

Volume 8 Issue 7

July 2017



ISSN 2156-5570(Online)

ISSN 2158-107X(Print)



[www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)

# 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!**

**Managing Editor**  
**IJACSA**  
**Volume 8 Issue 7 July 2017**  
**ISSN 2156-5570 (Online)**  
**ISSN 2158-107X (Print)**  
**©2013 The Science and Information (SAI) Organization**

# 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

**Chao-Tung Yang**

**Department of Computer Science, Tunghai University, Taiwan**

*Domain of Research: Software Engineering and Quality, High Performance Computing, Parallel and Distributed Computing, Parallel Computing*

**Elena SCUTELNICU**

**"Dunarea de Jos" University of Galati, Romania**

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

**Krassen Stefanov**

**Professor at Sofia University St. Kliment Ohridski, Bulgaria**

*Domains of Research: e-Learning, Agents and Multi-agent Systems, Artificial Intelligence, Big Data, Cloud Computing, Data Retrieval and Data Mining, Distributed Systems, e-Learning Organisational Issues, e-Learning Tools, Educational Systems Design, Human Computer Interaction, Internet Security, Knowledge Engineering and Mining, Knowledge Representation, Ontology Engineering, Social Computing, Web-based Learning Communities, Wireless/ Mobile Applications*

**Maria-Angeles Grado-Caffaro**

**Scientific Consultant, Italy**

*Domain of Research: Electronics, Sensing and Sensor Networks*

**Mohd Helmy Abd Wahab**

**Universiti Tun Hussein Onn Malaysia**

*Domain of Research: Intelligent Systems, Data Mining, Databases*

**T. V. Prasad**

**Lingaya's University, India**

*Domain of Research: Intelligent Systems, Bioinformatics, Image Processing, Knowledge Representation, Natural Language Processing, Robotics*

## Reviewer Board Members

- **Aamir Shaikh**
- **Abbas Al-Ghaili**  
Mendeley
- **Abbas Karimi**  
Islamic Azad University Arak Branch
- **Abdelghni Lakehal**  
Université Abdelmalek Essaadi Faculté  
Polydisciplinaire de Larache Route de Rabat, Km 2 -  
Larache BP. 745 - Larache 92004. Maroc.
- **Abdul Razak**
- **Abdul Karim ABED**
- **Abdur Rashid Khan**  
Gomal University
- **Abeer Elkorany**  
Faculty of computers and information, Cairo
- **ADEMOLA ADESINA**  
University of the Western Cape
- **Aderemi A. Atayero**  
Covenant University
- **Adi Maaita**  
ISRA UNIVERSITY
- **Adnan Ahmad**
- **Adrian Branga**  
Department of Mathematics and Informatics,  
Lucian Blaga University of Sibiu
- **agana Becejski-Vujaklija**  
University of Belgrade, Faculty of organizational
- **Ahmad Saifan**  
yarmouk university
- **Ahmed Boutejdar**
- **Ahmed AL-Jumaily**  
Ahlia University
- **Ahmed Nabih Zaki Rashed**  
Menoufia University
- **Ajantha Herath**  
Stockton University Galloway
- **Akbar Hossain**
- **Akram Belghith**  
University Of California, San Diego
- **Albert S**  
Kongu Engineering College
- **Alcinia Zita Sampaio**  
Technical University of Lisbon
- **Alexane Bouënard**  
Sensopia
- **ALI ALWAN**  
International Islamic University Malaysia
- **Ali Ismail Awad**  
Luleå University of Technology
- **Alicia Valdez**
- **Amin Shaqrah**  
Taibah University
- **Amirrudin Kamsin**
- **Amitava Biswas**  
Cisco Systems
- **Anand Nayyar**  
KCL Institute of Management and Technology,  
Jalandhar
- **Andi Wahyu Rahardjo Emanuel**  
Maranatha Christian University
- **Anews Samraj**  
Mahendra Engineering College
- **Anirban Sarkar**  
National Institute of Technology, Durgapur
- **Anthony Isizoh**  
Nnamdi Azikiwe University, Awka, Nigeria
- **Antonio Formisano**  
University of Naples Federico II
- **Anuj Gupta**  
IKG Punjab Technical University
- **Anuranjan misra**  
Bhagwant Institute of Technology, Ghaziabad, India
- **Appasami Govindasamy**
- **Arash Habibi Lashkari**  
University Technology Malaysia(UTM)
- **Aree Mohammed**  
Directorate of IT/ University of Sulaimani
- **ARINDAM SARKAR**  
University of Kalyani, DST INSPIRE Fellow
- **Aris Skander**  
Constantine 1 University
- **Ashok Matani**  
Government College of Engg, Amravati
- **Ashraf Owis**  
Cairo University
- **Asoke Nath**

St. Xaviers College(Autonomous), 30 Park Street,  
Kolkata-700 016

- **Athanasios Koutras**
- **Ayad Ismaeel**  
Department of Information Systems Engineering-  
Technical Engineering College-Erbil Polytechnic  
University, Erbil-Kurdistan Region- IRAQ
- **Ayman Shehata**  
Department of Mathematics, Faculty of Science,  
Assiut University, Assiut 71516, Egypt.
- **Ayman EL-SAYED**  
Computer Science and Eng. Dept., Faculty of  
Electronic Engineering, Menofia University
- **Babatunde Opeoluwa Akinkunmi**  
University of Ibadan
- **Bae Bossoufi**  
University of Liege
- **BALAMURUGAN RAJAMANICKAM**  
Anna university
- **Balasubramanie Palanisamy**
- **BASANT VERMA**  
RAJEEV GANDHI MEMORIAL COLLEGE, HYDERABAD
- **Basil Hamed**  
Islamic University of Gaza
- **Basil Hamed**  
Islamic University of Gaza
- **Bhanu Prasad Pinnamaneni**  
Rajalakshmi Engineering College; Matrix Vision  
GmbH
- **Bharti Waman Gawali**  
Department of Computer Science & information T
- **Bilian Song**  
LinkedIn
- **Binod Kumar**  
JSPM's Jayawant Technical Campus, Pune, India
- **Bogdan Belean**
- **Bohumil Brtnik**  
University of Pardubice, Department of Electrical  
Engineering
- **Bouchaib CHERRADI**  
CRMEF
- **Brahim Raouyane**  
FSAC
- **Branko Karan**
- **Bright Keswani**  
Department of Computer Applications, Suresh Gyan  
Vihar University, Jaipur (Rajasthan) INDIA
- **Brij Gupta**

University of New Brunswick

- **C Venkateswarlu Sonagiri**  
JNTU
- **Chanashekhhar Meshram**  
Chhattisgarh Swami Vivekananda Technical  
University
- **Chao Wang**
- **Chao-Tung Yang**  
Department of Computer Science, Tunghai  
University
- **Charlie Obimbo**  
University of Guelph
- **Chee Hon Lew**
- **Chien-Peng Ho**  
Information and Communications Research  
Laboratories, Industrial Technology Research  
Institute of Taiwan
- **Chun-Kit (Ben) Ngan**  
The Pennsylvania State University
- **Ciprian Dobre**  
University Politehnica of Bucharest
- **Constantin POPESCU**  
Department of Mathematics and Computer  
Science, University of Oradea
- **Constantin Filote**  
Stefan cel Mare University of Suceava
- **CORNELIA AURORA Gyorödi**  
University of Oradea
- **Cosmina Ivan**
- **Cristina Turcu**
- **Dana PETCU**  
West University of Timisoara
- **Daniel Albuquerque**
- **Dariusz Jakóbczak**  
Technical University of Koszalin
- **Deepak Garg**  
Thapar University
- **Devena Prasad**
- **DHAYA R**
- **Dheyaa Kadhim**  
University of Baghdad
- **Djilali IDOUGH**  
University A.. Mira of Bejaia
- **Dong-Han Ham**  
Chonnam National University
- **Dr. Arvind Sharma**

Aryan College of Technology, Rajasthan Technology  
University, Kota

- **Duck Hee Lee**  
Medical Engineering R&D Center/Asan Institute for  
Life Sciences/Asan Medical Center
- **Elena SCUTELNICU**  
"Dunarea de Jos" University of Galati
- **Elena Camossi**  
Joint Research Centre
- **Eui Lee**  
Sangmyung University
- **Evgeny Nikulchev**  
Moscow Technological Institute
- **Ezekiel OKIKE**  
UNIVERSITY OF BOTSWANA, GABORONE
- **Fahim Akhter**  
King Saud University
- **FANGYONG HOU**  
School of IT, Deakin University
- **Faris Al-Salem**  
GCET
- **Firkhan Ali Hamid Ali**  
UTHM
- **Fokrul Alom Mazarbhuiya**  
King Khalid University
- **Frank Ibikunle**  
Botswana Int'l University of Science & Technology  
(BIUST), Botswana
- **Fu-Chien Kao**  
Da-Y eh University
- **Gamil Abdel Azim**  
Suez Canal University
- **Ganesh Sahoo**  
RMRIMS
- **Gaurav Kumar**  
Manav Bharti University, Solan Himachal Pradesh
- **George Pecherle**  
University of Oradea
- **George Mastorakis**  
Technological Educational Institute of Crete
- **Georgios Galatas**  
The University of Texas at Arlington
- **Gerard Dumancas**  
Oklahoma Baptist University
- **Ghalem Belalem**  
University of Oran 1, Ahmed Ben Bella
- **gherabi noreddine**

- **Giacomo Veneri**  
University of Siena
- **Giri Babu**  
Indian Space Research Organisation
- **Govindarajulu Salendra**
- **Grebenisan Gavril**  
University of Oradea
- **Gufan Ahmad Ansari**  
Qassim University
- **Gunaseelan Devaraj**  
Jazan University, Kingdom of Saudi Arabia
- **GYÖRÖDI ROBERT STEFAN**  
University of Oradea
- **Hadj Tadjine**  
IAV GmbH
- **Haewon Byeon**  
Nambu University
- **Haiguang Chen**  
ShangHai Normal University
- **Hamid Alinejad-Rokny**  
The University of New South Wales
- **Hamid AL-Asadi**  
Department of Computer Science, Faculty of  
Education for Pure Science, Basra University
- **Hamid Mukhtar**  
National University of Sciences and Technology
- **Hany Hassan**  
EPF
- **Harco Leslie Henic SPITS WARNARS**  
Bina Nusantara University
- **Hariharan Shanmugasundaram**  
Associate Professor, SRM
- **Harish Garg**  
Thapar University Patiala
- **Hazem I. El Shekh Ahmed**  
Pure mathematics
- **Hemalatha SenthilMahesh**
- **Hesham Ibrahim**  
Faculty of Marine Resources, Al-Mergheb  
University
- **Himanshu Aggarwal**  
Department of Computer Engineering
- **Hongda Mao**  
Hossam Faris
- **Huda K. AL-Jobori**  
Ahlia University
- **Imed JABRI**

- **iss EL OUADGHIRI**
- **Iwan Setyawan**  
Satya Wacana Christian University
- **Jacek M. Czerniak**  
Casimir the Great University in Bydgoszcz
- **Jai Singh W**
- **JAMAIAH HAJI YAHAYA**  
NORTHERN UNIVERSITY OF MALAYSIA (UUM)
- **James Coleman**  
Edge Hill University
- **Jatinderkumar Saini**  
Narmada College of Computer Application, Bharuch
- **Javed Sheikh**  
University of Lahore, Pakistan
- **Jayaram A**  
Siddaganga Institute of Technology
- **Ji Zhu**  
University of Illinois at Urbana Champaign
- **Jia Uddin Jia**  
Assistant Professor
- **Jim Wang**  
The State University of New York at Buffalo,  
Buffalo, NY
- **John Sahlin**  
George Washington University
- **JOHN MANOHAR**  
VTU, Belgaum
- **JOSE PASTRANA**  
University of Malaga
- **Jui-Pin Yang**  
Shih Chien University
- **Jyoti Chaudhary**  
high performance computing research lab
- **K V.L.N.Acharyulu**  
Bapatla Engineering college
- **Ka-Chun Wong**
- **Kamatchi R**
- **Kamran Kowsari**  
The George Washington University
- **KANNADHASAN SURIYAN**
- **Kashif Nisar**  
Universiti Utara Malaysia
- **Kato Mivule**
- **Kayhan Zrar Ghafoor**  
University Technology Malaysia
- **Kennedy Okafor**  
Federal University of Technology, Owerri
- **Khalid Mahmood**  
IEEE
- **Khalid Sattar Abdul**  
Assistant Professor
- **Khin Wee Lai**  
Biomedical Engineering Department, University  
Malaya
- **Khurram Khurshid**  
Institute of Space Technology
- **KIRAN SREE POKKULURI**  
Professor, Sri Vishnu Engineering College for  
Women
- **KITIMAPORN CHOOCHOTE**  
Prince of Songkla University, Phuket Campus
- **Krasimir Yordzhev**  
South-West University, Faculty of Mathematics and  
Natural Sciences, Blagoevgrad, Bulgaria
- **Krassen Stefanov**  
Professor at Sofia University St. Kliment Ohridski
- **Labib Gergis**  
Misr Academy for Engineering and Technology
- **LATHA RAJAGOPAL**
- **Lazar Stošić**  
College for professional studies educators  
Aleksinac, Serbia
- **Leanos Maglaras**  
De Montfort University
- **Leon Abdillah**  
Bina Darma University
- **Lijian Sun**  
Chinese Academy of Surveying and
- **Ljubomir Jerinic**  
University of Novi Sad, Faculty of Sciences,  
Department of Mathematics and Computer Science
- **Lokesh Sharma**  
Indian Council of Medical Research
- **Long Chen**  
Qualcomm Incorporated
- **M. Reza Mashinchi**  
Research Fellow
- **M. Tariq Banday**  
University of Kashmir
- **madjid khalilian**
- **majzoob omer**
- **Mallikarjuna Doodipala**  
Department of Engineering Mathematics, GITAM  
University, Hyderabad Campus, Telangana, INDIA

- **Manas deep**  
Masters in Cyber Law & Information Security
- **Manju Kaushik**
- **Manoharan P.S.**  
Associate Professor
- **Manoj Wadhwa**  
Echelon Institute of Technology Faridabad
- **Manpreet Manna**  
Director, All India Council for Technical Education,  
Ministry of HRD, Govt. of India
- **Manuj Darbari**  
BBD University
- **Marcellin Julius Nkenlifack**  
University of Dschang
- **Maria-Angeles Grado-Caffaro**  
Scientific Consultant
- **Marwan Alseid**  
Applied Science Private University
- **Mazin Al-Hakeem**  
LFU (Lebanese French University) - Erbil, IRAQ
- **Md Islam**  
sikkim manipal university
- **Md. Bhuiyan**  
King Faisal University
- **Md. Zia Ur Rahman**  
Narasaraopeta Engg. College, Narasaraopeta
- **Mehdi Bahrami**  
University of California, Merced
- **Messaouda AZZOUZI**  
Ziane Achour University of Djelfa
- **Milena Bogdanovic**  
University of Nis, Teacher Training Faculty in Vranje
- **Miriampally Venkata Raghavendra**  
Adama Science & Technology University, Ethiopia
- **Mirjana Popovic**  
School of Electrical Engineering, Belgrade University
- **Miroslav Baca**  
University of Zagreb, Faculty of organization and  
informatics / Center for biometrics
- **Moeiz Miraoui**  
University of Gafsa
- **Mohamed Eldosoky**
- **Mohamed Ali Mahjoub**  
Preparatory Institute of Engineer of Monastir
- **Mohamed Kaloup**
- **Mohamed El-Sayed**  
Faculty of Science, Fayoum University, Egypt
- **Mohamed Najeh LAKHOUA**  
ESTI, University of Carthage
- **Mohammad Ali Badamchizadeh**  
University of Tabriz
- **Mohammad Jannati**
- **Mohammad Alomari**  
Applied Science University
- **Mohammad Haghighat**  
University of Miami
- **Mohammad Azzeh**  
Applied Science university
- **Mohammed Akour**  
Yarmouk University
- **Mohammed Sadgal**  
Cadi Ayyad University
- **Mohammed Al-shabi**  
Associate Professor
- **Mohammed Hussein**
- **Mohammed Kaiser**  
Institute of Information Technology
- **Mohammed Ali Hussain**  
Sri Sai Madhavi Institute of Science & Technology
- **Mohd Helmy Abd Wahab**  
University Tun Hussein Onn Malaysia
- **Mokhtar Beldjehem**  
University of Ottawa
- **Mona Elshinawy**  
Howard University
- **Mostafa Ezziyani**  
FSTT
- **Mouhammd sharari alkasassbeh**
- **Mourad Amad**  
Laboratory LAMOS, Bejaia University
- **Mueen Uddin**  
University Malaysia Pahang
- **MUNTASIR AL-ASFOOR**  
University of Al-Qadisiyah
- **Murphy Choy**
- **Murthy Dasika**  
Geethanjali College of Engineering & Technology
- **Mustapha OUJAOURA**  
Faculty of Science and Technology Béni-Mellal
- **MUTHUKUMAR SUBRAMANYAM**  
DGCT, ANNA UNIVERSITY
- **N.Ch. Iyengar**  
VIT University
- **Nagy Darwish**

Department of Computer and Information Sciences,  
Institute of Statistical Studies and Researches, Cairo  
University

- **Najib Kofahi**  
Yarmouk University
- **Nan Wang**  
LinkedIn
- **Natarajan Subramanyam**  
PES Institute of Technology
- **Natheer Gharaibeh**  
College of Computer Science & Engineering at  
Yanbu - Taibah University
- **Nazeeh Ghatasheh**  
The University of Jordan
- **Nazeeruddin Mohammad**  
Prince Mohammad Bin Fahd University
- **NEERAJ SHUKLA**  
ITM UNiversity, Gurgaon, (Haryana) India
- **Neeraj Tiwari**
- **Nestor Velasco-Bermeo**  
UPFIM, Mexican Society of Artificial Intelligence
- **Nidhi Arora**  
M.C.A. Institute, Ganpat University
- **Nilanjan Dey**
- **Ning Cai**  
Northwest University for Nationalities
- **Nithyanandam Subramanian**  
Professor & Dean
- **Noura Aknin**  
University Abdelamlek Essaadi
- **Obaida Al-Hazaimeh**  
Al- Balqa' Applied University (BAU)
- **Oliviu Matei**  
Technical University of Cluj-Napoca
- **Om Sangwan**
- **Omaima Al-Allaf**  
Asesstant Professor
- **Osama Omer**  
Aswan University
- **Ouchtati Salim**
- **Ousmane THIARE**  
Associate Professor University Gaston Berger of  
Saint-Louis SENEGAL
- **Paresh V Virparia**  
Sardar Patel University
- **Peng Xia**  
Microsoft

- **Ping Zhang**  
IBM
- **Poonam Garg**  
Institute of Management Technology, Ghaziabad
- **Prabhat K Mahanti**  
UNIVERSITY OF NEW BRUNSWICK
- **PROF DURGA SHARMA ( PHD)**  
AMUIT, MOEFDRE & External Consultant (IT) &  
Technology Tansfer Research under ILO & UNDP,  
Academic Ambassador for Cloud Offering IBM-USA
- **Purwanto Purwanto**  
Faculty of Computer Science, Dian Nuswantoro  
University
- **Qifeng Qiao**  
University of Virginia
- **Rachid Saadane**  
EE departement EHTP
- **Radwan Tahboub**  
Palestine Polytechnic University
- **raed Kanaan**  
Amman Arab University
- **Raghuraj Singh**  
Harcourt Butler Technological Institute
- **Rahul Malik**
- **raja boddu**  
LENORA COLLEGE OF ENGINEERNG
- **Raja Ramachandran**
- **Rajesh Kumar**  
National University of Singapore
- **Rakesh Dr.**  
Madan Mohan Malviya University of Technology
- **Rakesh Balabantaray**  
IIIT Bhubaneswar
- **Ramani Kannan**  
Universiti Teknologi PETRONAS, Bandar Seri  
Iskandar, 31750, Tronoh, Perak, Malaysia
- **Rashad Al-Jawfi**  
Ibb university
- **Rashid Sheikh**  
Shri Aurobindo Institute of Technology, Indore
- **Ravi Prakash**  
University of Mumbai
- **RAVINA CHANGALA**
- **Ravisankar Hari**  
CENTRAL TOBACCO RESEARCH INSTITUE
- **Rawya Rizk**  
Port Said University

- **Reshmy Krishnan**  
Muscat College affiliated to Stirling University.U
- **Ricardo Vardasca**  
Faculty of Engineering of University of Porto
- **Ritaban Dutta**  
ISSL, CSIRO, Tasmania, Australia
- **Rowayda Sadek**
- **Ruchika Malhotra**  
Delhi Technological University
- **Rutvij Jhaveri**  
Gujarat
- **SAADI Slami**  
University of Djelfa
- **Sachin Kumar Agrawal**  
University of Limerick
- **Sagarmay Deb**  
Central Queensland University, Australia
- **Said Ghoniemy**  
Taif University
- **Sandeep Reddivari**  
University of North Florida
- **Sanskriti Patel**  
Charotar University of Science & Technology,  
Changa, Gujarat, India
- **Santosh Kumar**  
Graphic Era University, Dehradun (UK)
- **Sasan Adibi**  
Research In Motion (RIM)
- **Satyena Singh**  
Professor
- **Sebastian Marius Rosu**  
Special Telecommunications Service
- **Seema Shah**  
Vidyalankar Institute of Technology Mumbai
- **Seifedine Kadry**  
American University of the Middle East
- **Selem Charfi**  
HD Technology
- **SENGOTTUVELAN P**  
Anna University, Chennai
- **Senol Piskin**  
Istanbul Technical University, Informatics Institute
- **Sérgio Ferreira**  
School of Education and Psychology, Portuguese  
Catholic University
- **Seyed Hamidreza Mohades Kasaei**  
University of Isfahan
- **Shafiqul Abidin**  
HMR Institute of Technology & Management  
(Affiliated to GGSIP University), Hamidpur, Delhi -  
110036
- **Shahanawaj Ahamad**  
The University of Al-Kharj
- **Shaidah Jusoh**
- **Shaiful Bakri Ismail**
- **Shakir Khan**  
Al-Imam Muhammad Ibn Saud Islamic University
- **Shawki Al-Dubae**  
Assistant Professor
- **Sherif Hussein**  
Mansoura University
- **Shriram Vasudevan**  
Amrita University
- **Siddhartha Jonnalagadda**  
Mayo Clinic
- **Sim-Hui Tee**  
Multimedia University
- **Simon Ewedafe**  
The University of the West Indies
- **Siniša Opic**  
University of Zagreb, Faculty of Teacher Education
- **Sivakumar Poruran**  
SKP ENGINEERING COLLEGE
- **Slim BEN SAOUD**  
National Institute of Applied Sciences and  
Technology
- **Sofien Mhatli**
- **sofyan Hayajneh**
- **Sohail Jabbar**  
Bahria University
- **Sri Devi Ravana**  
University of Malaya
- **Sudarson Jena**  
GITAM University, Hyderabad
- **Suhail Sami Owais Owais**
- **Suhas J Manangi**  
Microsoft
- **SUKUMAR SENTHILKUMAR**  
Universiti Sains Malaysia
- **Süleyman Eken**  
Kocaeli University
- **Sumazly Sulaiman**  
Institute of Space Science (ANGKASA), Universiti  
Kebangsaan Malaysia

