Using Machine Learning Algorithms: A Survey," *International Journal Of Scientific & Technology Research* , Vol. 9, Issue. 02, pp. 1796-1802, February 2020.
- [51] B. Cakir and E. Dogdu, "Malware Classification Using Deep Learning Methods," *ACMSE '18: Proceedings of the ACMSE 2018 Conference*, March 2018 .
- [52] S. Fathima, S. and A. K. , "Comparative Analysis of Malware Classification using Machine Learning Algorithms," *International Journal of Engineering Research and Applications*, Vol. 10, Issue 12, pp. 64-68, December 2020.
- [53] E. Raff, J. Barker, J. Sylvester and R. Brandon, "Malware Detection by Eating a Whole EXE," *The Workshops of the Thirty-Second AAAI Conference on Artificial Intelligence*, pp. 268-276, October 2017.
- [54] L. W. B. S. Y. B. & Z. Q. X. Liu, "Automatic malware classification and new malware detection using machine learning," *Frontiers of Information Technology & Electronic Engineering*, Vol.18, pp. 1336-1347, 27 October 2017.
- [55] W. Jansen and T. Grance, "Guidelines on Security and Privacy in Public Cloud Computing," *NIST Special Publication:800-144*, December 2011.
- [56] B. Balakrishnan and R. R. Varuni, "Cloud Security Monitoring," *SANS Institute*, 8 Mar 2017.
- [57] A. Amazon Web Services, *Amazon CloudWatch Developer Guide*, 2010.
- [58] B. Beach, S. Armentrout , R. Bozo and E. Tsouris, "Virtual Private Cloud," in *Pro PowerShell for Amazon Web Services*, USA, Apress, Berkeley, CA, pp. 85-115, 22 September 2019.
- [59] J. Dykstra , "Digital forensics for infrastructure-as-a-service cloud computing," Ph.D dissertation, *Faculty of the Graduate School of the University of Maryland, Baltimore County*, 2013.
- [60] A. Pichan, M. Lazarescu and S. T. Soh, "Cloud forensics: Technical challenges, solutions and comparative analysis," *Digital Investigation*, Elsevier, Vol.13, pp. 38-57, 23 March 2015.
- [61] A. Anand, "Managing Infrastructure in Amazon using EC2,CloudWatch, EBS, IAM and CloudFront," *International Journal of Engineering Research & Technology (IJERT)*, Vol.6, Issue.03, pp. 373-378, March 2017.
- [62] Д. Ранделович, К. Кук, В. Боровик, Д. Младенович and Д. Эрлевайн, "Influence Of The Operating System On The Forensics Tools," *Thematic Conference proceedings Archibald Reiss*, Vol.343, 2016.
- [63] B. Carrier, *File system forensic analysis*, US: Pearson Education, Inc., 2005.

Reputation Measurement based on a Hybrid Sentiment Analysis Approach for Saudi Telecom Companies

Bayan Abdullah¹, Nouf Alosaimi², Sultan Almotiri³
College of Computer and Information Systems
Umm Al-Qura Univeristy
Makkah, Saudi Arabia

Abstract—Thousands of active people on social media daily share their thoughts and opinions about different subjects and different issues. Many social media platforms used to express the feeling or opinion and at top of it is Twitter. On Twitter, many opinions are expressed in many fields such as movies, events, products, and services; this data considered a valuable resource for companies and decision-makers to help in making decisions. This study was based on using a hybrid approach to extract the opinions from an Arabic tweet to measuring service providers' reputation. In this study, the Saudi telecom companies used as a case study. This research concentrates on determining peoples' opinions more accurately by utilizing the Retweet and Favorite. The number resulting from positive and negative tweets after applying the polarity equation was used to estimate reputation scores. The result indicated that the STC company represents a high reputation compared to other companies. The proposed approach shows promising results to expand existing knowledge of sentiment analysis in the domain of measure reputation.

Keywords—Reputation; sentiment analysis; Arabic language; social media

I. INTRODUCTION

Data analysis and mining are some of the most important areas studied in computer science [1]. Due to the importance of data in various fields and the amount of data that is generated during the day, many companies have paid great attention to data. The amount of data is increasing day by day due to a large number of social media platforms [2]. A large amount of content on social media is processed periodically. This data can provide information about users' loyalty, complaints, potential customers, or measure the reputation. To deliver several services and better understanding the customers companies rely on different social media platforms such as Facebook and Twitter [3].

Twitter is one of the popular platforms where people express their thoughts and opinions in 140 characters. The number of active users in this micro-blog is more than 328 million users, who vary between presidents, academics, organizations, and bodies [4]. Statistics indicate that the number of Twitter users in Saudi Arabia has reached 41%, which is a large percentage compared to other countries [5]. In addition, the Arabic language is considered one of the rich languages that have broad audiences, so the focus was on it and Saudi society.

With the millions of opinions that are disclosed on social media about some topic, it is out of the question that these opinions will be biased [6]. People tend to refer to people's comments and opinions about products before they buy them, the movies they will watch or the countries they are thinking of traveling to. The opinions expressed by online users are a kind of electronic word of mouth (eWOM) [7]. The opinions are greatly influencing the decisions a customer will make and its behavior [8]. Where these opinions are considered as more trustworthy and credible way because it's independent of marketers' selling efforts [9].

The spread of sentiment on the Internet has created a new field of text analysis. It attracted many studies in many fields such as education [10], stock market [11] and others. User-generated comments are contributing to the establishment of digitalized WOM forms, which in turn contributes to creating an online reputation. Reputation is a factor that influences the desirability of a product or service, so it needs investment and good management [12][13]. Reputation affects companies' capabilities as well as increases their chances of sustainability in the market [14][15].

The objective of this study is to measure reputation based on a hybrid sentiment analysis approach that combines textual and non-textual features. The analysis improved by including the number of retweets and favorite tweets. In the end, the reputation was calculated accurately for the service providers through the Beta reputation equation.

The rest of the paper is structured as: Section 2 investigates the related works of customer feedback analysis in different sectors. Then, Section 3 describes the methodology proposed which added a new step into the SA approach: the polarity equation, and reputation calculation by equation. Section 4 examines the supervised ML algorithm to determine which algorithm was the best and examines the reputation equation with two different experiments. Then describes the results after utilized the proposed methodology. Section 5 concludes the final notes of the work with recommended guidelines for future developments and research.

II. RELATED WORKS

Feedback from customers is important and it may be a complaint, recommendation, evaluation, etc. Different data mining techniques are used on the feedback to provide valuable

information. Analysis of customer feedback has been recently an active area. It covered several areas of the field of airline services, the smartphone industry, hotels, banking services, and many others.

The sources of feedback varied, some researchers resorted to using traditional methods such as the survey. The study of [16] was concerned about the customer satisfaction with the South Korean airline service with the two different customer groups the low-cost carrier and the full-service carrier. To investigate customer satisfaction, they related it to six different factors positive emotion, negative emotion, social words, comparison, risky values, and monetary values. They are analyzing more than 133000 of the customer feedback after they experience the airline service. The positive emotion and the social words were considered the most impactable factor on customer satisfaction while the monetary value has always a negative impact. The limitation of this study was they didn't pay attention to the demographic factor of the customer also they concern only with the English feedback.

Also, the work of [17] was based on the predictive analysis they investigated the customer intention to return visit the services provided by the airline services. By using the customer feedback comment and satisfaction rating for the previous use of the services they applied a machine learning approach. They used seven classifiers to analyze the sentiment in the comment. Their work was different from the previous work in that they lookup for the actual intention of the customer to return the visit. Also, they analyzed a huge data set what is contain with 309331 customer comment. All the long comment indicates higher accuracy. Their results indicated that customer feedback is very important to determine the actual intention for returned visits and also, it helps to understand the customer behavior over time.

Traditional methods are not recommended due to the sample that may participate in this survey. Also, the survey may be mandatory, which means the respondents may answering only for completion. Therefore, researchers have resorted to some alternative methods, which are based on social media or websites. These methods are more effective as they involve different samples and sectors.

Their work [18] was based on a sentiment analysis which they referred to as opinion mining. They come up with a new idea to analyzing customer feedback from Amazon. So, they designed a framework that received any kind of data from the user and preprocessing it, and then process the data. The result shows the amount of the sales product and the market share as a figure. Also, shows the count of the positive and negative reviews for each feature of the sold phones. Which is give the company a better picture of its product and helps it to succeed in the market.

Their work [19] was focused on the complaints of the customer they analyzed the data set from the Customer Relationship Management system (CRM). Which is a system that deals with the complaints of the customer for the Metropolitan Transportation Authority (MTA). The MTA provides several services that including a subway and buses and a railroad. From the correlation matrix they come up with the agency and subject matter attributes that are positively

correlated and there is no negative correlation. Also, they performed more analysis on details of the complaints to improve the quality of the services. For the classification, they applied different modeling: naive Bayes, KNN, random tree, and ID3. For each model, the input is the complement details that give the agency name as an output. Their work was helpful to determine the factor that led to the customer complaint and also helped the company to improve the quality of the provided services.

The comments that posted in different social media platform could be considered as customer feedback. Mini and Poulouse study [20] to discover the customer intention about the eco-tourism. They took a property in Kerala in India as a case study. They collected the data from Facebook and performed a sentiment analysis on it. They were able to determine the satisfaction and dissatisfaction of the customers which will help improve the quality and improvement process. In their study the focus only on collecting the comment from Facebook that was written in English. Also, the multiple emotion in the collected comments was not easy to extract it.

The study of [21] was based on user-generated content (UGC) to investigate the airport service quality. Their study was in London Heathrow airport services. Sentiment analysis was performing on the collected tweets. They were able to evaluate the quality of the services and determined they good services and the services that need to be improved. UGC is considered a better resource more than the traditional way. The limitation of this study they ignore the demographic characteristics and it's applied to one airport.

UGC on social media plays a significant role in measuring reputation. As the work is done by [22] they measure reputation for four companies Apple, Twitter, Google, and Microsoft by analyzing the tweet. Their proposed approach a distinct work on dealing with the misspelling in the tweet and can apply to all languages also, it does not need a text preprocessing stage. They sentimentally classify each tweet to one of the companies and determine the emotion in this tweet. By relying on using N-gram approaches their results to show better accuracy more than the Bayesian algorithm and a neural network.