- **Sumit Goyal**  
National Dairy Research Institute
- **Supareerk Janjarasjitt**  
Ubon Ratchathani University
- **Suresh Sankaranarayanan**  
Institut Teknologi Brunei
- **Susarla Sastry**  
JNTUK, Kakinada
- **Suseendran G**  
Vels University, Chennai
- **Suxing Liu**  
Arkansas State University
- **Syed Ali**  
SMI University Karachi Pakistan
- **T C.Manjunath**  
HKBK College of Engg
- **T V Narayana rao Rao**  
SNIST
- **T. V. Prasad**  
Lingaya's University
- **Taiwo Ayodele**  
Infonetmedia/University of Portsmouth
- **Talal Bonny**  
Department of Electrical and Computer Engineering, Sharjah University, UAE
- **Tamara Zhukabayeva**
- **Tarek Gharib**  
Ain Shams University
- **thabet slimani**  
College of Computer Science and Information Technology
- **Totok Biyanto**  
Engineering Physics, ITS Surabaya
- **Touati Youcef**  
Computer sce Lab LIASD - University of Paris 8
- **Tran Sang**  
IT Faculty - Vinh University - Vietnam
- **Tsvetanka Georgieva-Trifonova**  
University of Veliko Tarnovo
- **Uchechukwu Awada**  
Dalian University of Technology
- **Udai Pratap Rao**
- **Urmila Shrawankar**  
GHRCE, Nagpur, India
- **Vaka MOHAN**  
TRR COLLEGE OF ENGINEERING
- **VENKATESH JAGANATHAN**
- **ANNA UNIVERSITY**
- **Vinayak Bairagi**  
AISSMS Institute of Information Technology, Pune
- **Vishnu Mishra**  
SVNIT, Surat
- **Vitus Lam**  
The University of Hong Kong
- **VUDA SREENIVASARAO**  
PROFESSOR AND DEAN, St.Mary's Integrated Campus, Hyderabad
- **Wali Mashwani**  
Kohat University of Science & Technology (KUST)
- **Wei Wei**  
Xi'an Univ. of Tech.
- **Wenbin Chen**  
360Fly
- **Xi Zhang**  
illinois Institute of Technology
- **Xiaojing Xiang**  
AT&T Labs
- **Xiaolong Wang**  
University of Delaware
- **Yanping Huang**
- **Yao-Chin Wang**
- **Yasser Albagory**  
College of Computers and Information Technology, Taif University, Saudi Arabia
- **Yasser Alginahi**
- **Yi Fei Wang**  
The University of British Columbia
- **Yihong Yuan**  
University of California Santa Barbara
- **Yilun Shang**  
Tongji University
- **Yu Qi**  
Mesh Capital LLC
- **Zacchaeus Omogbadegun**  
Covenant University
- **Zairi Rizman**  
Universiti Teknologi MARA
- **Zarul Zaaba**  
Universiti Sains Malaysia
- **Zenzo Ncube**  
North West University
- **Zhao Zhang**  
Deptment of EE, City University of Hong Kong
- **Zhihan Lv**

Chinese Academy of Science

- **Zhixin Chen**  
ILX Lightwave Corporation
- **Ziyue Xu**  
National Institutes of Health, Bethesda, MD

- **Zlatko Stacic**  
University of Zagreb, Faculty of Organization and  
Informatics Varazdin
- **Zuraini Ismail**  
Universiti Teknologi Malaysia

# CONTENTS

**Paper 1: A Better Comparison Summary of Credit Scoring Classification**

*Authors: Sharjeel Imtiaz, Allan J. Brimicombe*

**PAGE 1 – 4**

**Paper 2: MobisenseCar: A Mobile Crowd-Based Architecture for Data Acquisition and Processing in Vehicle-Based Sensing**

*Authors: Lionel Nkenyereye, Jong Wook Jang*

**PAGE 5 – 18**

**Paper 3: Impedance Matching of a Microstrip Antenna**

*Authors: Sameh Khmailia, Hichem Taghouti, Sabri Jmal, Abdelkader Mami*

**PAGE 19 – 23**

**Paper 4: Development of A Clinically-Oriented Expert System for Differentiating Melanocytic from Non-melanocytic Skin Lesions**

*Authors: Qaisar Abbas*

**PAGE 24 – 29**

**Paper 5: Feature Selection and Extraction Framework for DNA Methylation in Cancer**

*Authors: Abeer A. Raweh, Mohammad Nassef, Amr Badr*

**PAGE 30 – 36**

**Paper 6: Research Pathway towards MAC Protocol in Enhancing Network Performance in Wireless Sensor Network (WSN)**

*Authors: Anitha K, Usha S*

**PAGE 37 – 44**

**Paper 7: Reduced-Latency and Area-Efficient Architecture for FPGA-Based Stochastic LDPC Decoders**

*Authors: Ghania Zerari, Abderrezak Guessoum*

**PAGE 45 – 51**

**Paper 8: Towards an SOA Architectural Model for AAL-PaaS Design and Implimentation Challenges**

*Authors: El murabet Amina, Abtoy Anouar, Abdellah Touhafi, Abderahim Tahiri*

**PAGE 52 – 56**

**Paper 9: Core Levels Algorithm for Optimization: Case of Microwave Models**

*Authors: Ali Haydar, Ezgi Deniz Ülker, Kamil Dimililer, Sadık Ülker*

**PAGE 57 – 64**

**Paper 10: A Proposed Adaptive Scheme for Arabic Part-of Speech Tagging**

*Authors: Mohammad Fasha*

**PAGE 65 – 75**

**Paper 11: The Performance of the Bond Graph Approach for Diagnosing Electrical Systems**

*Authors: Dhia Mzoughi, Abderrahmene Sallami, Abdelkader Mami*

**PAGE 76 – 81**

Paper 12: *Application of the Tabu Search Algorithm to Cryptography*

Authors: Zakaria KADDOURI, Fouzia OMARY

PAGE 82 – 87

Paper 13: *Hearing Aid Method by Equalizing Frequency Response of Phoneme Extracted from Human Voice*

Authors: Kohei Arai, Takuto Konishi

PAGE 88 – 93

Paper 14: *New Divide and Conquer Method on Endmember Extraction Techniques*

Authors: Ihab Samir, Bassam Abdellatif, Amr Badr

PAGE 94 – 100

Paper 15: *Real-Time Analysis of Students' Activities on an E-Learning Platform based on Apache Spark*

Authors: Abdelmajid Chaffai, Larbi Hassouni, Houda Anoun

PAGE 101 – 109

Paper 16: *Reducing Dimensionality in Text Mining using Conjugate Gradients and Hybrid Cholesky Decomposition*

Authors: Jasem M. Alostad

PAGE 110 – 116

Paper 17: *An Efficient Distributed Traffic Events Generator for Smart Highways*

Authors: Abdelaziz Daaif, Omar Bouattane, Mohamed Youssfi, Oum El Kheir Abra

PAGE 117 – 127

Paper 18: *The Role of Strategic Information Systems (SIS) in Supporting and Achieving the Competitive Advantages (CA): An Empirical Study on Saudi Banking Sector*

Authors: Nisreen F. Alshubaily, Abdullah A. Alfameem

PAGE 128 – 139

Paper 19: *The Optimization of Query Processing in Seabase Cloud Databases based on CCEVP Model*

Authors: Abdulkadir ÖZDEMİR, Hasan Asil

PAGE 140 – 143

Paper 20: *A Copula Statistic for Measuring Nonlinear Dependence with Application to Feature Selection in Machine Learning*

Authors: Mohsen Ben Hassine, Lamine Mili, Kiran Karra

PAGE 144 – 154

Paper 21: *Security in OpenFlow Enabled Cloud Environment*

Authors: Abdalla Alameen, Sadia Rubab, Bhawna Dhupia, Manjur Kolhar

PAGE 155 – 163

Paper 22: *Analysis of Received Power Characteristics of Commercial Photodiodes in Indoor Los Channel Visible Light Communication*

Authors: Syfaul Fuada, Angga Pratama Putra, Trio Adiono

PAGE 164 – 172

**Paper 23: An Efficient Spectral Amplitude Coding (SAC) Technique for Optical CDMA System using Wavelength Division Multiplexing (WDM) Concepts**

*Authors: Hassan Yousif Ahmed, Medien Zeghid*

**PAGE 173 – 179**

**Paper 24: Efficient Key Agreement and Nodes Authentication Scheme for Body Sensor Networks**

*Authors: Jawaid Iqbal, Noor ul Amin, Arif Iqbal Umar, Nizamud Din*

**PAGE 180 – 187**

**Paper 25: Design of Efficient Pipelined Router Architecture for 3D Network on Chip**

*Authors: Bouraoui Chemli, Abdelkrim Zitouni, Alexandre Coelho, Raoul Velazco*

**PAGE 188 – 194**

**Paper 26: UHF RFID Reader Antenna using Novel Planar Metamaterial Structure for RFID System**

*Authors: Bouraoui Chemli, Abdelkrim Zitouni, Alexandre Coelho, Raoul Velazco*

**PAGE 195 – 199**

**Paper 27: Factors Influencing Users' Intentions to Use Mobile Government Applications in Saudi Arabia: TAM Applicability**

*Authors: Raed Alotaibi, Luke Houghton, Kuldeep Sandhu*

**PAGE 200 – 211**

**Paper 28: Image and AES Inspired Hex Symbols Steganography (IAIS) for Anti-Forensic Artifacts**

*Authors: Somyia M. Abu Asbeh, Sarah M. Hammoudeh, Arab M. Hammoudeh*

**PAGE 212 – 219**

**Paper 29: Network Traffic Classification using Machine Learning Techniques over Software Defined Networks**

*Authors: Mohammad Reza Parsaei, Mohammad Javad Sobouti, Seyed Raouf khayami, Reza Javidan*

**PAGE 220 – 225**

**Paper 30: A New Approach for Leukemia Identification based on Cepstral Analysis and Wavelet Transform**

*Authors: Amira Samy Talaat Abou Taleb, Amir F. Atiya*

**PAGE 226 – 232**

**Paper 31: A New Strategy of Validities' Computation for Multimodel Approach: Experimental Validation**

*Authors: Abdennacer BEN MESSAOUD, Samia TALMOUDI BEN AOUN, Moufida LAHMARI KSOURI*

**PAGE 233 – 241**

**Paper 32: Simulation and Analysis of Quality of Service (QoS) Parameters of Voice over IP (VoIP) Traffic through Heterogeneous Networks**

*Authors: Mahdi H. Miraz, Suhail A. Molvi, Muzafar A. Ganie, Maaruf Ali, AbdelRahman H. Hussein*

**Page 242 – 248**

**Paper 33: A Comparative Analysis of Quality Assurance of Mobile Applications using Automated Testing Tools**

*Authors: Haneen Anjum, Muhammad Imran Babar, Muhammad Jehanzeb, Maham Khan, Saima Chaudhry, Summiyah Sultana, Zainab Shahid, Furkh Zeshan, Shahid Nazir Bhatti*

**Page 249 – 255**

**Paper 34: Detection of Cardiac Disease using Data Mining Classification Techniques**

*Authors: Abdul Aziz, Aziz Ur Rehman*

**Page 256 – 259**

Paper 35: Image Encryption Technique based on the Entropy Value of a Random Block

Authors: Mohammed A. F. Al-Husainy, Daa Mohammed Uliyan

Page 260 – 266

Paper 36: Enhancing Lean Software Development by using Devops Practices

Authors: Ahmed Bahaa Farid, Yehia Mostafa Helmy, Mahmoud Mohamed Bahloul

Page 267 – 277

Paper 37: SEUs Mitigation on Program Counter of the LEON3 Soft Processor

Authors: Afef KCHAOU, Wajih EL HADJ YOUSSEF, Rached TOURKI

Page 278 – 285

Paper 38: Deep Learning based Computer Aided Diagnosis System for Breast Mammograms

Authors: M. Arfan Jaffar

Page 286 – 290

Paper 39: ODSA: A Novel Ordering Divisional Scheduling Algorithm for Modern Operating Systems

Authors: Junaid Haseeb, Muhammad Tayyab, Khizar Hameed, Samia Rehman, Muhammad Junaid, Agha Muhammad Musa Khan

Page 291 – 296

Paper 40: Customized Descriptor for Various Obstacles Detection in Road Scene

Authors: Haythem Ameer, Abdelhamid Helali, J. Ramirez, J. M. Gorriz, Ridha Mghaieth, Hassen Maaref

Page 297 – 306

Paper 41: A Mathematical Model for Comparing Memory Storage of Three Interval-Based Parametric Temporal Database Models

Authors: Nashwan Alromema, Mohd Shafry Mohd Rahim, Ibrahim Albidewi

Page 307 – 313

Paper 42: Semantic based Data Integration in Scientific Workflows

Authors: M. Abdul Rahman, Jamil Ahmed, Ahmed Waqas, Ajmal Sawand

Page 314 – 325

Paper 43: Comparative Analysis of Online Rating Systems

Authors: Mohammad Azzeh

Page 326 – 330

Paper 44: Method for System Requirements Approval

Authors: Lindita Nebiu Hyseni, Zamir Dika

Page 331 – 336

Paper 45: A Novel Modeling based Agent Cellular Automata for Advanced Residential Mobility Applications

Authors: Elarbi Elalaouy, Khadija Rhoulemi, Moulay Driss Rahmani

Page 337 – 343

Paper 46: Intelligent Diagnostic System for Nuclei Structure Classification of Thyroid Cancerous and Non-Cancerous Tissues

Authors: Jamil Ahmed Chandio, M. Abdul Rehman Soomrani

Page 344 – 352

**Paper 47: Mobility for an Optimal Data Collection in Wireless Sensor Networks**

*Authors: EZ-ZAIDI Asmaa, RAKRAK Said*

**Page 353 – 360**

**Paper 48: Estimating True Demand in Airline's Revenue Management Systems using Observed Sales**

*Authors: Alireza Nikseresht, Koorush Ziarati*

**Page 361 – 369**

**Paper 49: 2.5 D Facial Analysis via Bio-Inspired Active Appearance Model and Support Vector Machine for Forensic Application**

*Authors: Siti Norul Huda Sheikh Abdullah, Mohammed Hasan Abdulameer, Nazri Ahmad Zamani, Fasly Rahim, Khairul Akram Zainol Ariffin, Zulaiha Othman, Mohd Zakree Ahmad Nazri*

**Page 370 – 375**

**Paper 50: Efficient Feature Selection for Product Labeling over Unstructured Data**

*Authors: Zeki YETGIN, Abdullah ELEWI, Furkan GÖZÜKARA*

**Page 376 – 381**

**Paper 51: Intelligent System for Detection of Micro-Calcification in Breast Cancer**

*Authors: M. Abdul Rahman, Jamil Ahmed, Ahmed Waqas, Ajmal Sawand*

**Page 382 – 387**

**Paper 52: Financial Market Prediction using Google Trends**

*Authors: Farrukh Ahmed, Dr. Raheela Asif, Dr. Saman Hina, Muhammad Muzammil*

**Page 388 – 391**

**Paper 53: Ultra-Wideband Antenna Design for GPR Applications: A Review**

*Authors: Jawad Ali, Noorsaliza Abdullah, Muhammad Yusof Ismail, Ezri Mohd, Shaharil Mohd Shah*

**Page 392 – 400**

**Paper 54: Introducing Time based Competitive Advantage in IT Sector with Simulation**

*Authors: Rida Maryam, Junaid Haseeb, Muhammad Tayyab, Adnan Naseem, Khizar Hameed, Babar Shahzaad*

**Page 401 – 406**

**Paper 55: New Deep Kernel Learning based Models for Image Classification**

*Authors: Rabha O.Abd-elsalam, Yasser F.Hassan, Mohamed W.Saleh*

**Page 407 – 411**

**Paper 56: Low-Power Hardware Design of Binary Arithmetic Encoder in H.264**

*Authors: Ben Hamida Asma, Nedra Jarray, Zitouni Abdelkrim*

**Page 412 – 416**

**Paper 57: Dynamic Access Control Policy based on Blockchain and Machine Learning for the Internet of Things**

*Authors: Aissam OUTCHAKOUCT, Hamza ES-SAMAALI, Jean Philippe LEROY*

**Page 417 – 424**

**Paper 58: Eye Controlled Mobile Robot with Shared Control for Physically Impaired People**

*Authors: Muhammad Wasim, Javeria Khan, Dawer Saeed, Dr. Usman Ghani Khan*

**Page 425 – 430**

**Paper 59: Fast-ICA for Mechanical Fault Detection and Identification in Electromechanical Systems for Wind Turbine Applications**

*Authors: Mohamed Farhat, Yasser Gritli, Mohamed Benrejeb*

**Page 431 – 439**

**Paper 60: An Enhanced Approach for Detection and Classification of Computed Tomography Lung Cancer**

*Authors: Wafaa Alakwaa, Mohammad Nassef, Amr Badr*

**Page 440 – 446**

**Paper 61: Short Survey on Static Hand Gesture Recognition**

*Authors: Huu-Hung Huynh, Duc-Hoang Vo*

**Page 447 – 451**

**Paper 62: Mobility based Net Ordering for Simultaneous Escape Routing**

*Authors: Kashif Sattar, Aleksandar Ignjatovic, Anjum Naveed, Muhammad Zeeshan*

**Page 452 – 461**

**Paper 63: A Survey on User Interfaces for Interaction with Human and Machines**

*Authors: Mirza Abdur Razzaq, Muhammad Ali Qureshi, Kashif Hussain Memon, Saleem Ullah*

**Page 462 – 467**

**Paper 64: OSPF vs EIGRP: A Comparative Analysis of CPU Utilization using OPNET**

*Authors: Muhammad Kashif Hanif, Ramzan Talib, Nafees Ayub, Muhammad Umer Sarwar, Sami Ullah*

**Page 468 – 471**

**Paper 65: Anonymized Social Networks Community Preservation**

*Authors: Jyothi Vadisala, Valli Kumari Vatsavayi*

**Page 472 – 476**

**Paper 66: A Comparative Study for Performance and Power Consumption of FPGA Digital Interpolation Filters**

*Authors: Tim Donnelly, Jungu Choi, Alexander V. Kildishev, Matthew Swabey, Mark C. Johnson*

**Page 477 – 483**

**Paper 67: A Survey of Datasets for Biomedical Question Answering Systems**

*Authors: Muhammad Wasim, Dr. Waqar Mahmood, Dr. Usman Ghani Khan*

**Page 484 – 488**

**Paper 68: A Comprehensive Analysis on the Security Threats and their Countermeasures of IoT**

*Authors: Abdul Wahab Ahmed, Omair Ahmad Khan, Mian Muhammad Ahmed, Munam Ali Shah*

**Page 489 – 501**

**Paper 69: Ladder Networks: Learning under Massive Label Deficit**

*Authors: Behroz Mirza, Tahir Syed, Jamshed Memon, Yameen Malik*

**Page 502 – 507**

**Paper 70: Privacy-preserving Twitter-based Solution for Visually Impaired People**

*Authors: Dina Ahmed Abdraboo, Tarek Gaber, Mohamed El Sayed Wahed*

**Page 508 – 512**

**Paper 71: A Text based Authentication Scheme for Improving Security of Textual Passwords**

*Authors: Shah Zaman Nizamani, Syed Raheel Hassan, Tariq Jamil Khanzada, Mohd Zalisham Jali*

**Page 513 – 521**

**Paper 72: Impact of Pulse Voltage as Desulfator to Improve Automotive Lead Acid Battery Capacity**

*Authors: EL MEHDI LAADISSI, ANAS EL FILALI, MALIKA ZAZI*

**Page 522 – 526**

**Paper 73: The Dynamics of IT Workaround Practices - A Theoretical Concept and an Empirical Assessment**

*Authors: Ahmed Alojairi*

**Page 527 – 534**

# A Better Comparison Summary of Credit Scoring Classification

Sharjeel Imtiaz  
School of Computing  
Doctorate Data Science Student  
University of East London (UEL)  
London, United Kingdom

Professor Allan J. Brimicombe  
School of Computing Head,  
Centre for Geo-Information Studies  
University of East London  
London, United Kingdom

**Abstract**—The credit scoring aim is to classify the customer credit as defaulter or non-defaulter. The credit risk analysis is more effective with further boosting and smoothing of the parameters of models. The objective of this paper is to explore the credit score classification models with an imputation technique and without imputation technique. However, data availability is low in case of without imputation because of missing values depletion from the large dataset. On the other hand, imputation based dataset classification accuracy with linear method of ANN is better than other models. The comparison of models with boosting and smoothing shows that error rate is better metric than area under curve (AUC) ratio. It is concluded that artificial neural network (ANN) is better alternative than decision tree and logistic regression when data availability is high in dataset.

**Keywords**—Credit score data mining; classification; artificial neural network; imputation

## I. INTRODUCTION

In 2006 Taiwan faced credit flow crises as matter of fact to propel in market the bank has over issued cash and credit regardless of their ability to repayments, overconsumptions history. This situation blows consumer and finance rapport badly. The well down financial system crisis on downstream and risk management on upstream. The purpose of risk management is to enforce checks on consumer ability to repay bills thus reduce damage and consumer's credit repayment uncertainty.