The same data set was also analyzed in the study of [23]. Their study was focus on the user reputation by applying some factor to calculate to determine which user have a higher reputation than the others. Their proposed framework was first assigned a score for each tweet the positive tweet was indicated by 1 and the negative was indicated by -1. Then they extracted the ID of the tweet to retrieve the Twitter creator they extracted five factors for each creator: number of followers, number of following, Twitter lists, numbers of retweeted by the followers, and number of tweets posted by that user. Via these five factors, they calculated the reputation of the user. Then they calculate the accumulated weight by multiplying the sentiment score with the user reputation. The tweet from the higher reputation user will have a higher accumulated weight. The proposed method would be helpful to minimize the effort and the size of data by extracting the tweet that only relevant to the higher reputation users.

Online reputation may be useful in determining and measuring profits, according to a study of [24]. They emerged Booking review and Financial Analysis Made Easy (FAME) and used Latent semantic analysis on the textual content. They were able to determine which attributes are associated with financial profitability. From the positive review, they were able to determine that the hotel location and the room quality was indicated the highest impact attribute on the financial profitability. The study was conducted only on the UK hotels and they extract the review from the Booking platform only.

In the tourism sector, there have been several studies and research, such as the work that was done by [25]. They measured Marbella's reputation as a tourist destination. Their study focused on the factors that attract tourists and determine the geographical distribution of the countries from which they came. They analyzed visitor reviews according to sociodemographic characteristics of the city over three years. Their study was focusing only on TripAdvisor reviews.

Also, the study of [26] that took a different sector focused on the political sector. They examine the reputation of the members of the National Assembly (MNAs) of Pakistan on Facebook and Twitter. They were able to determine the credibility of the members and which platforms have a more supportive audience. The comment was chosen based on the systematic random sampling technique. The supportive audience was on Facebook more than Twitter.

Many kinds of research have been dedicated to the field of analyzing customer feedback. Regrettably, few studies consider the Arabic language.

The study of [27] was aimed at determining the polarity of the text only. Therefore, it dealt with several topics in politics, sport and education. They collected data from Twitter and Facebook and then applied three classifiers, SVM, Naïve Bayes, and K-NN using Rapid miner. The size of the data set was small, so no promising results were shown.

The study of [28] was also on the same area they concerned about the telecom Saudi company STC, Mobile, and Zain. Their study focused on the assignment of the polarity scores of the tweets by SWN. The determination of polarity in SWN was based on a lexicon as the results showed that most of the tweets tended to be natural. Defining polarity helped them better understand their customers and their needs. They were also able to identify areas of clients who had complained about some problems.

Previous studies have indicated determining the polarity of sentiment in the telecom industry, both of which were based on a lexicon-based approach. All previous studies focused on text analysis only, they did not include the number of retweets and favorites. The following Table I gives an overview of the previous studies, their purpose, source of data, and the approach used in their methodology.

TABLE I. OVERVIEW ON RELATED ARTICLES

Cite	Analysis Objective	Dataset Sources	Approach
[16]	Measuring customer satisfaction	Survey	Structural equation modeling SEM
[18]	Measuring how to succeed in the market	Amazon review	Framework to receive data and sentimentally analyse it
[19]	Understanding the Complement of customer	CRM	Data mining Technique and Machine learning Algorithm for classification
[17]	Determine if the customer will return to visit	Survey	Machine learning for sentiment analysis
[20]	Discover the customer intention about the eco-tourism	Collected tweet	Hybrid-based Approach for sentiment analysis
[21]	Investigate service quality	Collected tweet	Online tools
[22]	Measuring companies reputation	Collected tweet	N-gram
[23]	Measuring user reputation	Collected tweet	Framework to sentimentally analyse tweet and calculate reputation
[24]	Determine the impact of online reputation on hotel profitability	Booking reviews	Latent semantic Analysis
[25]	Measuring the reputation of Marbella city	Trip Advisor reviews	A quantitative analysis
[26]	Measuring leaders reputation	Collected tweet	Quantitative sentiment analysis
[28]	Gain better customer insight	Collected tweet	Sentiment analysis

III. PROPOSED METHODOLOGY

The methodology that followed in this study contains five different phases. Fig. 1 shows an overview of the proposed methodology. Later, in this section, every phase will explain in detail.

The polarity of the collected tweets was determined to positive, negative, or natural. For this study, only positive and negative tweets were included. Most of the sentiment analysis that was performed was focus on polarity only. For more precise a hybrid approach that combines textual and non-textual features was applied. The hybrid approach was based on the polarity equation [29]. The integer number from the polarity equation for both the positive and negative classes was the input for the reputation equation [30].