A lender commonly makes two types of decisions: first, whether to grant credit to a new application or not, and second, how to deal with existing application, including whether to increase their credit limits or not [1]. Scoring model is better alternative for traditional model but model is not perfect-sometimes a bad application will receive high score and accepted, therefore, model is needed. There are various methods testing in past and in this literature that group into parameterized and non-parameterized methods. In this direction, [2]-[3] highlighted the importance of artificial neural network (ANN) and support vector machine (SVM) towards credit scoring as a better alternative of conventional approaches. Further from analysis point of view this literature is not focusing on ensemble methods or hybrid models' due to complexity of design. Although, admired the complexity of ensemble models and utilized for multiple tasks [4].

The non-parameterize model logistic regression is not perfectly suitable for the classification than parameterized models. These models are decision trees [5]-[7] and artificial neural networks [8]-[10]. Here, ultimate focus is to elaborate the importance of parameters of models with smoothing and boosting with the dataset imputation. This research will further interrogate the existing approaches to find the better smoothing and boosting model with respect to imputation technique.

## II. RELATED WORK

The work of author [11] has proven that area ratio (AUC) is better metric in accuracy than error rate compared with decision tree, logistic and ANN. It was concluded that ANN is better in performance in area ratio metric regardless of over consumed computation. Our approach further analyzed the existing work in depth to prove that boosting and smoothing of models worked well for these core approaches especially for ANN, which is better determiner of metric accuracy. That is the reason, why we contradict the existing approach because existing approach undertake the same dataset without treatment of missing values with imputation technique. Our work is continuation of the existing work to prove that neural network on similar dataset performed better when imputation technique applied over smoothing and boosting of models.

## III. OVERVIEW OF EXISTING WORK

### A. Decision Tree

Decision trees allow creating a tree-based classification model. Decision trees can graphically illustrate other choices that can be made and enable the decision maker is to identify the best situation in a circumstance. Common algorithms for decision tree induction include ID3, C4.5, CART, CHAID and QUEST [12]. In [13], author says that the decision rules should maximize a divergence measure of the difference in default risk between the two subsets. The splitting is repeated until no group can be split into two subgroups which are statistically different. According to [14], there are three major tasks of a classification tree: 1) how to partition the data at each step; 2) when to stop partitioning; and 3) how to predict the value of  $y$  for each  $x$  in partition.

### B. Logistic Regression

Logistic regression analysis is the multivariate technique, which allow to estimate the probability that an event occurs or

not, by predicting a binary dependent outcome from a set of independent variables. The logit model is a widely used statistical parametric model for modelling binary dependent variable. The logit model for credit scoring is presented with comparisons with other models including conventional one [15].

C. Artificial Neural Network (ANN)

ANN is an information processing model resembling connections structure in the synapses. It consists of many nodes (also called neurons or units) by links. The feed-forward neural network with back propagation (BP) is widely used for credit scoring, where the neurons receive signals from pre-layer and output them to the next layers without feedback.

According to [16], made a comparison of neural networks and linear scoring models in the credit union environment and the results indicated that neural network had better performance for correctly classifying bad loans than LR model. Besides, ANN need many training samples and long learning time. In [17], found that ANN has a higher accuracy rate by comparing with Logistic regression and discriminate analysis.

IV. PROPOSED SOLUTION AND DATASET

This section is discussing about boosting and smoothing criteria for each model with parameters. The purpose of the approach parametrization is to evaluate the different conventional models to improve the accuracy of classification. The dataset chosen from UCI website as shown below, it is about 2006 Taiwan faced credit flow crises thus result of repayments evidences are required. All the dependent and independent variables are given in Table 1. Later the models feed by opted strong and moderately correlated variables after data insight analysis technique, because, these variables are not necessarily correlated to each other thus prune from dataset eventually after correlational test. The categorical variables sex, education, marital status and age and continues variable Limit\_BAL eventually have weak correlation with other independent variables, therefore, not included as input to the models. The final set of 19 independent variables has given as input to the models from Table 1.

A. Artificial Neural Network (ANN)

According to [18], it has presented credit scoring by integrating back propagation neural networks with linear method. Linear model (1) and non-linear model (2) definitions are as follows:

$$y = h(x_i; w) = w^T x \tag{1}$$

$$a = \sigma \left( \sum_k \omega_{jk}^l a_k^{l-1} + b_j^l \right) \tag{2}$$

TABLE I. CREDIT SCORE RELETED VARIABLES

Variable	Description	Scale of variable
Limit_BAL	Amount of the given credit (NT dollar)	Continuous Interval
Sex	Gender (1 = male; 2 = female).	Categorical Nominal
Education	Education (1 = graduate school; 2 = university; 3 = high school; 4 = others).	Categorical Nominal
Marital Status	Marital status (1 = married; 2 = single; 3 = others).	Categorical Nominal
Age	Age (year).	Continues Interval
Pay_0 to Pay_6	April to September	Categorical
Bill_AMT1 to BILL_AMT6	Amount of bill statement (NT dollar)	Continues Interval
Pay_AMT1 to PAY_AMT6	Amount of previous payment (NT dollar)	Continues Interval
Y	Default payment (Yes = 1, No = 0)	Dichotomous

To find the best model, the gradual increment of hidden nodes with settings of 5 hidden followed by 2 hidden layers are involving better for the result accuracy. There are 19 input parameters of all three models, where, ANN-l is single layer default model, ANN-H model train with minimum hidden layers and ANN-L model is train with linear model without any activation function. The results are shown as in Table 2, where, ANN-L performed slightly better than all other models ANN-l and ANN-H with accuracy metrics error rate and area under curve (AUC).

B. Decision Tree (DT)

In [19] author says, the decision rules should maximize a divergence measure of the difference in default risk between the two subsets. The splitting is repeated until no group can be split into two subgroups which are statistically different. According to [20], there are three major tasks of a classification tree: (i) how to partition the data at each step, (ii) when to stop partitioning and (iii) how to predict the value of y for each x in partition. The decision tree [21] purpose is to find the optimal sub-tree that gives bad and good credit based on overall accuracy and error rate. This paper evaluated result using C4.5 classification, which formulate classification tree based on principle of entropy (1) and information gain principle (2).

$$E(S) = \sum_{i=1}^c -p \log_2 p_i \tag{1}$$

$$E(T, X) = \sum_{c \in X} P(c) E(c) \tag{2}$$

TABLE II. STATISTICS OF ANN

Model	Error Rate		AUC	
	Training	Testing	Training	Testing
ANN - 1	75.22	76.0	69.5	70.0
ANN -H	78.11	80.22	73.6	76.33
ANN -L	78.81	81.46	73.6	78.03

TABLE III. STATISTICS OF DT

Model	Error Rate		AUC	
	Training	Testing	Training	Testing
Default	79.75	78.48	75.89	73.25
Pruning	79.58	78.52	75.33	72.83
Boosting 10%	80.57	78.79	77.31	73.84
Boosting 100%	80.0	78.69	76.27	73.09

Table III: Default DT, Pruning DT, Boosting 10% , Boosting 100%

TABLE IV. STATISTICS OF LOGISTIC REGRESSION

Model	Error Rate		AUC	
	Training	Testing	Training	Testing
Default	77.76	81.22	73.4	77.91

Table IV: Regression Classification

The split of tree based on pure values evaluated by measures entropy and information gain. The result evaluated and compared among four model configurations as in Table 3. It was presented that boosting 10% of decision tree model is better in terms of accuracy than all other configuration such as booting 100%, pruning and default DT model.

C. Logistic Regression

According to [22], it was the first paper published investigates the logistic regression (LR) with discriminant analysis applied to credit scoring. Its results shows LR exhibiting higher accuracy rates, however, neither method was found to be sufficiently good to be cost effective for his problem. LR was also applied by [23] to a commercial loan evaluation process (exploring several models using random effects for bank branches).

$$P=(\alpha+\beta x) \quad \text{Eq. Logistic Regression}$$

In Table 4, logistic regression error rate metric considerably evaluated results better than AUC for bad and good credit. The area under curve (AUC) metric accuracy is 77.91, on the opposite, error rate metric test accuracy 81.22 is far better than AUC.

V. RESULT DISCUSSION

Researchers either consciously or by default in a statistical analysis drop the variables that have in-complete data. As an alternative to complete-case analysis, researchers may in a plausible value for the missing observations, such as using the mean of the observed cases on that variable [24]. But, here this research is focus on nearest distance based imputation technique. Besides, k-mean there are many statisticians recently advocated methods that are based on distributional models for the data (such as maximum likelihood and multiple imputation). More literature has been published in the statistical literature on missing data [25]-[27].

In [28], propose a new approach to clustering that divides the data features into observed features, which are known for all objects, and constraining features, which contain missing values. We generate a set of constraints based on the known values for the constraining feature. Based on our observation, we found high percentage of missing values in our dataset, therefore, we implemented similar technique of k-mean clustering for imputation to diminish the missing percentage in dataset to gain accuracy. We evaluated the result with all best models chosen after boosting and smoothing.

Fig. 1 below shows that k-mean imputation with ANN-linear model outperformed all other models in accuracy of error rate metric. Similarly, logistic regression slightly performed less ANN-Linear but, better than other model DT in error metric and AUC metric. Here, it was clearly notable that error metric is better metric for accuracy gain of the model in all comparisons. On the same note, we evaluated all models comparatively without imputation, which means contain more empty values.

No imputation results are shown in Fig. 2 below which comprehends the result for training dataset, it clearly reveals that DT shown significance over ANN. It was also evident that in test dataset ANN-Linear performed better than DT with no imputation technique; further below given table shows the test dataset comparison between models (Table 5).

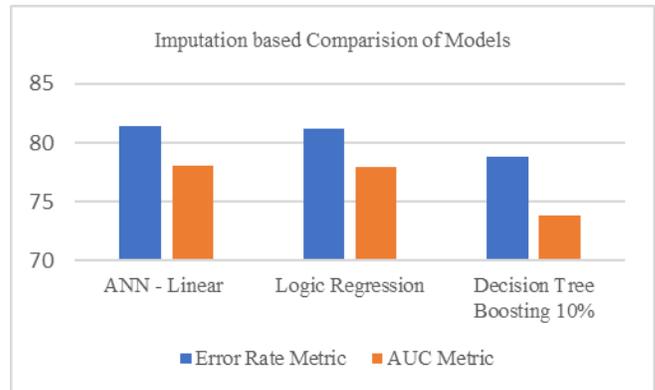


Fig. 1. The graph shows the comparisons between models with Error Rate Metric and AUC with imputation.

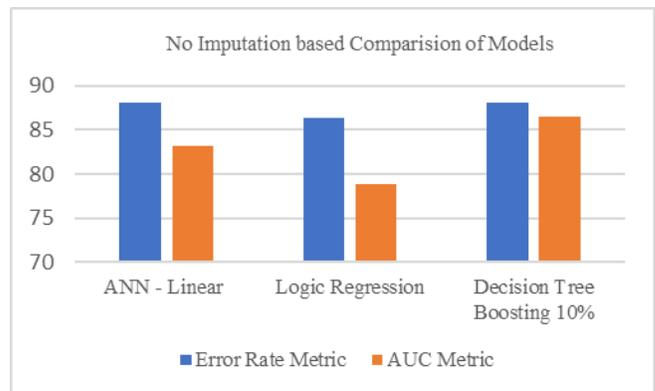


Fig. 2. The graph shows the comparisons between models with Error Rate Metric and AUC without Imputation.

TABLE V. NO IMPUTATION COMPARISONS

Model without Imputation	Error Rate		AUC	
	Training	Testing	Training	Testing
Decision Tree Boosting 10%	88.14	89.57	86.53	79.09
Logistic model	86.36	90.29	78.90	86.18
ANN-LINEAR	88.04	90.99	83.13	87.90

It was evident by all comparisons with imputation and without imputation technique that ANN performed better when data is in piece fully available for analysis into depth of model such as in case of imputation technique of K-mean. K-mean fill the values with nearest neighbor thus increase data availability that increases the accuracy rate of classification of ANN. But, in case of without imputation dataset lacks the volume in the form of missing values thus DT performed better than other model in training dataset but in test dataset ANN still performed better over low volume of dataset.

## VI. CONCLUSION

This paper examines the major classification techniques in data mining and compares the performance of classification. The novel imputation method of k-mean improvised to avoid the data loss, for the first time, is presented for the similar dataset and its comparison with no imputation technique. Obviously, error rate is more sensitive than AUC, because, it is more appropriate criterion to measure the classification accuracy of models. Artificial neural networks linear model performs classification accurately than the other models in comparison to imputation and without imputation. Artificial neural networks model is also shown the best performance in no imputation test dataset but performed second last in case of training dataset. It shown more accuracy in case of availability of data like in imputation based dataset accuracy is better than all models. From the perspective of risk control, estimating the client risks without imputation is more meaningful than imputation on classification. It was also concluded that, artificial neural networks model is more reliable to be employed for credit scoring for bad and good clients' awareness. In future, big scoring and its impact can be tested with larger dataset using ANN ensemble or hybrid approach to cater the multiple tasks 1) feature selection; 2) classification. Because conventional credit score techniques inherit with narrow scope that is not perfect model because it only analyzes customer payment history but unable to justify customer characters, nature and credibility by the help of external source for instance, social media.

## REFERENCES

[1] Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). *Credit scoring and its applications*. Society for Industrial and Applied Mathematics.

[2] Atiya, A. F., & Parlos, A. G. (2000). New results on recurrent network training: unifying the algorithms and accelerating convergence. *IEEE transactions on neural networks*, 11(3), 697-709.

[3] Bellotti, T., & Crook, J. (2009). Support vector machines for credit scoring and discovery of significant features. *Expert Systems with Applications*, 36, 3302-3308.

[4] Ala'raj, M., & Abbod, M. F. (2016). A new hybrid ensemble credit scoring model based on classifiers consensus system approach. *Expert Systems with Applications*, 64, 36-55.

[5] DAVIS, Rober Hunter, D. B. Edelman, and A. J. Gammerman. "Machine-learning algorithms for credit-card applications." *IMA Journal of Management Mathematics* 4.1 (1992): 43-51.

[6] Frydman, Halina, Edward I. Altman, and DUEN-LI KAO. "Introducing recursive partitioning for financial classification: the case of financial distress." *The Journal of Finance* 40.1 (1985): 269-291.

[7] Zhang, D., Zhou, X., Leung, S. C., & Zheng, J. (2010). Vertical bagging decision trees model for credit scoring. *Expert Systems with Applications*, 37(12), 7838-7843.

[8] Jensen, H. L. (1992). Using neural networks for credit scoring. *Managerial finance*, 18(6), 15-26.

[9] West, David. "Neural network credit scoring models." *Computers & Operations Research* 27.11 (2000): 1131-1152.

[10] West, D., Dellana, S., & Qian, J. (2005). Neural network ensemble strategies for financial decision applications. *Computers & operations research*, 32(10), 2543-2559.

[11] Yeh, I. C., & Lien, C. H. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, 36(2), 2473-2480.

[12] Rokach, Lior, and Oded Maimon. *Data mining with decision trees: theory and applications*. World scientific, 2014, pp. 168.

[13] Witzany, Jiří. *Credit risk management and modeling*. Praha: Oeconomica, 2010, pp. 46.

[14] Loh, Wei-Yin. "Classification and regression tree methods." *Encyclopedia of statistics in quality and reliability* (2008), pp. 315-323

[15] Huang, S. C., & Day, M. Y. (2013, August). A comparative study of data mining techniques for credit scoring in banking. In *Information Reuse and Integration (IRI), 2013 IEEE 14th International Conference on* (pp. 684-691). IEEE.

[16] V. S. Desai, J. N. Crook and G. A. Overstreet, "A Comparison of Neural Networks and Linear Scoring Models in the Credit Union Environment," *European Journal of Operational Research*, Vol. 95, No. 1, 1996, pp. 24-47. doi:10.1016/0377-2217(95)00246-4.

[17] H. Abdou, J. Pointon and A. E. Masry, "Neural Nets Versus Conventional Techniques in Credit Scoring in Egyptian Banking," *Expert Systems with Applications*, Vol. 35, No. 2, 2008, pp. 1275-1292. doi: 10.1016/j.eswa.2007.08.030.

[18] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5(3),1.

[19] Witzany, Jiří. *Credit risk management and modeling*. Praha: Oeconomica, 2010, pp. 46.

[20] Loh, Wei-Yin. "Classification and regression tree methods." *Encyclopedia of statistics in quality and reliability* (2008), pp. 315-323.

[21] Breiman, L., Friedman, J. H., & Olshen, R. A. (2009). *Stone, cj (1984) classification and regression trees*. Wadsworth, Belmont, California.

[22] Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, 15(03), 757-770.

[23] McFadden, D. and G. Leonard (1993). *Issues in the contingent valuation of environmental goods: Methodologies for data collection and analysis*. In *Contingent Valuation: a critical Assessment*, pp. 165-215. New York: North-Holland: Hausman.

[24] Pigott, T. D. (2001). A review of methods for missing data. *Educational research and evaluation*, 7(4), 353-383.

[25] Little, Roderick JA. "Regression with missing X's: a review." *Journal of the American Statistical Association* 87.420 (1992): 1227-1237.

[26] Little, Roderick JA, and Donald B. Rubin. "Statistical Analysis with Missing Data." (1987).

[27] Graham, John W., et al. "Analysis with missing data in prevention research." *The science of prevention: Methodological advances from alcohol and substance abuse research* 1 (1997): 325-366.

[28] Wagstaff, Kiri. "Clustering with missing values: No imputation required." *Classification, Clustering, and Data Mining Applications* (2004): 649-658.

# MobisenseCar: A Mobile Crowd-Based Architecture for Data Acquisition and Processing in Vehicle-Based Sensing

Lionel Nkenyereye  
Department of Computer Engineering  
Dong-Eui University  
Busan, Republic of Korea

Jong Wook Jang  
Department of Computer Engineering  
Dong-Eui University  
Busan, Republic of Korea

**Abstract**—The use of wireless technology via smartphone allows designing smartphone applications based on OBD-II for increasing environment sensing. However, uploading of vehicle's diagnostics data via car driver's tethered smart phone attests a long Internet latency when a large number of concurrent users use the remote mobile crowdsensing server application simultaneously, which increases the communication cost. The large volume of data would also challenge the traditional data processing framework. This paper studies design functionalities of mobile crowdsensing architecture applied to vehicle-based sensing for handling a huge amount of sensor data collected by those vehicle-based sensors equipped with a smart device connected to the OBD-II interface. The proposed MobiSenseCar uses Node.js, a web server architecture based on single-thread event loop approach and Apache Hive platform responsible for analyzing vehicle's engine data. The Node.js is 40% faster than the traditional web server side features thread-based approach. Experiment results show that MapReduce algorithm is highly scalable and optimized for distributed computing. With this mobile crowdsensing architecture it was possible to monitor information of car's diagnostic system condition in real time, improving driving ability and protect the environment by reducing vehicle emissions.

**Keywords**—Mobile crowdsensing; data processing; web services; hadoop; hiveQL; OBD-II

## I. INTRODUCTION

Mobile crowdsensing refers to a new paradigm that allows a certain number of individuals to collectively share data and extract information so as to measure and map phenomena of common interests [1]. Improvements in terms of smart device capabilities and communication technologies allow crowdsensing solutions to emerge as significant strategies to revolutionize environment sensing. If vehicular mobility is adopted, the sensing capabilities can be further increased by connecting smart devices to vehicles using the On Board Diagnostic interface (OBD-II) or provided directly by smart vehicles through vehicle-to-vehicle or vehicle-to-infrastructure communications [2].

Remote On-line Vehicle Diagnostics (ROVD) is such a telematics service that provides opportunities to constantly monitor the vehicle diagnostics system remotely [3]. ROVD integrates the capability of computing unit to identify a fault or a possibility of a fault in an automobile and transmit vehicle

trouble codes to a remote central processing center or public cloud computing [3]. Thus, ROVD replaces wireless technology over short distances or cable between on-board vehicle's connector and vehicle's monitoring system by mobile network based data [4]. These vehicle monitoring systems move into a remote data center owned by car manufacturers for instance. ROVD also adds the event-driven intelligence in which the vehicle determines when it is necessary to inform changes in the vehicle's status [4].

The number of mobile crowdsensing vehicle's diagnostics applications (MCVDAs) has increased in the automotive market [3]. These MCVDAs ensure that the messages received from the vehicle's engine reach the smartphone through the Bluetooth wireless interface and the smartphone's wireless communication modules transmit the received packets to the remote data center. Thus, these applications may play an important role in reducing the number of accidents and allowing the automotive industry to use data generated from an automotive electronic system to make new models of cars in the future. Another advantage is that it allows Original Equipment Manufacturer (OEM) or its service partner to inform the driver whether he can continue his journey safely or whether he requires assistance from a near repair shop. This helps prevent vehicle breakdown by detecting the vehicle problems at an early stage.

The transmission of vehicle's diagnostics data via car driver's tethered smart phone increases the cost in communication by looking at the amount of generated data in order of a few gigabytes per hour per vehicle [5]. As a consequence, the vehicle commercial company or car manufacturers and automobile ecosystem would have to support the lifetime cost of diagnostics data. It is not reasonable for the car manufactures or OEM to support communication cost when they cannot control the usage of the mobile application delivered to the car owner. The data transmission may not only increase cost communication but also internet latency when the architecture of the remote web server may not process concurrent requests from a large number of drivers uploading their vehicle's diagnostics data simultaneously. For example, the driver would interrupt the uploading of a vehicle's engine diagnostics data when he (she) uses other applications installed on his (her) smart phone due to the unexpected communication latency. In this case, the OEM

or its service partner would not be able to start a diagnosis session remotely, and therefore, leave the driver with possibility of not receiving any assistance. Furthermore, the enormous quantities of vehicle's diagnosis data generated by the MCVDA may challenge its analysis making it difficult for the car manufacturers' technicians to figure out the cause of vehicle breakdown quickly, for instance.

The connectivity inside the vehicle may be established. However, due to the mobility of vehicles and the lack of a reliable connection, a real-time uploading would be unsuccessful. For this, the pushing technique from a temporary storage should be used on the mobile application to the remote server for realizing data replication after a certain distance of driving.