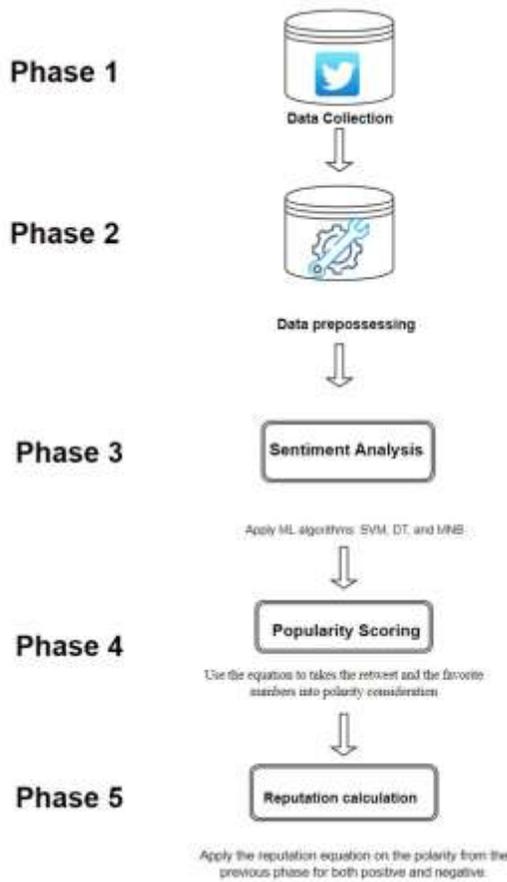


Fig. 1. Overview of the Proposed Methodology with the Five Phases.

A. Phase One: Data Collation

The data for this study was extracted through Twitter API. The analysis focused on the telecom service providers sector in Saudi Arabia. The providers were the data obtained from STC, Mobily, and Zine. The data was collected from the official account of the three providers: @stc_ksa, @Mobily, @ZainKSA. Each of these accounts is active accounts with thousands of tweets exchanged daily.

For each account, 5000 tweets were collected and the total amount of the data was approximately 15,000. Due to time constraints, data were collected within February only. The collected tweet was in the Arabic language mostly, while the tweets that contain some English words or letter were removed later as showed in Section III-B.

For each tweet, the time for its creation, tweet ID, text, the number of their retweet and favorite, user location, screen name, followers, and friends were extracted. For this study, only text, the number of their retweet, and favorite were used. Fig. 2 shows the statistical descriptive for the numerical values in the dataset.

B. Phase Two: Data Preprocessing

To prepared data that has been gathered for the sentiment analysis, it has to be cleaned and preprocessing. Preprocessing data means removing all the noise, unwanted parts, and missing values even the duplicate data. Performing all these procedures leads to get quality data and reliable results. The steps that followed on preprocessing data are:

1) Removed unwanted parts:

- Any English words, letter, or numbers.
- The retweet status, user mentions, URLs.
- Punctuation's, and the Tashkeel.

2) Normalized Arabic words: Arabic letter could be represented in many forms such (ا, إ, إ) all of it refer to (ا) letter. So, the three forms were replaced with (ا). Also, (ة) was replaced with (ه) and (ي) with (ي) .

3) Normalizing repeated letters if it was more than two letters.

4) Removed Arabic stop words: Same as any other language the Arabic language contains multilabel stop words. Stop words in Arabic are meaningless same articles (a, an, the) in English. The stop word such (لم عن الذي) removed based on two different corpora.

5) Removed line breaks and extra white spaces.

6) Removed Duplicates tweets.

	count	mean	std	min	25%	50%	75%	max
tweet_id	4656.0	1.360000e+18	0.000000	1.360000e+18	1.360000e+18	1.360000e+18	1.360000e+18	1.360000e+18
rt_count	4656.0	8.488543e+01	597.045859	0.000000e+00	0.000000e+00	0.000000e+00	1.400000e+01	2.822500e+04
fav_count	4656.0	6.432607e+01	22.626225	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	1.575000e+03
followers	4656.0	2.252855e+04	324997.451510	0.000000e+00	2.800000e+01	1.750000e+02	6.480000e+02	1.438033e+07
friends	4656.0	9.648485e+02	4322.111729	0.000000e+00	1.107900e+02	3.515000e+02	6.890000e+02	1.857690e+05

(a) STC Statistical Descriptive.

	count	mean	std	min	25%	50%	75%	max
tweet_id	4573.0	1.360000e+18	0.000000	1.360000e+18	1.360000e+18	1.360000e+18	1.360000e+18	1.360000e+18
rt_count	4073.0	8.070119e+01	702.810130	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	6.411000e+03
fav_count	4073.0	7.802132e+02	0.843522	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	5.400000e+01
followers	4073.0	1.101250e+03	16066.829386	0.000000e+00	1.700000e+01	6.600000e+01	2.410000e+02	7.823460e+05
friends	4073.0	1.216744e+03	2402.115086	0.000000e+00	1.950000e+02	6.150000e+02	1.653000e+03	1.057230e+05

(b) Mobily Statistical Descriptive.

	count	mean	std	min	25%	50%	75%	max
tweet_id	4947.0	1.362519e+18	4.776207e+10	1.360000e+18	1.360000e+18	1.360000e+18	1.370000e+18	1.370000e+18
rt_count	4947.0	3.259733e+01	1.898904e+02	0.000000e+00	0.000000e+00	0.000000e+00	1.000000e+00	2.668000e+03
fav_count	4947.0	1.338540e+01	6.011419e+01	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	3.400000e+01
followers	4947.0	5.917047e+04	2.051923e+05	0.000000e+00	1.800000e+01	1.400000e+02	5.730000e+02	1.545855e+05
friends	4947.0	1.298761e+03	5.685033e+03	0.000000e+00	4.300000e+01	2.090000e+02	6.880000e+02	1.845170e+05

(c) Zain Statistical Descriptive.

Fig. 2. Statistical Descriptive for Numerical Values for the Three Companies.

The collected tweets are considered unlabeled data. To use some of the ML algorithms an online tool was used to label data. The Mazajak tool¹ created by Ibrahim Abu Farha and Dr. Walid Magdy [31] classified the text polarity to Positive, Negative, or Neutral. However, there may be chances where some word polarity is not detected correctly. For example, the appearance of words that contains a prayer to Allah gives a positive labeled result. The tweet (ياربي متي ت عدل ون نكلم؟) labeled as positive, which indicated a negative opinion. Fig. 3 shows the class distribution for the three companies.

After selected Positive and Negative tweets the amount of data became 2,770 for STC, 1,552 for Mobily, and 2,553 for Zain. Therefore, the total amount of the three companies was 6,875.

C. Phase Three: Sentiment Analysis

This sentiment analysis stage refers to the primary stage. Where the analysis is performing on text-only. The binary sentiment analysis works only with Positive and Negative tweets, Tweets that rated Neutral were excluded from this analysis.

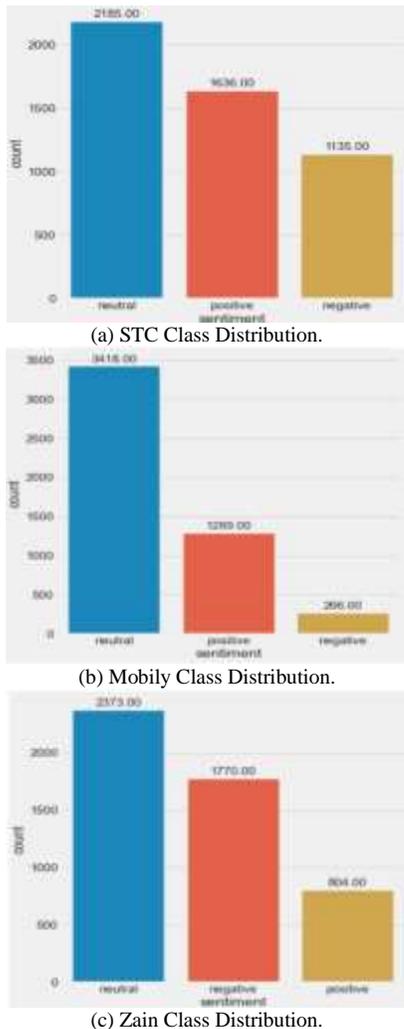


Fig. 3. Class Distribution for the Three Companies.

To prepare the target variable for classification the text must convert into a numerical record. Bag-of-Words representation works fine with a short text, is based on collect unique words in the dataset as features. Therefore, for the tweets in the proposed case Count Vectorizer was used for features extraction.

The dataset was split into training and testing sets. The training part was used to build the classification algorithms model, and the testing part to evaluate the model. Three Machine Learning (ML) algorithms were applied to the dataset Support Vector Machines (SVM), Decision Tree (DT), and Multinomial Naive Bayes (MNB).

D. Phase Four: Popularity Scoring

The popularity scoring equation takes the numbers of retweet and the favorite into polarity consideration. The popularity scoring equation can be considered as an advanced sentiment analysis stage. Since the primary stage dealt with text, the advanced stage deal with a retweet and the favorite.

A retweet means reposting someone's tweet or message in your timeline [32]. Where the popular tweets are reposted many times. Any person following you was able to see the retweeted tweet. Favorite the tweet means to keep that tweet on your account under the favorite tab. To favorite the tweet only, the user must press the heart-shaped icon. Both retweet and favorite are considered as a sign of agreement and admiration on the content.

The Popularity Scoring Equation that applied adopted from a study of [29]. It consists of two equations dependent on each other. Where the outputs of the first equation are inputs to the second equation.

The first popularity scoring equation 1 takes the sum of the classified tweet Ct (e.g. Positive classified tweet) plus the number of its retweet #RT number plus the number of its favorite #F av.

$$\text{PopularityScoring} = Ct + \#RT + \#Fav \quad (1)$$

The output from equation 1 which is the popularity scoring of the classified tweet (e.g. Positive classified tweet) multiplied by 100 and divided by the total number of the popularity scoring for all the classified tweet 2.

$$\text{DevelopedPopularityscoring} = \frac{\pi 100 * \text{PopularityScoring}}{2 \sum |P \text{opularityScoring}} \quad (2)$$

E. Phase Five: Reputation Calculation

Most reputation measurement systems rely on the opinion generated by customers, where this opinion reflects the general community's opinion [33]. From various principles of reputation measurement systems, this study is based on Bayesian systems. The Bayesian systems using the binary classes as input [34], and the reputation scores calculated based on statistical updating of beta probability density functions (PDF). To measure the reputation of service providers, the equation presented in a study was adopted [30]. The input in the sentiment analysis system is the positive and negative classes.

¹<http://mazajak.inf.ed.ac.uk:8000/>

In the proposed approach, the output from equation 2 used as input in the reputation equation. After the polarity calculated of both the positive and negative class, with the numbers of retweeting and favorite. The polarity number (e.g. positive class) took as a positive representation in the reputation equation. The equation 3 represent reputation equation:

$$R = \frac{\alpha}{\alpha + \beta} \tag{3}$$

The α represent $\alpha = r + 1$ and $\beta = s + 1$, where r is the output of popularity scoring number for positive classes and s the output of popularity scoring number for negative ones.