This study proposed a mobile crowdsensing architecture for designing MCVDA that relies on the tethered connectivity model, component-based mobile applications, Client-Server Architecture and data processing framework for processing volume of both structured and unstructured data. Therefore, the tethered connectivity model carries the smart phone inside the vehicle. The smart phone is used as a modem via wired, Bluetooth or Wi-Fi as communications technologies. The component-based mobile application and Client-Server Architecture feature have the ability to process a large number of requests from multiple drivers simultaneously. The proposed Client-Server Architecture prevents blocking and long execution requests that may increase communication cost. To solve the limitation of blocking running requests, the Client-Server architecture handles them using a single-thread event loop web server-side called Node.js [6], [7]. The data processing model is implemented using Big data technology which provides infrastructure that can manage and process vehicle-based sensing data, thus enabling to enhance safety and driving experience. The implantation of Big data includes Hadoop, an open source distributed platform for storing and processing data. This Hadoop includes the Hadoop Distributed File System (HDFS) which provides Application Program Interfaces (APIs) for MapReduce applications to read and write data in a parallel manner [8].

The main contribution of this work is as follows:

- Design components based mobile client server computing architecture is introduced for implementing the MCVDA application called MobiSenseCar which is based on Android. The design of the MobiSenseCar, a mobile application allows the managing of the issue of concurrent clients' requests and replication of vehicle's diagnostic data on a temporary SQLite database on a mobile application to solve the issue of unreliable connectivity.
- A novel analytical model is based on big data technology for monitoring vehicle's diagnostics data. In-vehicle's diagnostic data is stored in HDFS; analyzing jobs are executed by queries in HiveQL language, and the queries will be transferred into MapReduce progress.
- Investigate the performance of data acquisition platform which consists of a Client-Server Architecture that

includes backend server, web programming framework and database that can support a large and increasing number of concurrent clients' requests. The performance metrics are throughput, response time and error rate to compare web applications developed using JavaScript and JavaServlet.

The rest of this paper is organized as follows. First, the related work is discussed in Section II, then focus on the background of technology of event-driven approach, Node.js and Big Data technology in Section III. This section elaborates a general system design model for MobiSenseCar based on web server with Node.js and Hadoop. Next, the paper takes an isolated look at different stages of this model and gets to know their relevant components approaches to adopt in Section IV. In Section V, it discusses the results of the implementation of a MobiSenseCar mobile application based on Android that features different components of the proposed mobile client server computing architecture. In Section VI, it briefs the simulation, experimentation and analysis of results of the proposed mobile crowdsensing architecture. The conclusion in Section VII sums up the architecture for a diagnosis of the status of a vehicle's engine and the advantages of deploying applications based on tethered connectivity model for leveraging heterogeneous crowdsourced data from MCVDA.

## II. RELATED WORK

Currently there are few ad hoc solutions to smart phone based sensing vehicle monitoring systems. Therefore, prototype model includes on-board computer, wireless communication link, vehicle monitor server, and vehicle status browser. Below the paper proceeds by describing briefly the different smart phone based mobile crowdsensing architecture for monitoring the car as a sensing platform.

Prashanth et al. [9] made an architecture that leverages sensors besides GPS-accelerometer and microphone, in particular to glean rich information such as the quality of the road or the noisiness or traffic. They proposed algorithms to virtually reorient a disoriented accelerometer along a canonical set of axes and then use simple threshold-based heuristics to detect bumps and potholes and braking. Their study focus on the use of sensing components in the smart phone applied to proposed algorithms for detecting potholes. A similar application is PotHole [10] which can identify holes in streets using the crowdsourced vibration and position data collected from smartphone.

Recently, both Derick et al. [11], Eren et al. [12] and Mohamed et al. [13] propose a driving style recognition application using smartphone as a sensor platform. The evaluation in [11] proves that classical Dynamic Time Warping algorithm can accurately detect events with a very limited training set. In [12], safe or unsafe optimal path detection algorithm and Bayesian classification applied to vehicle data detect the driver behavior, and then increase safety while driving.

Jules et al. [14] present a formal model for accident detection that combines sensors and context data. They showed how smartphone sensors, network connections, and web services can be used to provide situational awareness to first

responders. The contribution of their work provides empirical results demonstrating the efficacy of different approaches employed by smartphone accident detection systems to prevent false positives.

There is another interesting work which proposes driver behavior profiling using smartphone [15]. This work analyzed how smartphone sensors can be used for identifying maneuvers and propose a platform that is able to detect risky driving events independently from the mobile device and vehicle. In this work, the fuzzy logic event detection mechanism is implemented in an Android application. The authors state that the approach intends to use the DoIP protocol to perform vehicle diagnostics data exchange synchronously over a TCP connection over wireless communication and information infrastructures.

The design prototype proposed in the above studies has some drawbacks. The first drawback is that the design prototype proposed in [10], [11], [14], [15] does not take into consideration the impact of a large amount of generated data. How and where to preprocess and aggregate Controller Area Network (CAN) data is also an important question which was not addressed. Furthermore, there is no performance evaluation of Client-Server Architecture designed to handle a large number of requests submitted from participatory vehicle – based sensing platform. The second drawback is about the integration of the data processing framework for analyzing a huge amount of vehicle’s diagnostics data generated by MCVDA that would enhance decision making.

At one of the range, there is rich commercial vehicle’s telematics application as mentioned in the white paper published by Oracle. In that white paper, the Oracle for the Connected Vehicle highlights how their Octo Telematics solution architecture would turn data generated by vehicles into business, therefore enhancing the value of vehicle’s diagnostics data [16]. Thus, the proposed Octo telematics solution relies on the embedded “clear box” of sensors used to collect on-board vehicle’s data, and then uploads them to the central message system where a mechanism of validation meshes unto the insurance company’s framework.

### III. BACKGROUND OF TECHNOLOGY

The purpose of this section is to provide an introduction and background to areas immediately tied to this work. It describes the characteristics of some important fields that should be taken into consideration when conducting this kind of studies. First, it describes the characteristics of the connectivity models inside vehicles that might provide a feasible integration of solution based on them. Secondly, it provides the reader with a short introduction to server-side scripting to respond to network and concurrent requests. Finally, it describes the big data technology and its related framework to process large data sets of information.

#### A. Communication Solution Inside Vehicles

The wireless communication technologies built-in or brought in the vehicle will enable the automotive world to

afford new applications such as navigation, billing facilities, and fleet management. The success of new in-car telematics applications and services can be realizable through the V2I (Vehicle-to Infrastructure) data exchange with network operators [17]. The connectivity inside the vehicle may be established by the network operator in three ways: embedded solution, tethered solution and integrated solution [17].

- The embedded solution: it includes both connectivity and intelligence that are built into the vehicle. For this kind of connectivity, the in-car telematics services proposed by the auto makers will force them to build TCU (Telematics Control Units) into the vehicle within a SIM (Subscriber Identity Module) card to connect to the network in order to make calls, and receive texts for instance.
- The tethered solution: The tethered connectivity model stands on the obligation of carrying the smart phone inside the vehicle. This smart phone is used as a modem via wires, Bluetooth or Wi-Fi.
- The integrated solution: integrated solution is considered as an integration of the smart phone application into the vehicle to enable the driver to access telematics or cloud-based applications and their features safely while driving.

The research work done on the growth of connectivity model inside the vehicle and published by the World Leading Knowledge Partners to the Automotive industry [17] states that tethered solutions will grow more and eventually peak in developed USA and Europe market by the end of the decade [17]. As embeddedness is considered to be the seamless and reliable solution for the future in-car telematics service, USA and Europe markets are interested in tethered solution for the only reason that tethered solutions are considered as short-term solutions to customers’ unwillingness to pay a second communication cost for vehicle-related connectivity in the car on the top of their mobile phone communication bill [17].

#### B. Server-side Scripting Approach

The architecture of web server-side and its scripting approach must inherit features that allow responding to an increasing number of network requests from the end-users. Thread-based scripting approach has been used to implement web application to respond to the clients’ requests. However, web server based on the thread-based approach might perform inefficiently as the number of incoming network requests increases. Therefore, many industries such as eBay, LinkedIn, etc. have started to adopt event-driven programming as an option to respond to a large number of concurrent requests and achieve scalability more operationally.<sup>1</sup> In this sub section, it first describes the thread-based scripting approach and its limitations. As an option to overcome limitations of thread-based scripting, it discusses the key features of the event-driven scripting approach and the benefits of the recent server-side platform based on event-driven scripting called Node.js.

<sup>1</sup> <https://www.codefactoryacademy.com/posts>

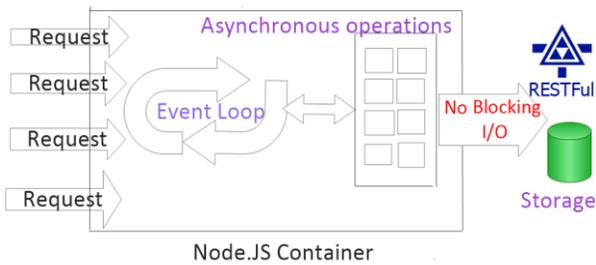


Fig. 1. A Conceptual Model for an event-driven architecture. Each incoming client request is handled by the single-thread event loop. Event handlers do trigger I/O actions that result in a new event later asynchronously.

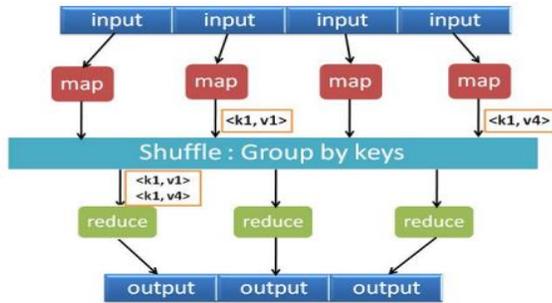


Fig. 2. Operation phases in Map/Reduce programming mode [22].

Thread based approach identifies each incoming request or tasks using a separate – thread. This thread-based architecture relies on the per-connection process model where a dedicated process for handling a connection is setup [18]. The thread descriptor is shared among all processes and each process stops up for a new connection, treats the connection and then delays for the next connection [18]. As a consequence, at a given time, the server finds itself in situation having the same number of threads as the number of requests [19]. Therefore, the application server would not scale efficiently when there are many threads or network requests [19].

Event-driven scripting approach attests as an option to synchronous blocking I/O (Input/Output). As shown in Fig. 1, the event-driven approach queues both new requests and blocking I/O requests. The single-thread executes an event loop by setting up a simple mapping of all requests. The event loop gradually dequeuing request from the queue, then processing the request, finally taking the next request or waiting for new requests submitted. The event-driven script is referred to asynchronous programming. This means that the statements inside the scripts are not necessarily executed in the order of being written in the code. Usually, no single statement will ever block the next line of code. In fact, even if the next line statement takes a long time to complete, the rest of the program will continue to run normally. At that time, the program will wait for some resource to complete its long-running tasks and when it is done, a callback function is called. For instance, in the case of a server-side web application, this paradigm allows for handling enormous load capabilities because it does not need to wait for a long running request to finish. Instead, the server can start beginning performing the next request and return to finish the previous request when callback result occurs. The server is never blocked, so it is suitable to handle a high number of concurrent requests.

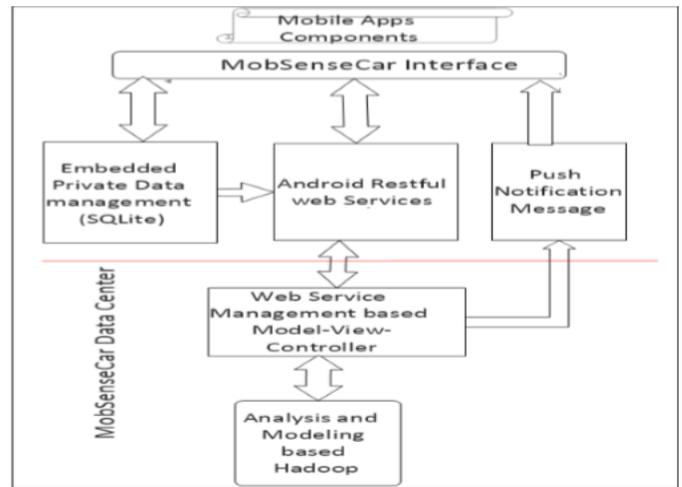


Fig. 3. Overview of the functional components-based Mobile Crowd Sensing and computing architecture for MobiSenseCar application. The components are divided in two type. The first type of components are designed to be implement on the Smartphone The second type of components is to manage the MobiSenseCar Data center-Wired SOA infrastructure.

One of the advantages of Node.js over thread-based framework is that it has a built-in single-thread event loop and non-blocking model [20]. The second advantage is that Node.js allows event-driven paradigm which is the key on which interactive Node.js applications are constructed. Node.js features the event-handler that creates events and the main loop executes the appropriate event. The event handlers in Node.js are known as callback functions. Therefore, callback functions are eventually executed on completion of the non-blocking operation. So, when the event loop in Node.js receives the completion feedback, it executes the callbacks.

### C. Apache Hadoop for Big Data

Yahoo, Google, and Facebook have extended their services to web-significance due to the amount of data collected on a daily basis. The data collected on line have overpowered the capabilities of the traditional IT architecture [21]. In order to extract the valuable data for decision making, they published open access papers and released code for core infrastructure responsible for distributed storage and processing into open source. Among these components, Apache Hadoop [22] has rapidly emerged on the top of components capable of aggregating, transforming, and analyzing server logs and a large volume of unstructured data.

Apache Hadoop is an open source distributed software platform for storing and processing data. It is written in Java, and runs on a cluster of industry-standard servers configured with direct-attached storage [21]. The structure and principal components of Apache Hadoop framework are described in details in the book “Hadoop The Definitive Guide by Tom White(foreword by Doug Cutting)” [23]. The distributed processing framework known as MapReduce (Fig. 2) is central to the scalability of Apache Hive. MapReduce helps programmers solve data-parallel problems. It splits the input data-set into multiple chunks, each of which is assigned a map task that can process the data in parallel. Map tasks functions read the input as a set of (key, value) pairs and produces a set of (key, value) pairs as result. The MapReduce framework

shuffles and sorts outputs of the map tasks, forwarding the intermediate (key, value) pairs to the reduce tasks, which joins together and produces final results. MapReduce uses JobTracker and TaskTracker mechanisms to schedule tasks, monitor them, and restart any task that fails [21].

The Apache Hadoop platform also includes the HDFS (Hadoop Distributed File System). The HDFS is designed for scalability and fault tolerance. HDFS stores large files by splitting them into blocks (64MB). Beside MapReduce and HDFS, Apache Hadoop also includes many other components, some of them are very useful for analysis and modeling data: 1) Apache Flume<sup>2</sup> is a distributed system for collecting, aggregating, and moving large amounts of data from multiple sources into HDFS or another central data store; 2) Apache Sqoop [24], is a tool for transferring data between Hadoop and relational databases. You can use Sqoop to import data from a MySQL into HDFS, run MapReduce on the data, and export them back into a Relational Database Management System (RDBMS); 3) Apache Hive [24] and Apache Pig [24] hold programming languages that make simpler development of applications using the MapReduce framework. HiveQL is a dialect of Structured Query Language (SQL) that supports a subset of SQL as query syntax. Although slower in running, HiveQL scripts are being actively enhanced for low-latency queries on Apache HBase [24] and HDFS. In contrast, Pig Latin is a procedural programming language that provides high-level abstractions for MapReduce. Open Database Connectivity/Java Database Connectivity (ODBC/JDBC)<sup>3</sup> Connectors for HBase and Hive are proprietary components expected in distributions for Apache Hadoop software. They provide connectivity with SQL applications by translating traditional SQL queries into HiveQL commands that run on the data set in HDFS or HBase.

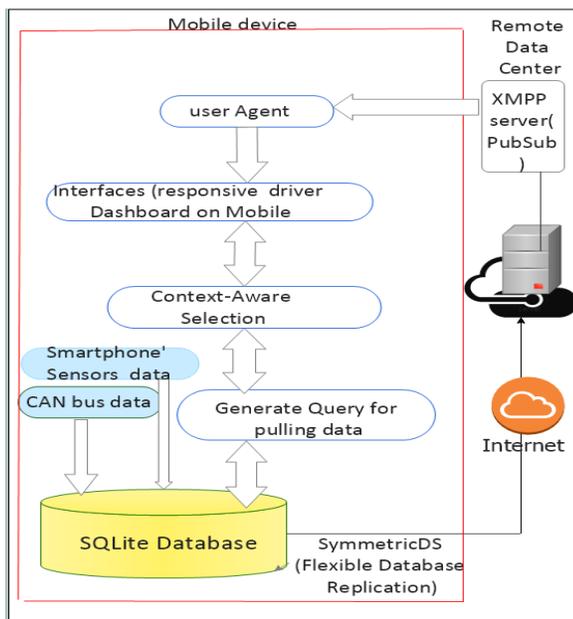


Fig. 4. Application Architecture delineating the various processing modules and interaction flow between them.

#### IV. DESIGN OF MOBISENSECAR APPLICATION FOR VEHICLE DATA ACQUISITION AND PROCESSING IN VEHICLE-BASED SENSING

This section covers the key functional components of designing functional components for vehicle data acquisition and processing in vehicle-based sensing. It outlines the models for building and deploying MobiSenseCar on a smart phone within Android.

##### A. Tethered Solution for the Proposed MobiSenseCar

The design is based on the choice of database, web server architecture, the manner in which vehicle's diagnosis data are transmitted to the remote data center through a service based on Representation State Transfer (REST) called RESTful framework, and the mechanisms used for event-driven approach to handle multiple Hypertext Transfer Protocol(HTTP) requests from the MobiSenseCar application.

The tethering solution reinforce the implementation of MobiSenseCar based Mobile Crowd Sensing and Computing (MCSC) [25]. This means that the driver can access the Internet anywhere, connect to the remote data center based cloud computation and services via the internet over wireless communication technology. In this study, the tethered solution based on MCSC includes several approaches that provide the essential platforms for allowing monitoring of car's diagnostic system, thus enhancing driving convenience. The tether application on the driver's smart phone will work through normal communication cost subscription SIM to transmit vehicle's diagnosis data to the remote diagnosis data center via wireless communication technology. In order to prevent continuous transmission of vehicle's data continuously to the MobiSenseCar data center, the MobiSenseCar application would have the functionality to store the engine's status data into embedded application's SQLite database on Android [26].

The architecture of the MobiSenseCar's web service server is event-driven and can improve scalability for efficient handling of several requests simultaneously. The aim of adopting event-driven server-side architecture is to prevent blocking and long running requests that may increase cost. Certainly, the adoption of event-driven web architecture to build web-based mobile application would permit car drivers to safely open and access existing mobile apps on their smartphone while the MobiSenseCar tasks are running in background such as uploading of vehicle's diagnosis system data stored in the embedded application's SQLite database to the MobiSenseCar data center.

##### B. Overview of Functional Components for MobiSenseCar

The functional components of the architecture proposed in this study are developed around a set of high-level functional areas common to most developers of mobile application based MCSC. Fig. 3 shows the overview of the functional components-based mobile client server computing architecture for MobiSenseCar. The overview of Fig. 3 is explained as follows. First, the MobiSenseCar interface enables registration of car owner's profile and the unique identification of the OBD scan tool<sup>4</sup> and several services in order to enable the real

<sup>2</sup> Apache Flume, <https://flume.apache.org>

<sup>3</sup> Sqoop connector, <https://sqoop.apache.org>

<sup>4</sup> ScanToolnet, <https://www.scantool.net/>

monitoring of the status of his (her) vehicle's engine. Second, Android RESTful web service routes the collection of vehicle's diagnosis data to the MobiSenseCar data center.

The Push Notification message handles warning notifications about OBD-II (check Engine Light) Trouble Codes. Third, the SQLite temporary stores the vehicle data while the car is driving in order to prevent the real-time transmission of this data. Many mobile applications rely on distributed key-value stores like SQLite for low latency access to data [27]. However, this SQLite database has the advantage of storage, consultation of SQLite's tables of vehicle's diagnostics data, GPS location, and notification message tables as fast as possible. Thus, drivers cannot be adversely impacted by the execution of other smart phone applications. In addition, the MobiSenseCar mobile apps has service that allows the car driver to use a non-persistent store for recording information related to its vehicle's diagnosis data such as SD memory card. This non-persistent store offers later monitoring against the corresponding fault of the vehicle's engine.

The SQLite's warning notification table maintains a mapping of unique identification-specific Id of mobile device that is used to identify the car driver for whom the warning message is concerned. The warning notification message is directly delivered using identification-specific Id of the mobile device. The GPS location table is essential when an urgent intervention is required in order to help drivers who face breakdown of their car. The collection of vehicle's diagnosis, warning message and GPS locations can be useful when evaluating the performance of vehicles sold, refining targeting, and adjusting deployment decisions on making new cars. Therefore, vehicle's diagnostics data are left entirely transient in the key-value store or kept more permanently in a movable hard disk. The decision is largely reflective to the car owner's tolerance for data loss.

On the MobiSenseCar data center-wired SOA infrastructure, Node.js web server based on the event-driven architecture interacts directly with NoSQL database (MongoDB)<sup>5</sup> through the framework designed for inserting and retrieving data. Hence, the effectiveness of fault detection algorithms run using MapReduce framework make it possible to analyze and process the huge volume of vehicle's diagnostics data. The storage module consists of NoSQL database such as MongoDB that avoids traditional table-based relational database structure. The particular suitability of a NoSQL database depends on the problem it is designed to manage. For big data and real-time web applications, the data structures used by NoSQL databases are also viewed as "more flexible" than the relational database tables.<sup>6</sup>

### C. Processing Modules for MobiSenseCar

The system design focus on the implementation of a MobiSenseCar Application using OBD-II scan tool, mobile device wireless communication technology, and OEM wireless SOA infrastructure. This OEM wireless SOA infrastructure integrates web service management, and big data analytics infrastructures. Fig. 4 shows the application architecture

delineating the various processing modules. Functional components of MobiSenseCar explained in the previous section constitute the basis of the proposed system design of MobiSenseCar. The user agent is responsible for controlling the whole application state and integrating push notification using XMPP server. The user agent authorizes the delivery of notification. The user agent initializes the configured smartphone's sensors. Each sensor is executed in a separate thread. Resource allocated to each sensor communicates with the corresponding sensors. The engine information and diagnostics troubles codes (CAN bus data) are temporarily stored in the MobiSenseCar application database (an android SQLite database). The context-aware selection module implemented the dynamic context-aware selection of which sensor data to be discarded. Based on the parameters from the context-aware selection module, a query is formed through it, the matched data stored on the SQLite database are replicated to the remote MobiSenseCar data center through SymmetricDS.<sup>7</sup> For example, when the car driver connects to the internet or activates internet connection through the smartphone's data plan, he (she) in turn transmits vehicle's engine data to the MobiSenseCar data center using SymmetricDS. Therefore, in order to monitoring the car's diagnostics system continuously and seamlessly, a web server-side solution based on Node.js enables the availability of vehicle's engine data for sharing purposes. The Node.js is responsible for handling data transmitted to a NoSQL database (MongoDB database).