To clarify the equation let consider this example: if α equals 9 and β equals 3. After applying the equation 3, the reputation score can be calculated as: $R = 9/(9 + 3) = 3/4 = 0.75$.

IV. RESULTS AND EVALUATION

This section indicates the results of the calculated reputation through a proposed methodology. By applying a hybrid sentiment analysis approach on 15,000 collected tweets.

A. Evaluation Metric

The proposed sentiment analysis that was applied was based on two stages: the primary stage and the advanced stage. On the primary stage SVM, DT, and MNB classification algorithms are used to capture the sentiment. For the training model 80% of the data was used and for testing 20% was used. Moreover, to measure the performance of classification four measurements are used in this study, which are the accuracy, precision, recall, and F-score.

Accuracy: is the degree of closeness of the classified outcomes to the true value. Measurement of the accuracy is significantly important because it reflects the percentage to realize the correct pattern and polarity. Where T P indicate True Positive, TN for True Negative, F P for False Positive, and FN for False Negative.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

Precision: The ability of the model to anticipate the positive classes. It's calculated by divided true positives over a total number of true positives and false positives.

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall: The values that the model was able to determine correctly. It's calculated by divided true positives over a total number of true positives and false negatives.

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

F-Score or the F-Measure: conveys the harmonic mean between the precision and the recall. It represents the integration of both precision and recalls into a single score.

$$F - Score = \frac{Precision * Recall}{Precision + Recall} \tag{7}$$

After the primary sentiment analysis stage performed the results of the evaluation model summarize Precision, Recall, F-score, and Accuracy in Tables II, III, and IV.

From the Tables II, III, and IV, the SVM gives a higher accuracy among the three classifiers. The accuracy of the three companies for STC, Mobily, and Zain equal 92%, 95%, and 89% respectively as Fig. 4 showed. DT mostly gives low results compared to the other classifiers because some classes are imbalanced.

TABLE II. THREE CLASSIFIER RESULT FOR STC COMPANY

STC			
Support Vector Machines classifier			
Data label	Precision	Recall	F-Score
Positive	0.95	0.91	0.93
Negative	0.88	0.93	0.90
Accuracy			
0.92			
Decision Tree classifier			
Data label	Precision	Recall	F-Score
Positive	0.86	0.91	0.88
Negative	0.86	0.78	0.82
Accuracy			
0.86			
Naive Bayes classifier			
Data label	Precision	Recall	F-Score
Positive	0.96	0.81	0.88
Negative	0.78	0.96	0.86
Accuracy			
0.87			

TABLE III. THREE CLASSIFIER RESULT FOR MOBILY COMPANY

Mobily			
Support Vector Machines classifier			
Data label	Precision	Recall	F-Score
Positive	0.94	1.00	0.97
Negative	1.00	0.70	0.82
Accuracy			
0.95			
Decision Tree classifier			
Data label	Precision	Recall	F-Score
Positive	0.95	0.98	0.97
Negative	0.91	0.77	0.84
Accuracy			
0.95			
Naive Bayes classifier			
Data label	Precision	Recall	F-Score
Positive	0.96	1.00	0.98
Negative	0.98	0.79	0.88
Accuracy			
0.96			

TABLE IV. THREE CLASSIFIER RESULT FOR ZAIN COMPANY

Zain			
Support Vector Machines classifier			
Data label	Precision	Recall	F-Score
Positive	0.93	0.70	0.79
Negative	0.87	0.97	0.92
Accuracy			
0.89			
Decision Tree classifier			
Data label	Precision	Recall	F-Score
Positive	0.68	0.71	0.70
Negative	0.86	0.85	0.86
Accuracy			
0.80			
Naive Bayes classifier			
Data label	Precision	Recall	F-Score
Positive	0.87	0.65	0.74
Negative	0.85	0.95	0.90
Accuracy			
0.86			



Fig. 4. Accuracy Comparison of SVM for the Three Companies.

B. Evaluation of Reputation

As mentioned before the analysis only included the positive and negative labels. After eliminating the natural labeled, duplicated, and blanked the dataset was reduced to 6,875. Two tests were applied to capture the impact of the Popularity Scoring equation on mustering reputation.

The first test was applied without the Popularity Scoring equation, so the number of retweets and favorites did not include. Only the total number of the classified tweet (e.g. Positive classified tweet) which represents α in the reputation equation. For example, the positive classified tweet for STC is 1636, where the negative is 1134. By applying the reputation equation 3 the reputation scores equal to 0.59.

The second test by applying the Popularity Scoring equation where the number of retweets and favorites counted. The result after applying the reputation equation was higher. The reputation score for STC was 0.97.

However, taking into consideration the number of retweets and favorites for each tweet showed an improvement in

calculating reputation. Table V and Fig. 5, 6 shows the difference in reputation score between the two testing.

TABLE V. COMPARISON OF THE REPUTATION SCORES OF THE THREE COMPANIES

Without Popularity Scoring	With Popularity Scoring
STC	
0.59	0.97
Mobily	
0.82	0.74
Zain	
0.31	0.19

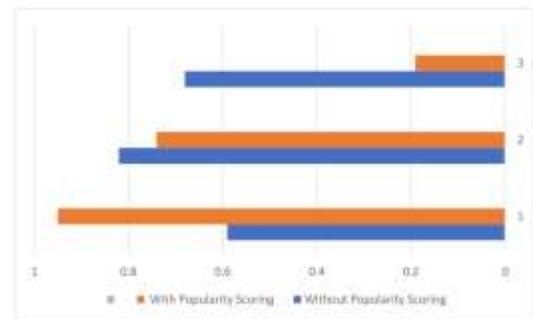
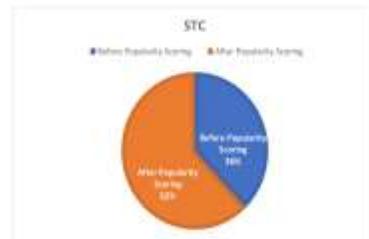


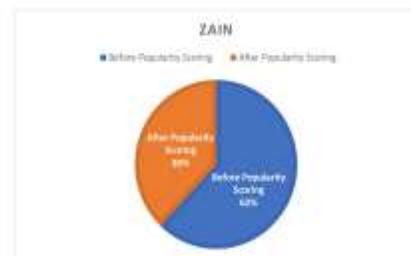
Fig. 5. Comparison of the Reputation Scores of the Three Companies.



(a) STC Reputation Score.



(b) Mobily Reputation Score.



(c) Zain Reputation Score.

Fig. 6. Reputation Scores before and after Applying Popularity Scoring Equation.

From the result, the STC reputation score was higher after applying the Popularity Scoring equation. While Mobily and Zain indicate the lowest reputation score after applying the Popularity Scoring equation. The reputation scores of Mobily and Zain decreased, and this is not due to defective or erroneous results, but after applying a Popularity Scoring equation gave more accurate results that included the re-tweet calculation and favorites.

V. CONCLUSION AND FUTURE WORK

In this study, a new opinion review presented based on a hybrid approach focused on the undertaking of sentence-level sentiment analysis to calculate reputation scores from Arabic tweets. The divergence of opinions between the customers of Telecom service providers in Twitter causes a need for a new approach to determine the reputation score of service providers involving developing a new way to compute the polarity of Arabic sentiment.

A new step was added to the traditional sentiment analysis process to enhance its accuracy. First, the significance value of counting the retweets and favorites numbers in the polarity score explains, which represents a non-verbal opinion. Second, developing a reputation approach based on the polarity score of sentiment. The first step could assist the classifier to understand and make a better accuracy about the sentiment analysis in Arabic text.

Different measures to evaluate the performance and efficacy of the classification were used in this study: the accuracy, precision, recall, and F-score. The result indicates that the SVM is the best-performed classifier, while the lowest-performing classifier is the Decision Tree. Also, the degree of reputation indicates that the STC company represents the highest reputation among its customers compared to other companies.

For future work, the approach will expand to consider multilevel word polarity instead of binary level. Also, there is a need to study the demographic characteristics of customers. In the data label, there may be a possibility where some word polarity is not noticed correctly. Currently, those instances are not handled so, it's better to use a lexicon-based approach besides labeling data by the Maza-jak tool. Moreover, the spam tweets should distinguish and eliminate to deliver more reliable reputation scores.