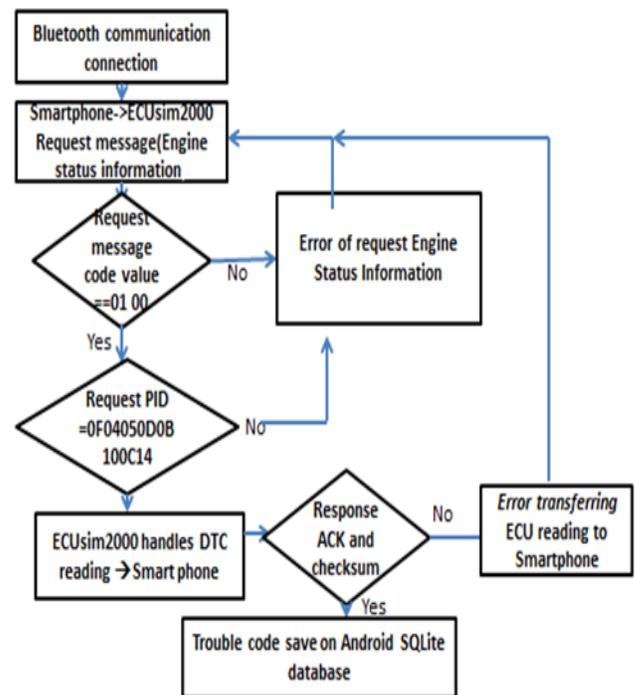


Fig. 5. Flowchart of engine status information collection algorithm.

<sup>5</sup> <https://github.com/mongodb/mongo-hadoop/releases>

<sup>6</sup> <http://www.allthingsdistributed.com/2012/01/amazon-dynamodb.html>

<sup>7</sup> [https://www.itcentralstation.com/comparisons/ibm-infosphere-database\\_vs\\_jumpmind-symmetricds](https://www.itcentralstation.com/comparisons/ibm-infosphere-database_vs_jumpmind-symmetricds)

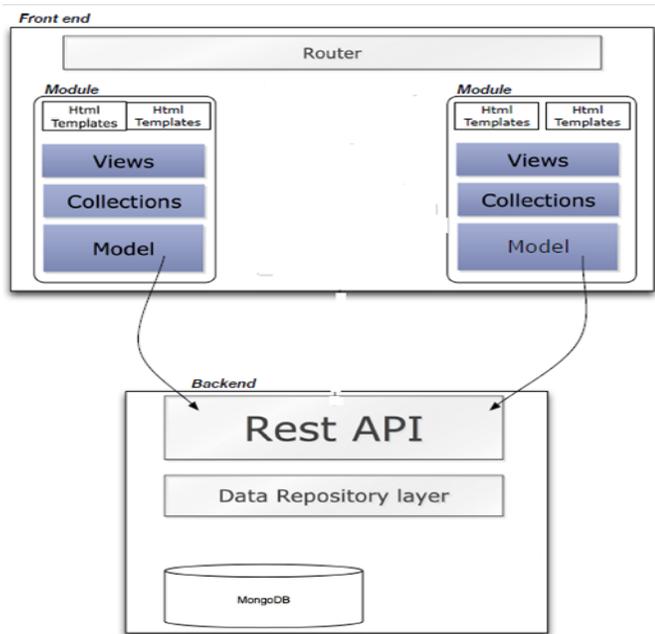


Fig. 6. Main Software Components of handling vehicle-based sensing data on the MobiSenseCar data center.

#### D. Bluetooth OBD-II Protocol Structure for Engine Status Information Collection

The system design of the MobiSenseCar service allows the driver confirming the status of engine in real-time via his (her) mobile device or web browser using a portal web site designed to retrieve the current vehicle's status information. Drivers may always check the status of the engine with a smart phone. For business purpose with the car manufacturer and vehicle commercial companies, the driver can transfer the current vehicle's status information to the MobiSenseCar data center. The engine status information collection algorithm is shown in Fig. 5.

In this study, it used an On-Board Diagnostic simulator called OBD-II simulator (ECUsim2000).<sup>8</sup> With this ECUsim2000 simulator, the communication test is conducted in the same way as an actual car was developed and tested. The hardware architecture includes the ECUsim 2000 OBD-II ECU simulator for reading the performance using OBD-PIDs code make up in the android application such as speed, revolutions per minute (RPM), intake temperature, coolant temperature. As shown in Fig. 5, the proposed system architecture includes Bluetooth communication between the ECU simulator (ECUsim 2000) and the Bluetooth interface that supports all OBD-II protocols and receiver (mobile) devices. The ECUsim2000 is designed to ensure integrity of supporting all OBD-II protocol in order to read OBD-II parameters clearly. Hence, the flow of data acquisition is followed by receiving byte for OBD-PID request when the message address matches. Subsequently, the value of request OBD-II parameter is transmitted via Bluetooth communication. At the receiver side, a smart phone plots the received engine's diagnostics data. The different value of OBD-II parameters are calculated in a human readable form and displayed on the MobiSenseCar mobile

application console designed for displaying current OBD-II parameters read out. The OBD-II message format consists of 1-byte priority, target address, source address header, 7 byte data, and checksum. It is basically used as a protocol for SAE-J1850 and ISO [28]. The CAN OBD message format consists of ID bits (11 or 29), DLC, 7 data bytes, and checksum<sup>9</sup> (CRC-15 processing method).

#### E. Deployment of Web-Side Based Node.js and MongoDB Database for Handling Vehicle-Based Sensing Data

The Node.js presents several advantages in terms of processing multiple connections or tasks concurrently. In our study, Node.js has influenced our choice of the web server architecture. Therefore, the MobiSenseCar application consists of collecting the vehicle sensing and transferring them to the remote MobiSenseCar center for further processing. The car's diagnosis system data need to be stored or hosted on a computer which is connected to the internet known to us as a Web Server. It serves to handle incoming requests from the web browsers (clients), and then responds by sending the required data through a web server program. As shown in Fig. 6, the router is the component that organizes routing between the MobiSenseCar's main pages. The router is configured to listen to every event so that when such an event is triggered, the router is notified which in effect tells to issue the URL request to the server. The models implement both business logic and data attributes.

Considering that the application has several collections of models, for example a list of current CAN bus data on the vehicle sensing data page, multiple rows of CAN bus data in the vehicle sensing data page, etc. It makes sense to encapsulate these models in separate modules (collections). This way, a collection is a holder for multiple coherent models. The motivation for this is that the collections can also work as active records, in that they can be responsible for fetching and maintaining a particular set of vehicle sensing data or CAN bus data from the database, regarding the collection of models they control. The view handles all the user-interactions that happen inside the HTML, it represents. The view is responsible for rendering its HTML templates.

Let now see how Node.js handles blocking I/O requests since MobiSenseCar application involves an important number of concurrent requests. These requests consist of transferring car sensing data such as vehicle's diagnostic data to the MobiSenseCar center. The Node.js event model uses event handler. This event handler comes on with great results until you run into the development of functions that involve blocking I/O. The blocking I/O is defined as a request that stops the execution of the current thread and delays for a response before continuing. With Node.js, the event model work is scheduled in advance as a function with a callback to the event queue.

MongoDB is a technology that is revolutionizing database usage. Together, the two tools (Node.js and MongoDB) are a powerful combination to the fact that they both employ JavaScript and JSON. We will need to install Mongoose, which is the library that Node.js uses to communicate with

<sup>8</sup> <https://www.scantool.net/ecusim-2000.html>

<sup>9</sup> [http://smartdata.usbid.com/datasheets/usbid/2000/2000-q4/j1850\\_wp.pdf](http://smartdata.usbid.com/datasheets/usbid/2000/2000-q4/j1850_wp.pdf).

MongoDB. The web service functions on Node.js would collect data from the vehicle when the OBD connector is paired with the MobiSenseCar application based on Android platform. Thus, through 3 or 4G LTE and HTTP protocols, data are transferred over the internet in JSON exchange format for easily processing on the web server with Node.

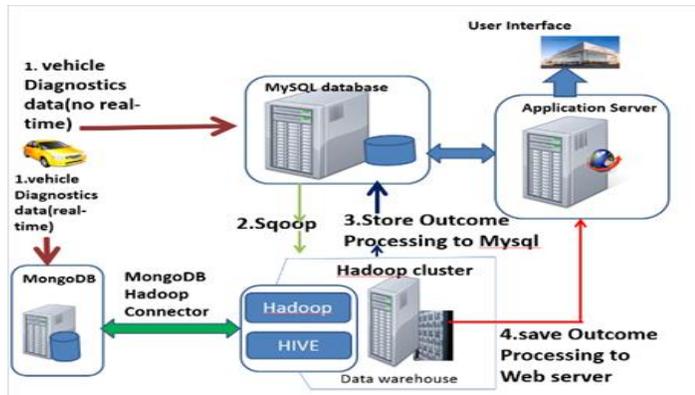


Fig. 7. Architecture and components of analysis model Based Big data analytics.

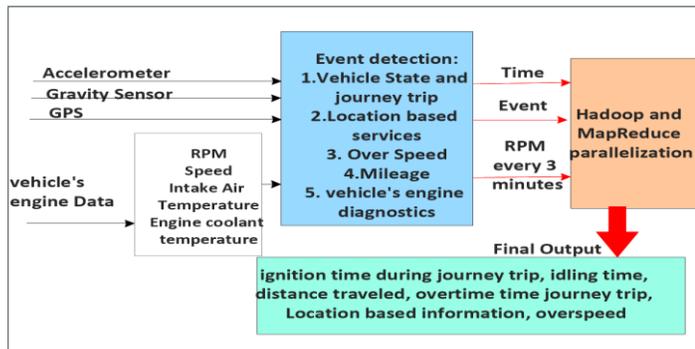


Fig. 8. Design of processing model for Monitoring and Analysing Vehicle' data based sensing using Hadoop and MapReduce parallelization.

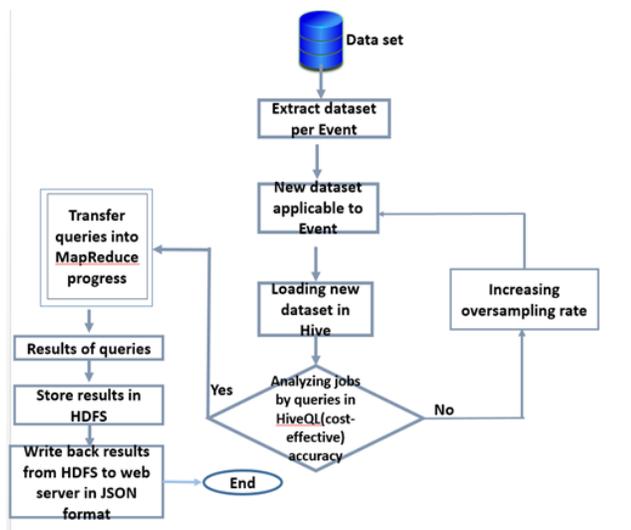


Fig. 9. Flow chart of the procedure followed during the implementation of the processing model of vehicle-based sensing data.

### F. Formal Representation of the Proposed Analysis Model Big Data and its Architecture

The proposed analysis model using Hadoop framework is based on a data-driven approach. This approach consists of collecting data sets uploaded from vehicles. Those data are then monitored based on different aspects of activity of the vehicles that we quote as “Events”. The first event relates the vehicle’s movement and journey trip. The second event is the collection of vehicle’s diagnostic data while the driver is driving.

The processing of the data acquired from the MobiSenseCar is divided into four phases as shown in Fig. 7:

- 1) Import data from MySQL to Hadoop clusters.
- 2) Loading data from HDFS to Apache Hive.
- 3) Analysis through Hadoop MapReduce framework.
- 4) Upload outcome files in JSON format from HDFS to the web server.

Firstly, the import data from MySQL to HDFS consists of importing data from MySQL into Sqoop. Sqoop is an open-source tool that allows users to extract data from a relational database into Hadoop for further processing [24]. The MongoDB Connector for Hadoop provides the ability to use MongoDB as input and/or an output destination [29]. Secondly, the loading data from HDFS to Hive is performed by Apache Sqoop. It keeps parallelizing import across multiple mappers. The import data into Apache Hive relies on the sake of efficiency that has a post processing step where Hive table is created and loaded. When the data is loaded into HIVE from HDFS directory, Hive moves the Sqoop replication table which is viewed as a directory into its warehouse rather than just copying data. Thirdly, Hive queries feature join patterns algorithms are implemented in the MapReduce jobs to execute SQL applications and queries [29] and finally, the Apache Sqoop writes back the output results to MySQL Database or MongoDB.

The description model of the proposed analytics framework associates for both events includes the vehicle’s movement, journey trip, location based service, over speed, mileage and diagnostics of vehicle’s engine events which is an appropriate subset of information. The subset of information are journey data, Global Positioning System (GPS) data, driver behavior data based on the smartphone’s in-built accelerometer, engine data and car diagnostics data. As shown in Fig. 8 for instance, the analysis process takes the vehicle’s movement and journey trip event, and then associates the RPM value of the vehicle to detect if the engine is running, current data, and accelerometer data to detect vehicle’s movement. For this event, the analysis process extracts the value of the RPM every three (3) minutes to detect the state of the vehicle, which is either in idle state or not. These data are then uploaded to HDFS. The data set on the HDFS serve as the basis for collecting useful information to submit to MapReduce functions for processing.

The Hadoop framework splits the input data-set into multiple chunks, each of which is assigned a map task that can process the data in parallel. Each map task reads the input as a set of key-value pairs and produces a transformed set of key-value pairs as the output. The framework shuffles and sorts

outputs of the map tasks, and send the intermediate key-value pairs to the reduce tasks, which group them into final results. For the vehicle's movement and journey trip the results are for example vehicle movement time, idling time, traveled time, and journey trip time. These results are then stored on the hosting database with an additional field to indicate on which vehicle's telematics applications the output are intended to be applied .

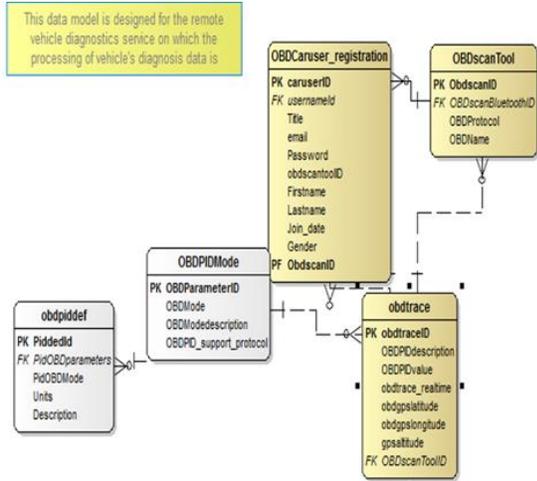


Fig. 10. Data model that includes OBD-PID II description, car user information, and the current vehicle data collected from vehicle.

```
Hive> INSERT OVERWRITE TABLE
> basichadoop.obdresultjson
> select concat(b.firstname,"-",b.lastname)as fullname,
> C.description_pid as DTC_description,max(a.pid_value)
> as measure_units FROM basichadoop.obdtrace a JOIN
> basichadoop.userregistration b ON a.obddevicesaddress
> =b.addressdevice
> right outer JOIN basichadoop.obdpiddef c ON
> a.pid = c.pid_code
> WHERE a.pid_desc is not null and a.obddeviceaddress
> is not null
> and c.pid_code in ('0C','0D','04','05','0F','0B','10',
> '14','RV') group by
> concat(b.firstname,"-",b.lastname),c.description_pid;
> INSERT OVERWRITE LOCAL DIRECTORY
> '/home/mysmart/workspacejee/studyBasicOBD/
> ApplyLicenceproject/hadoop/output.js
```

Fig. 11. HiveQL script for analyzing the current value of OBD-II parameters.

Fig. 9 presents a flow chart of the procedure followed during the implementation of the processing model for vehicle-based sensing data. The in-vehicle's diagnostic data is stored in HDFS; analyzing jobs are executed by queries in HiveQL language, and the queries will be transferred into Mapreduce progress. The results of queries are stored in HDFS as well. Thus, the result in Hive shall be transported from Hadoop to the web server. The aggregation analysis is based on the join of two tables, the "obdtrace" and the "obdpiddef". The data model of the MobiSenseCar application is shown on Fig. 10. The aggregation work is realized by integrating these two tables into Hive script is shown in Fig. 11.

## V. IMPLEMENTATION AND ITS RESULTS

In this section, it presents a performance measurement of the proposed design for implementing the MobiSenseCar mobile application. First, it presents the performance of Client-Server Architecture that features event-driven approach against thread-based approach at the sever-side layer. The goal of this measurement is to investigate which of the Client-Server Architecture configurations is able to support concurrent clients' requests. Secondly, it evaluates the efficiency of Hadoop MapReduce computing using join reduce side algorithm against HiveQL and Statistical Analysis System (SAS) framework [30].

To ensure integrity and reliability of the proposed mobile crowdsensing architecture for MobiSenseCar, different experiments are carried out to collect results. A physical real-time monitoring experiment is performed to ensure data transmission from the ECUsim 2000 equipped with Bluetooth to the mobile device. A MobiSenseCar application is built to assist the car owner in his daily monitoring of engine.

The engine's data synchronization and MobiSenseCar computing system are tested during the experiment to collect all the OBD-II data and seamlessly track the engine status. An event-driven web server Node.js platform is used to run JavaScript outside the browser. A MobiSenseCar Web page is developed for easy monitoring purposes. Finally, the simulation results for evaluating the performance of client-server architecture to handle concurrent requests as well as the performance of the MapReduce against traditional data processing framework are discussed in the second part of this section. Hadoop platform is used to analysis this engine's data and handling the result.

In order to evaluate the collection of car engine data of MobiSenseCar application based on Android's mobile device, the experimentation environment consists of three main components. First, Android mobile device version 4.4.2 is used in the implementation. Second, web server with Node.js, a core i3-3220 CPU within 3.3GH, 16 GB of RAM running Windows seven 64 bit and finally Hadoop multi-node cluster in a distributed environment using three systems (one master and two slaves , each of them is a core i5-6600Processor within 3.90GHz, 16 GB of RAM). Such an environment facilitates to run the collected data from vehicle-based sensing on a real cluster of servers.

### A. Performance Measurement of Event-Driven Approach With Node.js Against Thread-Based Approach with Apache

To measure the performance of event driven approach, the program Tsung was used.<sup>10</sup> The component under test was the main back end server. Tsung works by simulating multiple users making multiple requests to the server. Every user is run in a separate thread. To simulate normal conditions, Tsung allows user think-time and the arrival rate to be specified using a probability distribution. Several tests of each scenario were carried out under different cluster of client-server configurations and varying amount of loads generated by Tsung.

The test plan uses TSUNG to capture throughput, response time results of Node.js's single-thread event-loop against the traditional application based Java on the Apache Tomcat. The capacity and performance testing are required to show that a Client-Server Architecture consists of backend and database layers can run with acceptable responsiveness when a large number of concurrent users access the backend server-side and database simultaneously.

This study has considered three kinds of Client-Server Architectures that provides the backend for server-side implementation and database layers. The first Client-Server Architecture is that the server-side code resides on the Node.js web server and the database is MongoDB for use case. Thus, MongoDB fits perfectly for Node.js applications. Therefore Node.js and MongoDB allows writing JavaScript for the backend and database layer [31]. Furthermore, MongoDB is known for its schemaless nature that gives a better way to match the constantly evolving data structures in the MobiSenseCar application. The second Client-Server Architecture is that the server-side code resides on the Apache tomcat server. The application server that implements the http request is written using JavaServer Pages (JSP) technology [32]. JSP uses the Java programming language. With this model, a relational database MySQL is used as the database. The third Client-Server Architecture consists of Apache Tomcat on the server-side and MongoDB database. Here, it uses the Java API for MongoDB/BSON in Apache Tomcat [27]. For each of the Client-Server Architecture, the goal is not to test the web application but to listen to the http request sent from the mobile device in the same way an http request is submitted from a web browser. The application under test is based on the mobile client server computing that has a module of uploading the vehicle diagnostic data stored in a temporary data store like SQLite to the remote application server.

The test environment for client-server Architecture consists of web framework for Node.js, Node.js server-side for backend and MongoDB for database. This test model environment is configured with the architecture outlined below:

- 1) A user http request arrives over a SSL to the application server.
- 2) The application server forward calls to the server in the web tier.

- 3) The web tier runs in a computer 3.30 Ghz Intel core i3, 8GHZ on Windows server.
- 4) The web tier runs Node.js server, Express for Node.js and code of MongoDB object modeling for Node.js.
- 5) The data tier runs on a separate single virtual server, which hosts the MongoDB database. The rest of the Client-Server architecture have the same environment as the first except both the web tier and data tier configuration.

After defining the test environment for the three Client-Server Architectures for the MobiSenseCar server, it assumes that the goal of the tests is to establish the capacity of a server for handling vehicle diagnostics data from the mobile device by a large number of concurrent users starting simultaneously. Each Client-Server Architecture supports the expected peak load of concurrent users.

Tsung runs these tests locally from a different computer running 2.2GHz Intel core i3, 4GB DDR on OS windows seven. Scripts are created to determine the supported number of concurrent requests of uploading of vehicle diagnostics data and to simulate the concurrent users sessions.

The first tests against Node.js server side and the Apache Tomcat server were conducted simultaneously from 200 users up to 2000 users, 50 times. For each test, the "ramp-up" period is zero (0) which means that all the users start sending the http request simultaneously. We do however need to understand the capacity of the client-server architecture such that we can determine at what point uploading of vehicle diagnostics is completed successfully by a large number of concurrent users. The peak load testing scenarios state are shown in Table 1. The experiments requirements are defined in Table 2.