REFERENCES

- [1] G. Dubey, A. Rana, and N. K. Shukla, "User reviews data analysis using opinion mining on web," in 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), pp. 603–612, IEEE, 2015.
- [2] H. Akkineni, P. Lakshmi, and B. Babu, "Online crowds opinion mining it to analyze current trend: A review," International Journal of Electrical and Computer Engineering, vol. 5, no. 5, 2015.
- [3] W. He, S. Zha, and L. Li, "Social media competitive analysis and text mining: A case study in the pizza industry," International journal of information management, vol. 33, no. 3, pp. 464–472, 2013.
- [4] K. Arun, A. Srinagesh, and M. Ramesh, "Twitter sentiment analysis on demonetization tweets in india using r language," International Journal of Computer Engineering and Information Technology, vol. 9, no. 6, p. 119, 2017.
- [5] T. ministry of communication and information system, "Saudi arabia is the most twitter-crazy country in the world: Business insider."
- [6] M. M. Mostafa, "More than words: Social networks' text mining for consumer brand sentiments," Expert Systems with Applications, vol. 40, no. 10, pp. 4241–4251, 2013.
- [7] E. Jeong and S. S. Jang, "Restaurant experiences triggering positive electronic word-of-mouth (ewom) motivations," International Journal of Hospitality Management, vol. 30, no. 2, pp. 356–366, 2011.
- [8] P. Phillips, S. Barnes, K. Zigan, and R. Schegg, "Understanding the impact of online reviews on hotel performance: an empirical analysis," Journal of Travel Research, vol. 56, no. 2, pp. 235–249, 2017.
- [9] J. Nieto, R. M. Hernández-Maestro, and P. A. Muñoz-Gallego, "Marketing decisions, customer reviews, and business performance: The use of the top rural website by spanish rural lodging establishments," Tourism Management, vol. 45, pp. 115–123, 2014.
- [10] Z. Jianqiang and G. Xiaolin, "Comparison research on text pre-processing methods on twitter sentiment analysis," IEEE Access, vol. 5, pp. 2870–2879, 2017.
- [11] G. A. A. J. Alkubaisi et al., "The role of ensemble learning in stock market classification model accuracy enhancement based on naive bayes classifiers," International Journal of Statistics and Probability, vol. 9, no. 1, pp. 1–36, 2020.
- [12] K. Berezina, A. Bilgihan, C. Cobanoglu, and F. Okumus, "Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews," Journal of Hospitality Marketing & Management, vol. 25, no. 1, pp. 1–24, 2016.
- [13] G. E. Bolton, E. Katok, and A. Ockenfels, "How effective are electronic reputation mechanisms? an experimental investigation," Management science, vol. 50, no. 11, pp. 1587–1602, 2004.
- [14] Z. Xiang, U. Gretzel, and D. R. Fesenmaier, "Semantic representation of tourism on the internet," Journal of Travel Research, vol. 47, no. 4, pp. 440–453, 2009.
- [15] G. Viglia, R. Minazzi, and D. Buhalis, "The influence of e-word of-mouth on hotel occupancy rate," International Journal of Contemporary Hospitality Management, 2016.
- [16] E. Park, Y. Jang, J. Kim, N. J. Jeong, K. Bae, and A. P. Del Pobil, "Determinants of customer satisfaction with airline services: An analysis of customer feedback big data," Journal of Retailing and Consumer Services, vol. 51, pp. 186–190, 2019.
- [17] S. Hwang, J. Kim, E. Park, and S. J. Kwon, "Who will be your next customer: A machine learning approach to customer return visits in airline services," Journal of Business Research, vol. 121, pp. 121–126, 2020.
- [18] M. Gupta and S. Sebastian, "Framework to analyze customer's feedback in smartphone industry using opinion mining," International Journal of Electrical & Computer Engineering (2088-8708), vol. 8, no. 5, 2018.
- [19] A. Ghazzawi and B. Alharbi, "Analysis of customer complaints data using data mining techniques," Procedia Computer Science, vol. 163, pp. 62–69, 2019.
- [20] U. Mini et al., "Opinion mining for monitoring social media communications for brand promotion," International Journal of Advanced Research in Computer Science, vol. 8, no. 3, 2017.
- [21] L. Martin-Domingo, J. C. Martín, and G. Mandsberg, "Social media as a resource for sentiment analysis of airport service quality (asq)," Journal of Air Transport Management, vol. 78, pp. 106–115, 2019.
- [22] H. S. Manaman, S. Jamali, and A. AleAhmad, "Online reputation measurement of companies based on user-generated content in online social networks," Computers in Human Behavior, vol. 54, pp. 94–100, 2016.
- [23] A. Bukhari, U. Qamar, and U. Ghazia, "Urwf: user reputation based weightage framework for twitter micropost classification," Information Systems and e-Business Management, vol. 15, no. 3, pp. 623–659, 2017.
- [24] S. C. Anagnostopoulou, D. Buhalis, I. L. Kountouri, E. G. Manousakis, and A. E. Tsekrekos, "The impact of online reputation on hotel profitability," International Journal of Contemporary Hospitality Management, 2020.

- [25] M. J. Carrasco-Santos, A. M. Ciruela-Lorenzo, J. G. Méndez Pavón, and C. Cristófol Rodríguez, "An online reputation analysis of the tourism industry in marbella: A preliminary study on open innovation," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 7, no. 2, p. 111, 2021.
- [26] S. Shami, A. Ashfaq, and S. N. Khan, "Reputation management on social media: Analysis of audience feedback on posts and tweets of pakistani mnas.," *Journal of Political Studies*, vol. 26, no. 2, 2019.
- [27] R. M. Duwairi and I. Qarqaz, "Arabic sentiment analysis using supervised classification," in *2014 International Conference on Future Internet of Things and Cloud*, pp. 579–583, IEEE, 2014.
- [28] S. Khatoon, "Real-time twitter data analysis of saudi telecom companies for enhanced customer relationship management," *International Journal of Computer Science and Network Security*, vol. 17, no. 2, pp. 141–147, 2017.
- [29] A. S. Almalki, *A New Opinion Review Methodology for Arabic Twitter Sentiment Analysis*. The Catholic University of America, 2020.
- [30] H. Al-Hussaini and H. Al-Dossari, "A lexicon-based approach to build service provider reputation from arabic tweets in twitter," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 4, pp. 445–454, 2017.
- [31] I. A. Farha and W. Magdy, "Mazajak: An online arabic sentiment analyser," in *Proceedings of the Fourth Arabic Natural Language Processing Workshop*, pp. 192–198, 2019.
- [32] G. F. Khan, *Seven layers of social media analytics: Mining business insights from social media text, actions, networks, hyperlinks, apps, search engines, and location data*. Gohar Feroz Khan, 2015.
- [33] A. Jøsang and S. Pope, "Semantic constraints for trust transitivity," in *Proceedings of the 2nd Asia-Pacific conference on Conceptual modelling-Volume 43*, pp. 59–68, 2005.
- [34] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision support systems*, vol. 43, no. 2, pp. 618–644, 2007.