TABLE I. SCENARIOS CASES FOR THE EXPERIMENTATION

#Scenarios	Concurrent users	Ramp-up period(second)	Loop (times to run the similar sample)
Scenario 1	200	0	50
Scenario 2	1000	0	50
Scenario 3	2000	0	50
Scenario 4	800	200	50

TABLE II. REQUIREMENTS FOR SCENARIOS CASES FOR THE EXPERIMENTATION

Requirements for scenario 1, 2 and 3	Value	Additional information
Volume of load : upload from 200 users up 2000 users at the same time	Peak load : 1 scenario iteration every 0 second , 50 times	Equates to 100000 samples
How long should load be run	Less than 5 minutes	
Acceptance criteria	95% stored in the corresponding database successfully	
Metrics to be reported on	Response time, throughput, error rate	DB activity

<sup>10</sup> <http://tsung.erlang-projects.org>

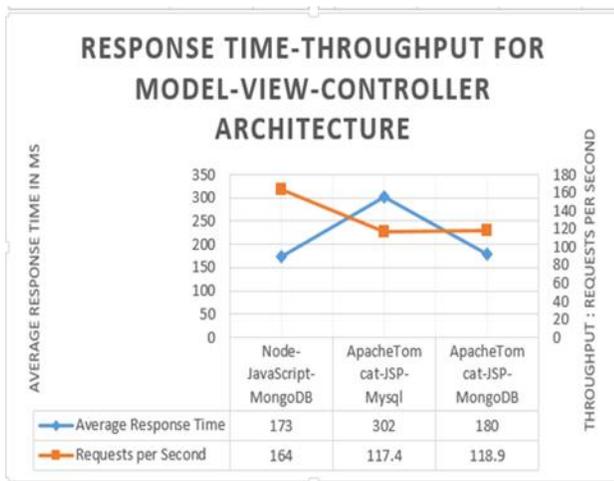


Fig. 12. The performance of three model-view-controller based on the throughput and response time metrics.

Fig. 12 shows the results of the three Client-Server Architecture configurations. To this performance, the results help for analyzing the throughput and response time metrics. The Node.js-JavaScript-MongoDB configuration outperforms. For this architecture, from 200 up to 2000 concurrent users (from 10000 to 100000 requests), the response time is high within 173ms but the throughput in comparison to the response time is less low with 164 requests per second. This signifies that this Client-Server Architecture is capable enough to sustain a large number of concurrent clients 'requests. The Apache Tomcat-JSP-MySQL has a higher response time but the throughput is much lower within 117 requests per second. This signifies that this Client-Server Architecture is not capable enough to execute concurrent requests. The third model that includes Apache at server-side and MongoDB as database outperforms less better in comparison to Node.js-JavaScript-MongoDB but better than Apache Tomcat-MySQL.

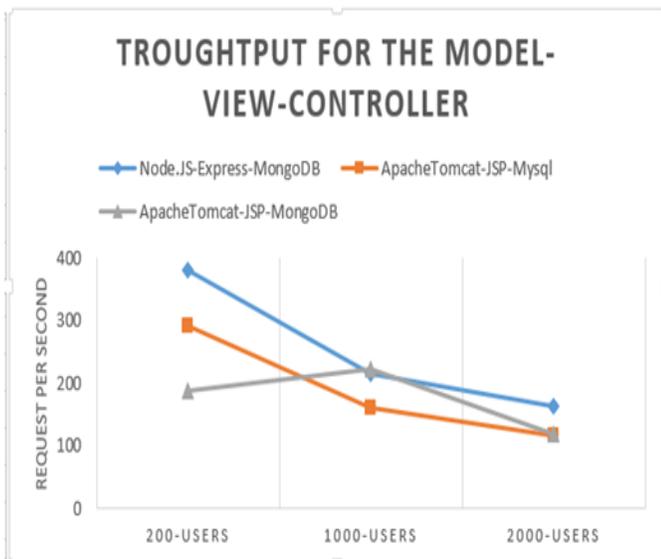


Fig. 13. Throughput of the three model-view controller for n concurrent users.

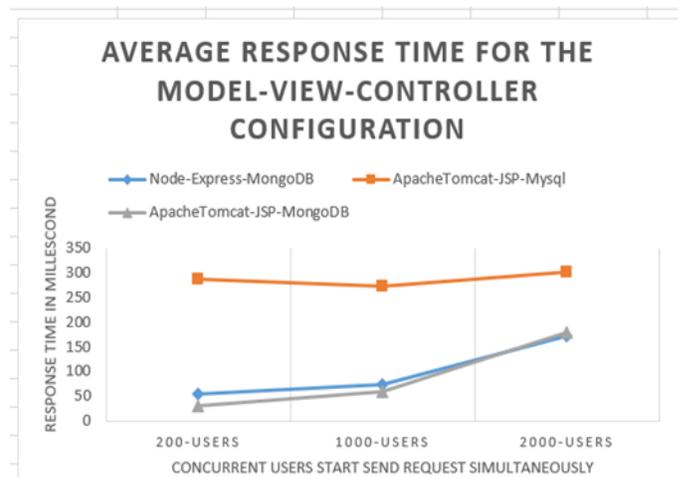


Fig. 14. The measurement of the response time on average for the three-view-controller architecture as the number of concurrent users increase.

Fig. 13 shows that the throughput deteriorates for the three models as the number of concurrent users increases. The results show that for Node.js associated to MongoDB, the throughput was 380 requests at 200 concurrent users, and 164 requests per second at 2000 concurrent users. The throughput of Node.js based event-driven approach has performed better than Apache-Tomcat based thread approach. The results were better using Apache-Tomcat-MongoDB than Apache Tomcat-MySQL and less good than Node.js-Mongo. With Apache-Tomcat-MongoDB configuration, the throughput increases up to 1000 concurrent users, then deteriorates as the number of concurrent users increases. There may be some tremendous opportunities for optimization that could enforce Node.js-MongoDB performance beyond Apache-Tomcat-MongoDB easily.

As shown in Fig. 14, the response time degrades as the number of concurrent requests increases. For example, Node.js-MongoDB was within a response time of 54ms on an average at 200 concurrent users, and 173ms on an average at 2000 concurrent requests. We can see that for Apache-Tomcat at the server side, the average response time has an almost linear correlation to the number of concurrent requests. This means that a thousand-fold increase in concurrent users' results in a hundredfold increase in response time. This that the number of concurrent users carried out by an Apache-Tomcat at server-side is not relatively constant. Therefore, Node.js is roughly 40% faster, for example, 173ms against 302ms for 2000 users that corresponds to one hundred thousand (100000) concurrent requests.

### B. Performance Measurement of Big data Analytics Framework for Processing Vehicle's Diagnostics Data

The performance evaluation of Big data analytics uses representative benchmarks that perform on the datasets from the MySQL tables and MongoDB. The set up experiment environment constitutes of a Hadoop multi-node cluster on a distributed environment using three systems (one master and

two slaves, each of them is a core i5-6600Processor within 3.90GHz, 16 GB of RAM). Such an environment facilitates to run the remote vehicle diagnosis event processing on a real cluster of servers. Diagnostic Trouble code (DTC) is collected into a relational database, and unstructured data are stored in NoSQL database (MongoDB), and then dumped directly into Hadoop cluster. When on-board diagnostics data are uploaded onto the database, Apache Sqoop performs a replication import of data required to run Map Reduce jobs.

When on-board diagnostics data are uploaded to the database, Apache Sqoop performs a replication import of data required to run Map Reduce functions. HIVE has an important role especially for data stored in a relational database. Sqoop generates a Hive table based on the table originally relational data source and also stores data on HDFS.

One of the most key Hadoop jobs in this study is to take the incoming on-board diagnostics and summarizes according to the useful information (see Fig. 8). It stores the processing outcome to the MySQL database or copies it into a format that can be used for further analysis or purposes on the web services. This is achieved by using HIVEQL and Map Reduce functions.

In order to compare the efficiency of Hadoop MapReduce computing using join algorithm particularly reduce side join [33], HiveQL, a higher-level framework in which join are integrated in their implementation and the traditional statistical analysis system (SAS) is used. This SAS @ 9.4 SPD Engine is used for storing data in the HDFS. It also has procedures that can replicate MapReduce's approach for data processing. Since data are in place, map reduce functions converted from HIVEQL can start analyzing them and turn processing outcome into valuable information.

The evaluation computes as well the same statistical values (mean) on Diagnostic Trouble codes stored in MySQL. Fig. 15 shows the processing time of three methods.

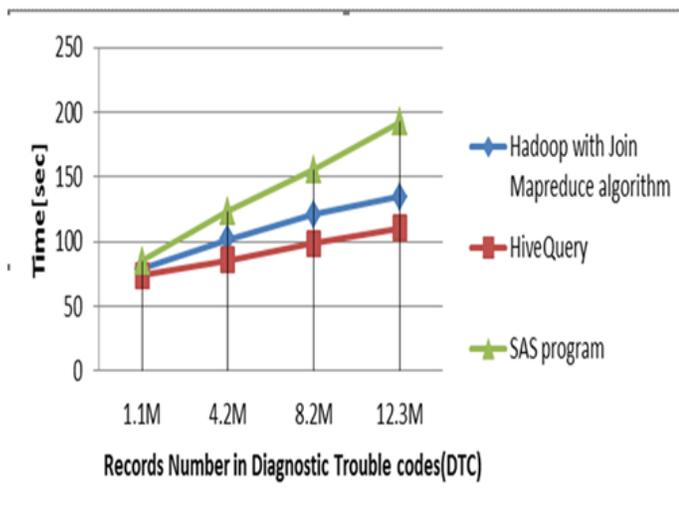


Fig. 15. Processing time of MapReduce framework features multi-way join based reduce-side cascade join.

Fig. 15 comes out with the following observations:

- Comparing to traditional statistical analysis system, Hadoop distributed parallel computing enhances processing speed when the size of dataset to be processed increases and is not significantly different over a lower volume of dataset.
- The Comparison of join algorithms using MapReduce framework to HiveQL, showed that join patterns based MapReduce increases computing speed over HiveQL but it takes time to implement. HiveQL is arguably one of the tools for developers and analysts with strong SQL skills but SQL is not suitable for every big data problem.

### C. Performance Measurement of Big data Analytics Framework for Processing Vehicle's Diagnostics Data

The Client-Server Architecture constitute of Node.js server side and MongoDB is 40% faster that the Java EE solution using Apache Tomcat at the server side with MySQL or MongoDB database for implementing mobile client server computing applications. In this study there were different concurrency models implemented using single-threaded event loop Node.js and multi-thread approach. To test whether Node.js is a higher concurrency level-where it is supposed to surpass multi-threading, other problems like increasing the number of requests. The reason is that Tsung is a 100% pure Erlang to evaluate the functional behavior and measure performance of the three Client-Server Architecture configuration. The experimentations were not able to run these tests beyond 4000 concurrent users, over 200,000 requests.

These study findings have shown what mobile application can do with RESTful responses and requests over a web server. Node.js is considered to perform greatly. In the future it should be taken in consideration for remotely monitoring applications based on mobile device where devices spread in different locations collect a large volume of information. While Node.js outperformed the competition in the scenario of transferring vehicle's diagnosis data, further work can point Node's strengths and weaknesses. Node accomplishes its goals of supporting highly-scalable and reliable web servers. It runs very quick on JavaScript engine. Therefore, Node.js is not designed to stand simply as a replacement of Apache.

Besides using Node.js for handling asynchronous I/O requests, Hadoop stands as a new form for processing a huge amount of data from all car owners subscribed to the MobiSenseCar application. Therefore, compare to traditional statistical analysis system, Hadoop distributed parallel computing enhances processing speed when the size of dataset to be processed increases and is not significantly different over a lower volume of dataset. Thus, comparing data join MapReduce algorithm to HiveQL, relational data join patterns in MapReduce increases computing speed over HiveQL but it takes time to implement. HiveQL is arguably one of the tools for developers and analysts with strong SQL skills but SQL is not suitable for every big data problem [24].

## VI. CONCLUSIONS AND FUTURE WORK

This paper argues the concerns about uprightness of data from vehicle-based sensed data are a major step for vehicle owners, authorities and businesses looking to take up mobile crowdsensing computing that enables value-added and others services. This paper presented a mobile crowd-based architecture which enables car's diagnostic system that features data remote monitoring event processing of vehicle's engine data.

The system design in this paper consists of MobiSenseCar application, an event-driven web server known as Node.js, and Hadoop platform. The MobiSenseCar allows the collection of vehicle's engine data and storage into an embedded application database known as SQLite. At the MobiSenseCar data center infrastructure, the web server Node.js enables MobiSenseCar mobile application requests to be processed asynchronously. Therefore, the request of transmitting vehicle's engine data is not interrupted when a large number of car driver start using the same application simultaneously. The use of Node.js helps to save communication cost and enables car drivers to use other mobile apps on their smartphone when transmitting vehicle's engine data. The Node.js has API (Application Programming Interface) which interact for instance with a NoSQL database (MongoDB).

Taking advantage of reducing communication cost, driver would still use their smartphone for other purpose while the Node.js handles and stores the generated car's status information to a MongoDB without waiting for the MobiSenseCar application to finish the uploading process submitted to Node.js. Thus, processing data from vehicles using Hadoop and making final results available, allows accessing of useful information via web services to the third party such as car manufacturer, transportation and road operators, car dealers, police, emergency services has been conducted on a single node cluster.

The outcome obtained from various Map Reduce functions managed after executing HIVEQL query indicate favorable results in term of time taken. It is unnecessary to receive and deal with all the data including even needless one, so the car owner may handle information satisfying car manufacturers or car repair shop needs only. Therefore, with this system, it was made possible that information of car's diagnostic system condition may be identified in real time

The proposed solution leverages the existing mobile crowdsensing architecture for data collection and processing on vehicle-based sensing. There are still several challenges that must be addressed for this kind of deployment model can be adopted. As future work will focus on the deployment and empirical validation for MobiSenseCar architecture with specific focus on the collection of vehicle-based sensed data stored in real time on the cloud computing based infrastructure as a Service.

## ACKNOWLEDGMENT

This research was supported by the Brain Busan 21 Project in 2017 and The Human Resource Training program for Regional Innovation and Creativity through the Ministry of

Educational and National Research Foundation of Korea (NRF-2015H1C1A1035898).

## REFERENCES

- [1] K. G. Raghun., Y. Fan, and, L. Hui, "Mobile Crowdsensing: Current State and Future Challenges", IEEE Communications Magazine , pp.32-39, 2011
- [2] T. Carlos, C. Celimuge, W. N. Enrico, and J. Francisco, "Crowdsensing and Vehicle-Based Sensing", Mobile Information Systems, 2016, pp. 1-2.
- [3] W. Jinn, C. Jinsong, and M. Tinghuai, "Real time services cloud computing enabled vehicle networks". International Conference on Wireless Communications and Signal Processing (WCSP), 2011, pp: 1-5.
- [4] I.E. Apetri, R. Ali, "Remote Connection of Diagnostic Tool", Master of Science Thesis in Communication Engineering", 2011, pp:1-104
- [5] M. Emanuele, A. Chaewon, and R. Carlo, "The Car as an Ambient Sensing Platform", Proceedings of the IEEE, Vol. 105, No. 01, 2017, pp: 3-7
- [6] S. Tilkov., S. Vinoski, "Node.js : Using Javascript to Build High-Performance Network Programs". Internet Computing, IEEE, 2010 STRIEGEL, GRAD OS F'11, PROJECT DRAFT 6.
- [7] S. Stephan, G. Eszter, R. Wolfgang, "Performance investigation of selected SQL and NoSQL databases", AGILE 2015-Lisbon, 2015, pp:1-5,
- [8] V. Madhavi, "Survey of Parallel Data processing in Context with MapReduce", AIAA 2011, Computer Science & Information Technology 03(CS & IT), 2011, pp:69-80
- [9] M. Prashanth., N.P. Venkata and R. Ramachandran, "Nericell: using mobile smartphone for rich monitoring of road and traffic conditions", Proceedings of the 6th ACM Conference on Embedded network sensor system, 2008, pp:357-371
- [10] J. Eriksson, L. Girod, and B. Hull, "The Pothole Patrol: Using a Mobile Sensor network for Road surface Monitoring", Proceedings of the 6th ACM Conference on Embedded network sensor system, 2008 , pp:29-39.
- [11] A. J. Derick and M. T. Mohan, "Driving Style Recognition Using a Smartphone as a Sensor platform", 14th International IEEE Conference on Intelligent Transportation Systems, 2011, pp:1609-1615
- [12] H. Eren, S. Makinist, F. Akin., and A. Yilmaz, "Estimating Driving Behavior by a smartphone", 2012 Intelligent Vehicles Symposium, 2012, pp:234-239
- [13] J. Eriksson, L. Girod, and B. Hull, "The Pothole Patrol: Using a Mobile Sensor network for Road surface Monitoring", Proceedings of the 6th ACM Conference on Embedded network sensor system, 2008, pp:29-39.
- [14] W. Jules, T. Chris., T. Hamilton, D. Brian, and C. S. Douglas, "Mobile Networks and Applications, Vol. 16, No. 3, 2011, pp:285-303
- [15] C. German, D. Thierry, F. Raphael, and E. Thomas, "Driver Behavior Profiling Using Smartphones : A Low-Cost Platform for Driving Monitoring", IEEE Intelligent Transportation Systems Magazine, 2015, pp:91-102
- [16] An Oracle White paper, "Oracle for the Connected Vehicle: Turning Data into Business", 2013, pp:1-21.
- [17] Gsmam Automotive, "Connecting Cars: Bring your Own Device-Tethering Challenges". Report on Intelligent Transportation system Report, 2013, pp: 1-20.
- [18] B. Erb, "Concurrent Programming for scalable Web Architecture", Diploma Thesis, Institute of Distributed Systems, 2012
- [19] Z. Yuhao, R. Daniel, H. Matthew, J. R. Vijay, "Microarchitectural implications of event-driven server-side web applications", Proceedings of the 48th International Symposium on Microarchitecture, 2015, pp: 762-774
- [20] S. S. Benjamin, L. Maude, "An Inside Look at the Architectural of NodeJS", available on line at [http://mcgill-csus.github.io/student\\_projects/Submission2.pdf](http://mcgill-csus.github.io/student_projects/Submission2.pdf), last access, January, 2016
- [21] C. Bin., M. Hong, C. Beng, "Big data: the driver for innovation in databases", National Science Review, Vol. 1, No. 1, 2014, pp:27-30
- [22] D. B. Arantxa., O. Aisling. "A big data methodology for categorising technical support requests using Hadoop and Mahout". Journal of Big Data, Vol. 1, No. 1, 2014, pp:1-8
- [23] T. White, "Hadoop: The definitive Guide, Third Edition", pp:17-44

- [24] S.P.Poonam, N.P.Rajesh., "Survey Paper on Big Data Processing and Hadoop components".International Journal of Science and Research(IJSR),Vol.3,No10,2014,pp:585-590
- [25] G.Bin, W.Zhu, Y.Zhiwen, G.Yu, Y.Neil.,H.Runhe, and Z.Xingshe, "Mobile Crowd Sensing and Computing : The Review of an Emerging Human-Powered Sensing Paradigm", ACM Computing Surveys, 2015, pp: 1-33
- [26] B. Oresti., V.Claudia, D.Miguel, G.Peter, P.Hector, R.Ignacio, "PhysioDroid: Combining Wearable Health Sensors and Mobile devices for a Ubiquitous, Continuous, and Personal Monitoring", The Scientific World Journal Soc, Article Id 490824, 2014, pp-1-14.
- [27] H.Florian ,P.Rene, "Performance optimization for querying social network data", Workshop Proceedings of the EDBT/ICDT,pp:232-239,2014
- [28] K.Minyoung, andJ. Jang-Wook," Design of Korea smart car driving information checking system". International Journal of Advanced Smart Convergence.1(1),2012,pp:38-42.
- [29] B.Mani, J. Balaraj, M.D.Oinam, "Comparison of Join Algorithms in MapReduce Framework",International Journal of Innovative Research in Computer and Communication Engineering,Vol.2,Special Issue 5, 2014
- [30] M.David, SASReduce-An implementation of MapReduce in BASE/SAS. Whitehound Limited, UK. Online athttp://support.sas.com/resources/papers/proceedings14/1507-2014.pdf, (2014), 1-16.
- [31] K.Brian,, "CS764 Project Report: Adventures in Moodle Performance Analysis", available on line at [http://pages.cs.wisc.edu/~bpkroth/cs764/bpkroth\\_cs764\\_project\\_report.pdf](http://pages.cs.wisc.edu/~bpkroth/cs764/bpkroth_cs764_project_report.pdf), pp:1-28,last access, March 2016
- [32] L.N.Glenn, "Tomcat Performance Tuning and Troubleshooting",ApacheConference, pp:1-10,2003
- [33] N.A.Foto, and D.U.Jeffrey, "Optimizing Joins in a Map-Reduce Environment",IEEE Trans.Knowl.Data Eng.23(9):2011,pp1282-1298

# Impedance Matching of a Microstrip Antenna

Sameh Khmailia

Dept. of Physics  
El Manar University, Faculty of Sciences  
Tunis, Tunisia

Sabri Jmal

Dept. of Physics  
El Manar University, Faculty of Sciences  
Tunis, Tunisia

Hichem Taghouti

Dept. of Physics  
El Manar University, Faculty of Sciences  
Tunis, Tunisia

Abdelkader Mami

Dept. of Physics  
El Manar University, Faculty of Sciences  
Tunis, Tunisia

**Abstract**—Microstrip patch antennas play a very significant role in communication systems. In recent years, the study to improve their performances has made great progression, and different methods have been proposed to optimize their characteristics such as the gain, the bandwidth, the impedance matching and the resonance frequency.

This paper presents a new method that allows to ameliorate the impedance matching, thus to increase the gain of a rectangular microstrip antenna.

This method is based on the adaptation technique using a simple “L” matching network.

The originality of this work is the exploitation of the principle of causality that permits to detect the problems of reflected waves and to obtain the suitable placement of components that constitute the matching circuit.

**Keywords**—Impedance matching; microstrip antenna; “L” matching network; bond graph model; principle of causality; wave matrix; scattering matrix; transmission and reflection characteristics

## I. INTRODUCTION

Microstrip patch antennas have been widely researched and developed in recent years. They are used in very specific communication applications such as radars, satellite, broadcasting, radio frequency identification [1].

In its simplest configuration, microstrip antenna consists of a radiating patch on one side of dielectric substrate, which has a ground plane on the other side [2], [3].

Compared with conventional antennas, microstrip patch antennas have better prospects and more advantages. They are lighter in weight, smaller in dimensions and they can be easily integrated with RF and microwave systems [4].

However, microstrip antenna have some drawbacks including narrow bandwidth, low power handling capability, low gain, low impedance matching [5]. But with technology advancements and extensive research into this area, these problems can be gradually overcome.

In this paper, we will present a new impedance matching

technique. This method is based on the principle of causality of bond graph and the “L” matching network and permits to improve the impedance matching of a rectangular microstrip antenna.

Firstly, we will present the notion of impedance matching and its different methods.

After that, we will choose the “L” matching network as the best method for improving the impedance matching and the gain of a microstrip antenna.

The principle of causality of bond graph will be exploited to detect the reflected waves and to deduce the appropriate structure of the “L” matching network.

## II. IMPEDANCE MATCHING

In the high frequency domain, impedance matching is one of the most important applications used to improve the systems performances [6]. It requires adding a matching network between the source and the load of a system in the aim to ensure a maximum power transfer and to avoid the reflected waves [7].

The impedance matching network can be realized using a transformer, a quarter wave line or an “LC” network as “L”, “T” or  $\Pi$  form. The most commonly used matching network in RF processing is the “L” network due to its simplicity, it is constituted only of two components; an inductance and a capacitance that can be controlled for adjusting the real and the imaginary part of the impedance without power loss.

The main principle of any impedance matching scheme is to force the load impedance to look like the complex conjugate of the source impedance, so a maximum power will be transferred to the load.

In RF domain, the first impedance matching concept was related to antenna matching with the aim to obtain desired characteristics such as the gain, the bandwidth and the resonance frequency without modifying the antenna geometry.

In the following, an “L matching network” will be used to improve the impedance matching of a microstrip antenna.



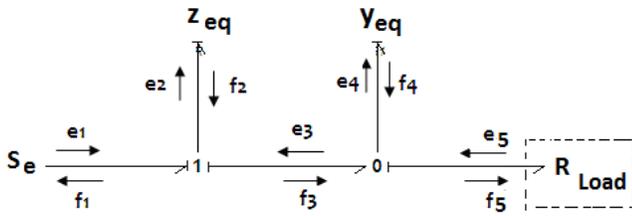


Fig. 4. Reduced bond graph model.

The reduced bond graph model is given in Fig. 4, it is constituted only by an impedance  $z_{eq}$  linked to 1-junction and an admittance  $y_{eq}$  linked to a 0-junction [19], [20].

$R_{Load}$  : is an imaginary load

The resistance of generator is replaced by reduced impedance  $z_{eq}$  linked to a 1-junction and the parallel elements R1, L1 and C1 are replaced by a reduced admittance  $y_{eq}$  linked to a 0-junction.

$$z_{eq} = r_g \quad (1)$$

$$y_{eq} = \frac{1}{r_1} + \frac{1}{\tau_{L1}} + \tau_{C1} \quad (2)$$

The normalized values of  $R_G$ , R1, L1 and C1 are given in the following equations:

$$r_g = \frac{R_G}{R_0} \quad (3)$$

$$r_1 = \frac{R1}{R_0} \quad (4)$$

$$\tau_{C1} = C1.R_0 \quad (5)$$

$$\tau_{L1} = \frac{L1}{R_0} \quad (6)$$

Basing on the bond graph theory, the impedance linked to the 1-junction can be represented by a wave matrix  $W_1$  :

$$W_1 = \begin{bmatrix} \frac{z_{eq} + 2}{2} & \frac{-z_{eq}}{2} \\ \frac{z_{eq}}{2} & \frac{2 - z_{eq}}{2} \end{bmatrix} \quad (7)$$

The admittance linked to 0-junction can be represented by a wave matrix  $W_2$  :

$$W_2 = \begin{bmatrix} \frac{y_{eq} + 2}{2} & \frac{y_{eq}}{2} \\ \frac{-y_{eq}}{2} & \frac{2 - y_{eq}}{2} \end{bmatrix} \quad (8)$$

The total wave- matrix  $W_T$  of the system is obtained by the product of  $W_1$  and  $W_2$  :

$$W_T = W_1.W_2 = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \quad (9)$$

To determine the transmission and reflection coefficients, we must determine the Scattering matrix called S- matrix. The following linear transformation that links the S and W matrices can be exploited [21], [22], [23].

$$\begin{cases} S_{11} = W_{22}W_{12}^{-1} \\ S_{12} = W_{21} - W_{22}W_{12}^{-1}W_{11} \\ S_{21} = W_{12}^{-1} \\ S_{22} = -W_{12}^{-1}W_{11} \end{cases} \quad (10)$$

Thus, the total S- matrix describing the antenna is given in the following equation:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} W_{22}W_{12}^{-1} & W_{21} - W_{22}W_{12}^{-1}W_{11} \\ W_{12}^{-1} & -W_{12}^{-1}W_{11} \end{bmatrix} \quad (11)$$

The adaptation level of the antenna is determined from the characteristic of the reflection coefficient S11 as function of frequency given in Fig. 5.

According to the variation characteristic of S11 as function of frequency, the impedance matching is not ensured. In fact, in the antenna theory, a microstrip antenna is considered adapted only if the absolute value of S11 is greater than or equal to 15 dB.

To remedy this problem, thus to ensure a maximum power transfer, we will use the "L" matching network. The technique is explained in the following paragraph.

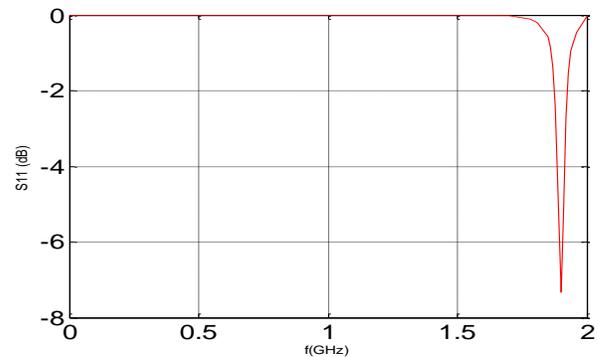


Fig. 5. Reflection coefficient S11 as function of frequency.

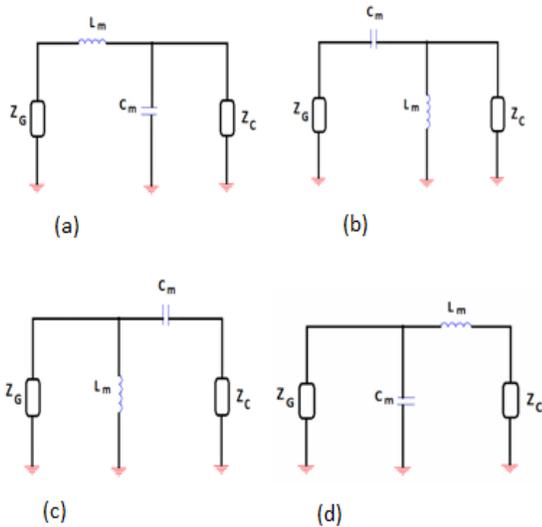


Fig. 6. (a, b, c, d): Four configurations of L matching network.

#### D. "L" Matching Network Application

The "L" match circuit gets its name because the circuit topology can look like the letter "L". It is one of the easiest lossless ways of matching the source impedance to the load impedance.

The four configurations containing "L" matching network given in Fig. 6 are possible.

$Z_G$  : is the impedance of the generator

$Z_C$  : represents the characteristic impedance of antenna

$L_m$  : Inductance of "L" matching network

$C_m$  : Capacitance of "L" matching network

The concept of causality of bond graph can be used to choice the more appropriate disposition of the capacitance  $C_m$  and the inductance  $L_m$  in the L matching network [24]. In fact, the principle of causality imposes that:

- For 0-junctions, one of the bonds sets the effort for the rest, so only one causal stroke is on the junction, while the others are away from it.
- For 1-junctions, one of the bonds sets the flow for the rest, and its effort is computed from them, so all but one of the causal strokes are on the junction, while the remaining one is away from it.

Using the bond graph model, the previous four representations of matched antenna can be represented such as given in Fig. 7(a, b, c, d).

According to the principle of causality, for 0-junction, only one causal stroke must be on junction, while others must be away from it. This principle is not respected in the models given in Fig. 7(a) and 7(b).

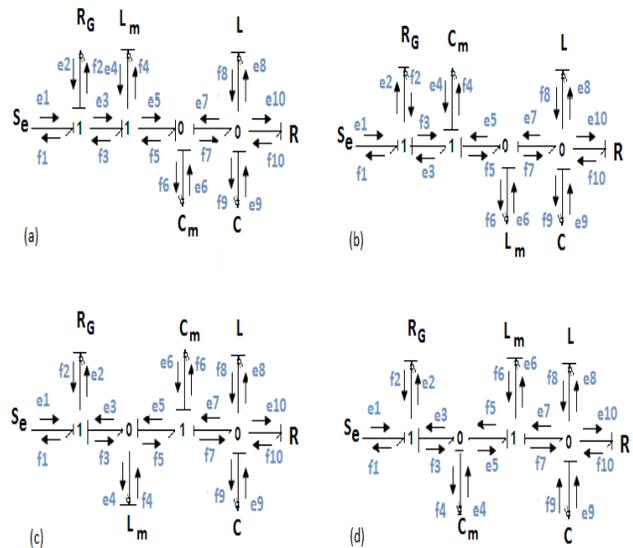


Fig. 7. (a, b, c, d): Bond graph models of four matched antenna configurations.

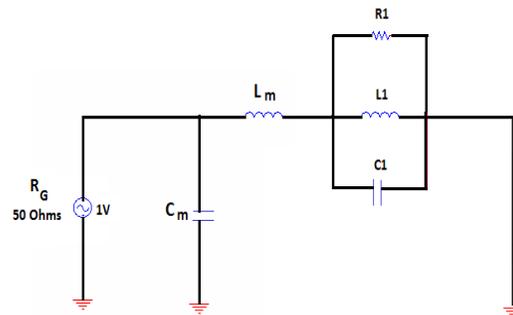


Fig. 8. Electric model of matched antenna.

Thus, we will choose between the structure given in Fig. 7(c) and 7(d).

If we choose the model given by Fig. 7(d), the electric model of matched antenna is given in Fig. 8.

To find the values of  $L_m$  and  $C_m$  at a resonance frequency  $f_r = 1.9GHz$ , we use the equations given below [25]:

$$C_m = \frac{1}{2\pi f_r} \sqrt{\frac{R_1 - R_2}{R_2}} \quad (12)$$

$$L_m = \frac{R_2}{2\pi f_r} \sqrt{\frac{R_1 - R_2}{R_2}} \quad (13)$$

Over a frequency range from 0 to 2GHz, we have compared the results of the variation of the reflection coefficient S11 as function of frequency with and without the "L" matching network. The result is shown in Fig. 9.

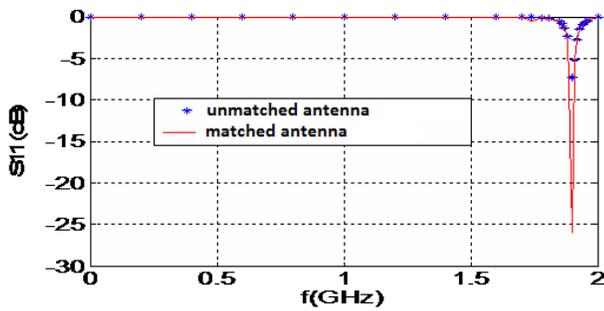


Fig. 9. S11 as function of frequency.

At the resonant frequency, the absolute value of S11 has gone up from 7.5 to 26 dB.

The “L” matching network permitted to improve considerably the gain of the microstrip antenna which is explained by a good impedance matching.

#### IV. CONCLUSION

Impedance matching is a widely used application in high frequency circuit design. It is concerned with matching one part of a circuit to another in order to achieve a maximum power transfer between the two parts and to minimize the reflected waves. There are different impedance matching network such as the “L” matching network, the transmission line matching circuit and the quarter wave impedance transformer.

In this work, we used the “L” matching network to match the impedance of a microstrip antenna, thus to improve its gain and resonant frequency. This choice is due to the simplicity, low cost and ease of design of an “L matching network”; it is only constituted by an inductance and a capacitance.

The principle of causality is used to detect the problem of reflected waves and to choose the more appropriate disposition of the “L” circuit components.

The method presented in the work of A. Mami and H. Taghouti is exploited to determine the scattering matrix of the antenna.

The simulation result of the reflection coefficient S11 as function of frequency with and without impedance matching shows that the gain of adapted antenna is more important.

#### REFERENCES

- [1] K. Rama Devi, A. Mallikarjuna Parasad and A. Jhansi Rani, “Design of Pentagon Microstrip Antenna for Radar Alimenter Application,” *International Journal of Web & Semantic Technology, IJWest*, Vol.3, No.4, October 2012.
- [2] Anushi Arora, Aditya Khemchan dani, Yash Rawat, Shashank Singhai, Gaurav Chaitanya, “Comparative Study of different Feedin Techniques for Rectangular Microstrip Patch Antenna,” *International Journal of innovative Research in Electrical, Electronics, Instrumentation and control Engineering* Vol 3, Issue 5, May 2015.
- [3] Amri Singh, “ Design of square patch microstrip antenna for circular polarization usgin IE3D Software”, National Institute of thechnology Rourkela Orissa India.

- [4] S.M. Yang, C.H. Huang, C.C. Hong, “Design and analysis of microstrip antenna arrays in composite Laminated Substrates” *Journal of Electromagnetic Analysis and Applications*, pp. 115–124, 2014.
- [5] J. Ankur Kaushal, Sachin Tyagi, “ Microstrip Patch antenna its Tyoe, merits, demerits and its applications” *International Journal of Engineering Sciences & Research Technology*, July 2015.
- [6] Thesis of Vishwanath Iyer, “Broad band impedance matching of antenna radiators,” The faculty of Worcester polytechnic institute in Electrical and Computer Engineering .
- [7] Thesis of Victor Freitas, “Etude et réalisation de réseaux d’adaptation d’impédance accordables linéaires et non linéaires sur PCB et silicium CMOS pour des applications en radiofréquences,” Grenoble University, 2006.
- [8] Wolff, I and Knoppik, N, “Rectangular and circular microstrip disk capacitors and resonators,” *IEEE Trans*, 1974, MTT-22, pp, 857- 864.
- [9] A.M.Niknejad” Two-port networks and amplifiers”, EECS 142.
- [10] F.Caspers”RF engineering basic concepts: S- parameters”, CERN, Geneva Switzerland .
- [11] H.Taghouti, and A.Mami, “Discussion around the Scattering Matrix Realization of a Microwave Filter using the Bond Graph Approach and Scattering Formalism”. *American Journal of Applied Sciences*, Vol. 9, Issue.4, 2012.
- [12] H.Taghouti, and A. Mami, “Modeling Method of a Low-Pass Filter Based on Microstrip T-Lines with Cut-Off Frequency 10 GHz by the Extraction of its Wave-Scattering Parameters from its Causal Bond Graph Model”. *Am. J. Eng. Applied Sci.*, Vol. 3, pp. 631-642, 2010.
- [13] Laurent Chusseau, “ Paramètres S- Antennes” , Centre d’électronique et de Microélectronique de Montpellier.
- [14] TAGHOUTI, H. et MAMI, A, “Extraction, Modelling and Simulation of the Scattering Matrix of a Chebychev Low- Pass Filter with cut-off frequency 100 MHz from its Causa and Decomposed Bond Graph Model”, *ICGST International Journal on Automatic Control and System Engineering*. Vol. 10, Issue I, pp. 29-37, 2010.
- [15] JANUSZ A. DOBROWOLSKI, “Microwave Network Design Using the Scattering Matrix”, Artech House.
- [16] Masson, S.J,” Feedback theory: Further properties of signal flow graphs”, *Monographs in modern engineering science* New York, Mc Graw Hill, 1964, pp.920-926.
- [17] BREEDVELD, P.C, “Computer-aids for Modelling and Conceptual Design” , IEEE SMC’93, Int. Conf. on Systems Man and Cybernetics. Le Touquet, Vol. 2, pp. 209-215, Oct. 1993.
- [18] Amara, M and Scavarda, S, ” A procedure to match bond graph and the scattering formalism”, *J. of the Franklin Inst*, Vol.328, N° 516, pp. 887-899, 1991.
- [19] S. Khmailia, H. Taghouti, R. Mehouchi and A. Mami, “Application of the reduced bond graph approach to determine the scattering parameters of a matching network of a planar inverted f antenna”, *International Journal of Advances in Engineering & Technology*, Vol. 3, Issue 2, pp. 77-89. May 2012.
- [20] C.Sueur and G.Dauphin- Tanguy, “Bond-graph approach for structural analysis of mimo linear systems”, *Journal of the Franklin institute*, Volume 328, Issue 1, 1991, Pages 55-70.
- [21] Sameh KHMALIA, Hichem TAGHOUTI & Abdelkader Mami, ” AStudy of A Band Stop Filter”, *European Journal Of Scientific Research*, Vol. 95 No 4 pp.582-588, February 2013.
- [22] Riadh MEHOUACHI, Hichem TAGHOUTI, Sameh KHMALIA, & Abdelkader Mami, ”Modeling and Simulation of the Patch Antenna by Using a Bond Graph Approach”, *International Journal of Advances in Engineering & Technology*, Vol. 2, Issue 1, pp.474-484, Jan 2012.
- [23] Raul G. Longoria, “Wave-scattering formalisms for multiport energetic systems” , *J.Franklin Inst*.vol.333(B).N°4,pp.539-564, 1996.
- [24] Rosenberg, R.c “Exploding Bond Graph Causality in physical Systems Models” *Trans. ASMEJ. OF Dyn. Sys. Meas. And Control*” 1987 vol.109, N 4, pp.206- 2012.
- [25] François de Dieuleveut, Olivier Romain “Eléctronique appliquée aux hautes fréquences”.

# Development of A Clinically-Oriented Expert System for Differentiating Melanocytic from Non-melanocytic Skin Lesions

Classification of PSLs by Abbas Q

Qaisar Abbas

College of Computer and Information Sciences,  
Al Imam Mohammad Ibn Saud Islamic University (IMSIU),  
Riyadh, Saudi Arabia

**Abstract**—Differentiating melanocytic from non-melanocytic (MnM) skin lesions is the first and important step required by clinical experts to automatically diagnosis pigmented skin lesions (PSLs). In this paper, a new clinically-oriented expert system (COE-Deep) is presented for automatic classification of MnM skin lesions through deep-learning algorithms without focusing on pre- or post-processing steps. For the development of COE-Deep system, the convolutional neural network (CNN) model is employed to extract the prominent features from region-of-interest (ROI) skin images. Afterward, these features are further purified through stack-based autoencoders (SAE) and classified by a softmax linear classifier into categories of melanocytic and non-melanocytic skin lesions. The performance of COE-Deep system is evaluated based on 5200 clinical images dataset obtained from different public and private resources. The significance of COE-Deep system is statistical measured in terms of sensitivity (SE), specificity (SP), accuracy (ACC) and area under the receiver operating curve (AUC) based on 10-fold cross validation test. On average, the 90% of SE, 93% of SP, 91.5% of ACC and 0.92 of AUC values are obtained. It noticed that the results of the COE-Deep system are statistically significant. These experimental results indicate that the proposed COE-Deep system is better than state-of-the-art systems. Hence, the COE-Deep system is able to assist dermatologists during the screening process of skin cancer.

**Keywords**—Skin cancer; melanocytic; non-melanocytic; dermoscopy; deep learning; convolutional neural network; stack-based autoencoders

## I. INTRODUCTION

Melanocytic and non-melanocytic (MnM) skin lesions [1] are the two major form of skin cancer. According to estimation in 2016, the skin cancer is rapidly increasing throughout the world and it is very common in white skin populations. Even in the United States, skin cancer is the most common form of cancer. For clinical experts, they have to first decide whether the lesion belongs to melanocytic or non-melanocytic (MnM) class. After identification of this step, the clinical experts then classify the melanocytic lesion is benign or malignant. Whereas in a case of non-melanocytic lesions, the experts have to further classify them as a basal cell carcinoma (BCC), squamous cell carcinoma (SCC) or seborrheic keratosis (SK) skin lesions.

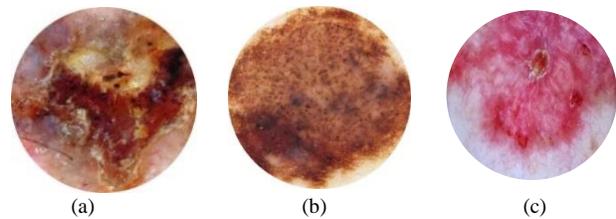


Fig. 1. An example of melanocytic skin lesions (a, b) and non-melanocytic skin lesion dermoscopy samples images, where figure (c) represents BCC - Basal Cell Carcinoma.

An example of these lesions is visually represented in Fig. 1. All these classes are known as pigmented skin lesions (PSLs). Among different types of pigmented skin lesions (PSLs), the malignant melanoma has the highest mortality rate. Despite this fact, the occurrence of melanoma and non-melanoma skin cancers are increasing with the highest rate. For early detection of skin cancer, it can definitely reduce the mortality of this disease. To diagnosis PSLs, the dermatologists are widely using digital dermoscopy with automatic image analysis computer-aided diagnostic (CADx) [2] system. In general, the dermoscopy equipped with CADx system is provided the most cost-effective non-invasive technique for early detection.

Over the last few years, the computer-aided diagnostic (CADx) systems are developed for automatic classification of pigmented skin lesions (PSLs). Those CADx systems were used for providing the second opinion to dermatologists and assist them in better diagnosis of skin cancer. For classification of CADx system into melanocytic and non-melanocytic categories, it is very crucial due to highest similarity among them. Compared to existing melanoma CAD system [3], the recognition rate of non-melanoma skin lesions is below than 75%.

To differentiate PSLs lesions, the authors developed many state-of-the-art CADx tools [4] because the diagnosis by clinical experts is based on subjective whereas, a CADx system is more objective and reliable. The current CADx tools [5], [6] are developed based on hand-crafted features combine with machine learning algorithms such as neural network

(NN), support vector machines (SVMs), AdaBoost and deep-learning to achieve very good performance on certain skin cancers such as melanoma. But they are unable to perform diagnosis [7] over bigger classes of skin diseases such as in the case of melanocytic and non-melanocytic (MnM) categories.

Human hand-crafted features are not providing a perfect solution for the development of CADx system for automatic diagnosis MnM skin lesions. In practice, the hand-crafted features required high expertise for domain-expert knowledge and it is suitable only for limited skin diseases. On the other hand, the deep learning algorithms are utilized in the few studies for the development of CADx tools. By using deep learning algorithms, the hand-crafted features are no need to define and it extracted automatically from an image. As a result, there is no need domain expert knowledge or pre- or post-processing steps to recognize PSLs lesions. Even for large-scale datasets, the deep-learning algorithms have displayed high performance compared to other algorithms such as NN, SVM or AdaBoost. Inspired by deep-learning algorithms, the convolutional neural network (CNN), stack-based autoencoders (SAE) and soft-max linear classifiers are integrated into this paper to get higher performance in terms of large-scale applicability of CADx tools to automatically diagnosis PSLs lesions.

The rest of the paper is organized as follows. Section 2 introduces the background about this research study and deep learning architectures. In Section 3, the dataset and the proposed methodology are technically described. Section 4 shows the experimental results on the performance of the deep-learning algorithms using different training settings. Conclusions and future works of this paper are given in Section 5.

## II. BACKGROUND

The past studies suggested that the researchers focused only the classification of melanocytic lesions (benign and melanoma) from dermoscopy images due to certain issues mentioned in the previous section. In practice, it is not so easy for clinical experts to differentiate among non-melanocytic lesions [8] such as SK, BCC or SCC compared with melanocytic lesions. Due to this reason, the differentiation between melanocytic and non-melanocytic (MnM) skin lesions is the first and important steps that are ignored currently by many computer-aided diagnostics (CADx) systems. As those CADx tools were trained and developed through melanocytic lesions and if we provided those non-melanocytic lesions then the results showed unreliably. In this case, if the CADx system is extended to work with non-melanocytic lesions then the system should have the capability to recognize them as well.

To develop those CADx systems, there are mainly four steps involved such as image enhancement, segmentation, feature extraction and selection, and recognition. As a result, it is very much difficult for a person to develop a CADx system without having expertized on complex image processing techniques. In addition to this, the segmentation of non-melanocytic lesions is very difficult to compare to melanocytic lesions due to rough and intensity variation

around the lesion border. Moreover, the old CADx tools were developed through old machine learning algorithms such as artificial neural network (ANN), support vector machines (SVMs) and AdaBoost classifiers to recognize only melanocytic lesions. However, those CADx tools required lots of pre- or post-processing steps and domain expert knowledge for features selection. Also, those CADx tools were only applied on a limited dataset. Therefore in this paper, a deep-learning modern machine learning algorithms are used to differentiate between melanocytic from non-melanocytic (MnM) pigmented skin lesions, which applies in a large-scale environment. According to my limited knowledge, there is no study available that classify MnM through deep learning algorithm.

There are few CADx tools developed in the past to recognize only melanocytic skin lesions based on deep learning architectures. At the beginning, the most famous architecture was used is CNN model to extract the features and then the decision of classification is performed based on softmax linear classifier. As mentioned above, the CNN model can be used to select features for multiple objects. Therefore, the use of simply CNN model is not suitable for differentiation between MnM skin lesions. Those CADx tools are mentioned in the subsequent paragraphs.

The support vector machines (SVM) and deep belief network (DBN) are combined together in [9] to recognize a limited number of dermoscopy images such as 100. This system is tested on a set of the limited data set so unsuitable for a large-scale environment. In [10], the hybrid version of AdaBoost-SVM and deep neural network are integrated to learn hand-crafted features for classification of melanoma skin lesion. Also in [11], the SVM is combined with deep learning and sparse encoder techniques to classify melanoma images on 2624 images and reported 91.2% accuracy. By using of deep convolutional neural networks (DC-NN) machine learning algorithm in [12], the authors developed a three pattern detectors approach on a set of 211 images and reported accuracy below than 85%. The CNN model used in [13] to extract features with pooling techniques to recognize PSLs skin lesions and achieved 85.8% accuracy. The deep-neural-network (DNN) is used to classify melanoma and achieved 89.3% accuracy. Similarly, the authors in [14] used CNN model to dermoscopy images to classify malignant melanoma skin lesions.

The above-mentioned CADx tools are just used to classify melanoma skin lesions instead of non-melanoma lesions that are the first step required by dermatologists. In the past approaches, there is only one study [15] developed for differentiation between melanocytic and non-melanocytic skin lesions but required pre- or post-processing steps.

Hence, this paper is focused on both categories and developed an automatic system through deep-learning algorithms. Deep learning algorithms are based on multilayer architecture and each is connected with other in a non-linear combination [16]. There are many variants of deep-learning algorithms such as convolution neural network (CNN), deep belief network (DBN), restricted Boltzmann machine (RBM) and state-based autoencoders (SAE). For differentiation

between melanocytic and non-melanocytic (MnM) skin lesions, the CNN, SAE are integrated together and the final decision is performed through softmax linear classifier [17]. In fact, the CNN model is used to best extract features from the pixels of the images and converted them into edges through its multilayer architecture approach. Afterward, the features are extracted by CNN model, are not optimized, therefore, the stack-based autoencoders (SAE) are employed to automatically select most discriminative features for better classification. As a result, the deep-learning algorithms are utilized to diagnosis pigmented skin lesions.

### III. METHODOLOGY

The clinically-oriented expert system through deep learning (COE-Deep) algorithms involve three main steps such as extraction of deep features, optimization of deep features and classification of these features into melanocytic and non-melanocytic skin lesions. The overall systematic diagram of COE-Deep system is shown in Fig. 2. These phases are explained in the following sub-sections.

#### A. Dataset Acquisition

Clinically-oriented expert system using deep learning (COE-Deep) algorithms is tested on 5200 dermoscopy images contains an equal number of melanocytic and non-melanocytic skin lesions. These images were obtained from many public and private resources. Among 2300 dermoscopy images, the 400 melanocytic and another 400 non-melanocytic skin lesions are collected from EDRA [18] as a CD-room. One more, the dataset was collected from the Department of Dermatology, University of Auckland (DermAuck) [19]. The DermAuck dataset contains 600 melanocytic and 600 non-melanocytic lesions. The total 1600 melanocytic and 1600 non-melanocytic skin lesions were collected from the International Skin Imaging Collaboration (ISIC) [20]. In total, the dataset of 5200 dermoscopy images is obtained from these three different sources along with different image sizes. All these images were resized to a standard size of (800 X 800) pixels resolution. Moreover, an expert dermatologist was requested to verify the images in all these two categories. The images contain skin lesion with other skin areas. Therefore from the center position of each image, the circular region-of-interest (ROI) of size (400 X 400) pixels is automatically selected. An example of this dataset is also displayed in Fig. 1.

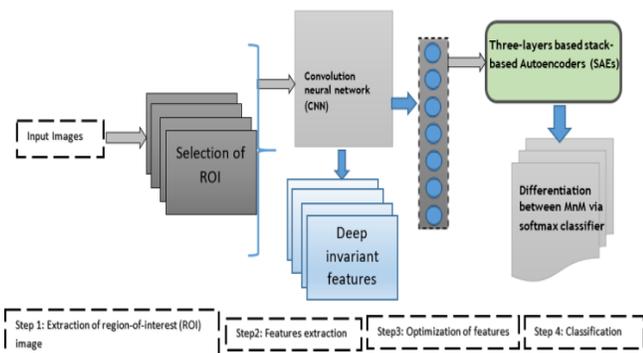


Fig. 2. A systematic flow diagram of proposed COE-Deep system for classification of melanocytic and non-melanocytic skin lesions.

#### B. Features Extraction

During last few decades, the discriminative features extraction and selection becomes one of the difficult and challenging tasks because the subsequent recognition step depends on this step. As mentioned above, the features selected required domain expert knowledge for defining handcraft features and there are lots of steps about pre- or post-processing. Therefore in this paper, the convolutional neural networks (CNNs) model [17] is used to automatically select features from the raw pixels of the image. The CNNs model is used because it is utilized as a major tool in the past studies for classification problems. The CNNs model is applied to the pixel of images and there is no need to manually perform features extraction technique to define handcrafted features set. If the CNNs model is used to extract the features then without overfitting, it can have possible train the deep network in a sensible amount of time.

In this article, the CNN model employs in the form 3-layers deep neural networks to solve the problem of features selection from dermoscopy images. The first layer is directly linked to the image pixels and generated features map after convolving layer filter. In the second layer, the similar features map are combined to generate edges that are presented in dermoscopy images. At last, the third layer is used to select mean activation function of the features from edge map. In this paper, the unsupervised approach of CNNs model is employed.

The mathematical description of the CNN model is defined on a set of  $k$  filters, filters element as  $Fltr_{c,l,k}$  and elements as  $Elmt_{c,x+l,y+k}$  with  $C$  channels of size  $(m \times n)$  with a set of  $N$  images with  $C$  channels of size  $(l \times k)$ . Based on this description, the first convolutional layer output  $Cnn_{i,k,x,y}$  is given as:

$$Cnn_{i,k,x,y} = \sum_{c=1}^C \sum_{l=1}^m \sum_{k=1}^n \dots Elmt_{c,x+l,y+k} Fltr_{c,l,k} \quad (1)$$

And the output of an entire image/filter in the convolutional process is defined in CNN model as pairs as follows:

$$Cnn_{i,k} = \sum_{c=1}^C Elmt_{c,k} * Fltr_{c,k} \quad (2)$$

Where  $*$  represents 2D correlation. Fig. 3 illustrates the utilization of CNNs model to extract the features from the dermoscopy images.

#### C. Optimization

The features defined by CNNs model is not optimized. To optimize the most discriminative deep-invariant features, the stack-based autoencoders (SAEs) [17] is applied. In this paper, the SAEs algorithm is selected because it depicts the behavior of the human-like brain. The best results described in the past studies, if the supervised SAEs algorithm and four layers were used to optimize the deep features. In practice, the SDAs algorithm hypotheses are tested through trained greedy layer-wise pre-training approach on the testing dataset. The main steps for the development of features optimization through SAEs are presented here.

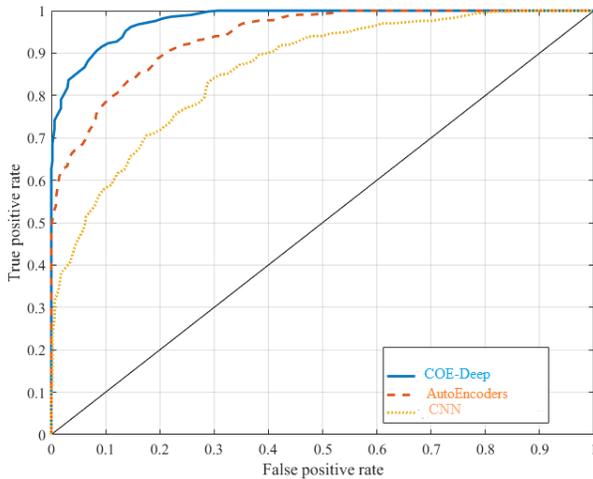


Fig. 3. Performance comparisons of proposed DermaDeep system with state-of-the-art classification systems in terms of Area under the Receiver operating curve.

In general, the pixels in an image that represents the feature vectors defined as an input hidden layer by autoencoders. However, the first input hidden layer in this paper is defined on features generated in the previous step. The second and third hidden layers transform those features into best representation, and an output final hidden layer matches the input layer for reconstruction. Autoencoders is assumed to be deep if the number of hidden layers is greater than one. Moreover, in this study, the original dimension of the hidden layers are defined small to perform features reduction step. Specifically, the autoencoders are developed through stochastic gradient descent method and trained by back propagation variants.

The mathematical description of autoencoders it to learn the code ( $H_y$ ) from the features data, ( $F_x$ ) and map with weights ( $W$ ) according to some sigmoid ( $Sgm$ ) function. It is defined as:

$$H_y = Sgm(W.F_x + b) \quad (3)$$

Where,  $b$  represented the biases of autoencoders. The code is then mapped back through a decoder into a reconstruction ( $R$ ) through the similar transformation as mentioned above and defined as:

$$R = Sgm(W_0.H_y + b_0) \quad (4)$$

And the reconstruction error is measured as:

$$L(F_x; R) = \|F_x - R\|^2 \quad (5)$$

To minimize this  $L(F_x; R)$  mean square reconstruction error, the stochastic gradient decent approach was used in the training process of an AutoEncoder. This minimization step is performed by searching the weights on the encoder and decoder's connection, and share those weights on the encoder and decoder that utilized the same weights. As a result, this step is definitely used to reduce the features by  $\frac{1}{2}$  without having any deficiency on the performance of autoencoders. The autoencoders with these four layers are not sufficient to take the final classification decision due to over-fitting

problem on this deep neural architecture. Therefore, the softmax linear classifier is used to take the final classification decision.

#### D. Classification

The softmax classifier is normally utilized in the past studies to recognize the objects or features through logic regression classifier in the form of binary representation. The softmax linear classifier [17] proceeds with a vector of random real-valued scores and compresses them into a vector of values between zero and one. The decision of differentiation is performed by softmax classifier based on normalize class probabilities and normally, this classifier is used to reduce the cross-entropy between estimated of class probabilities and the known distribution.

### IV. EXPERIMENTAL RESULTS

The proposed clinically-oriented deep-learning (CO-Deep) system was implemented in Matlab® 2016 and tested on Windows 10 platform on Core i7 CPU. The statistical analysis was performed through sensitivity (SE), specificity (SP), accuracy (ACC) and area under the receiver operating curve (AUC) on the dataset of 5200 dermoscopy images collected from different resources. In this selected dataset, the melanocytic and non-melanocytic lesions are in equal quantity to provide equal importance during testing and classification stages. For developing the CO-Deep system, the dataset is divided into 40% of training and 60% of testing through 10-fold cross validation test.

Some of the results are shown in Table 1 of the proposed COE-Deep system on 5200 melanocytic and non-melanocytic (MnM) skin lesions when diagnosis through digital dermoscopy images. This table describes the statistical analysis in terms of Sensitivity (SE), Specificity (SP), Accuracy (ACC), training errors (E) and area under the receiver operating curve (AUC). As a display in Table 1, the average values for SE of 92%, SP of 94%, ACC of 93%, AUC of 0.94 and E of 0.73 are obtained when tested on this dataset in the case of melanocytic skin lesions whereas in the case of non-melanocytic skin lesions, the SE of 88%, SP of 92%, ACC of 90%, AUC of 0.90 and E of 0.65 are achieved. From these results, it clears that the proposed COE-Deep system is getting significantly higher results in the case of melanocytic than non-melanocytic skin lesions. It is due to the fact that it is very difficult to recognize non-melanocytic lesions compared to melanocytic skin lesions.

TABLE I. THE AVERAGE VALUES OBTAINED BY COE-DEEP SYSTEM ON 5200 DERMOSCOPY IMAGES IN TERMS OF SENSITIVITY (SE), SPECIFICITY (SP), ACC: ACCURACY, E: TRAINING ERRORS AND AREA UNDER THE RECEIVER OPERATING CURVE (AUC)

No.	Results					
	Category dataset	SE <sup>a</sup>	SP <sup>b</sup>	ACC <sup>c</sup>	AUC <sup>d</sup>	E <sup>e</sup>
1.	Melanocytic	92%	94%	93%	0.94	0.73
2.	Non-melanocytic	88%	92%	90%	0.90	0.65
Average Results		90%	93%	91.5%	0.92	0.62

<sup>a</sup>. Sensitivity, <sup>b</sup> Specificity, <sup>c</sup> Accuracy, <sup>d</sup> Area under ROC curve, <sup>e</sup> Training errors

Therefore, according to limited knowledge, there is no effective study for differentiation between MnM skin lesions through deep-neural-network approach without the need of hand-crafted features and pre- or post-processing steps.

In the past studies, there was only one paper found [15], where the authors utilized domain expert knowledge of image processing and machine learning algorithms to perform this classification of MnM skin lesions but the system required lots of steps with pre- and post-processing stages. They represented classification results of melanocytic lesions on 548 lesions in terms of sensitivity of 98.0% and a specificity of 86.6% using a cross-validation test. These obtained results were mentioned on the small dataset and classifier may be over-fitted when applied on a large scale environment. Therefore, the proposed system is better compared to [15] in terms of large-scale applicability. Using the above-obtained results, it confirmed that the COE-Deep system based on the advanced deep learning algorithm is capable of classifying melanocytic and non-melanocytic skin lesions. This is the first and basic difficult step for dermatologists to draw a separate line between MnM skin lesions in the diagnosis process. As a result, the proposed method assists the clinical experts to draw this clear line.

The comparisons are also performed with the state-of-the-art deep-learning algorithms in terms of SE, SP, ACC, AUC and E-statistical analysis on this selected dataset. As calculated in Table 2, the convolutional neural network (CNN) with four layers on average obtained SE of 80%, SP of 84%, ACC of 82%, AUC of 0.81 and E of 0.75 values to different MnM skin lesions. If CNN is integrated with the softmax linear classifier then the recognition results are high significantly better. In the case of CNN and softmax classifiers, SE of 84%, SP of 88%, ACC of 86%, AUC of 0.87 and E of 0.73 values are achieved. In contrast with CNN, if the stack-based autoencoders (SAEs) are utilized then the SE of 85%, SP of 88%, ACC of 86.5%, AUC of 0.86 and E of 0.71 values on average are obtained. However, the significantly better results are obtained in the case of SAE and softmax linear classifiers.

TABLE II. THE AVERAGE COMPARISONS RESULTS WITH OTHER DEEP-LEARNING METHODS IN TERMS OF SENSITIVITY (SE), SPECIFICITY (SP), ACCURACY, E: TRAINING ERRORS AND AREA UNDER THE RECEIVER OPERATING CURVE (AUC) ON 5200 DERMOSCOPY IMAGES

No.	Results					
	Category dataset	SE <sup>a</sup>	SP <sup>b</sup>	ACC <sup>c</sup>	AUC <sup>d</sup>	E <sup>e</sup>
1.	CNN	80%	84%	82.0%	0.81	0.75
2.	CNN+ softmax	84%	88%	86.0%	0.87	0.73
3.	SAE	85%	88%	86.5%	0.86	0.71
4.	SAE+ softmax	89%	90%	89.5%	0.88	0.69
5.	COE-Deep	90%	93%	91.5%	0.92	0.62

<sup>a</sup> Sensitivity, <sup>b</sup> Specificity, <sup>c</sup> Accuracy, <sup>d</sup> Area under ROC curve, <sup>e</sup> Training errors

In that case, the SE of 89%, SP of 90%, ACC of 89.5%, AUC of 0.88 and E of 0.69 values on average are gained. But the higher significant results are obtained in the case of proposed COE-deep system when combined CNN, SAE and softmax classifiers to recognize melanocytic and non-melanocytic skin lesions.

All these above-mentioned results in Tables 1 and 2 were reported through 10-fold cross-validation test to classify MnM skin lesions. Fig. 3 has shown the corresponding receiving operating characteristic curve (ROC) for differentiation between MnM skin lesions. An area under the curve (AUC) shows the significant result of this COE-Deep system, which is greater than 0.5 compared to CNN and stack-based autoencoders (SAEs). The SAEs deep-learning algorithms are getting higher AUC value compared to CNN model but less than the proposed COE-Deep system. As displayed in Table 1, it can be noticed that in the case of melanocytic skin lesions, the best performance has been measured i.e., AUC: 0.94. This proposed system based on deep-learning algorithms significantly improves the performance with the average value of AUC: 0.92. It is because of designing an effective classification system through advanced concepts of deep-learning algorithms without focusing on features extraction and selection steps.

## V. CONCLUSIONS

A clinically-oriented expert system based on deep-learning (COE-Deep) algorithms is presented in this paper to automatically differentiate between melanocytic and non-melanocytic (MnM) skin lesions. The convolutional neural network (CNN) is employed to extract deep features and then most discriminative features are selected by stack-based autoencoders (SAEs) model. Finally, the recognition of decision is performed by Softmax linear classifier. On 5200 clinical dermoscopy images, the statistically significant results were obtained in terms of sensitivity (SE), specificity (SP), accuracy (ACC) and area under the receiver operating curve (AUC) when used 10-fold cross validation test. On average, the 90% of SE, 93% of SP, 91.5% of ACC and 0.92 of AUC values are obtained. Hence, the proposed COE-Deep system is best suited for classification of non-melanocytic skin lesions should improve the accuracy, reliability, and accessibility of pigmented skin lesions screening system. In the future work, this effort much added to get more accurate and an improved accuracy.

## REFERENCES

- [1] Barata, C., Celebi, M. E., & Marques, J. S. Development of a clinically oriented system for melanoma diagnosis. *Pattern Recognition*, 69, 270-285, 2017.
- [2] Daniel Ruiz, Vicente Berenguer, Antonio Soriano, Belén Sánchez, A decision support system for the diagnosis of melanoma: A comparative approach, *Expert Systems with Applications*, 38(12), pp.15217–15223, 2011.
- [3] Stephan Dreiseitl, Michael Binder, Do physicians value decision support? A look at the effect of decision support systems on physician opinion, *Artificial Intelligence in Medicine*, 33(1), pp. 25–30, 2005.
- [4] Laura K. Ferris, Jan A. Harkes, Benjamin Gilbert, Daniel G. Winger, Kseniya Golubets, Oleg Akilov, Mahadev Satyanarayanan, Computer-aided classification of melanocytic lesions using dermoscopic images, *Journal of the American Academy of Dermatology*, 73(5), pp. 769–776, 2015.

- [5] William V. Stoecker, Mark Wronkiewicz, Raed Chowdhury, R. Joe Stanleyb, Jin Xu, Austin Bangert, Bijaya Shrestha, David A. Calcara, Harold S. Rabinovitz et al. Detection of granularity in dermoscopy images of malignant melanoma using color and texture features, *Computerized Medical Imaging and Graphics*, 35, pp. 144–147, 2011.
- [6] Nabin K. Mishra, M. Emre Celebi, An Overview of Melanoma Detection in Dermoscopy Images Using Image Processing and Machine Learning, *Computer Vision and Pattern Recognition*, 2016.
- [7] Unlu E, Akay BN, Erdem C. Comparison of dermatoscopic diagnostic algorithms based on calculation: The ABCD rule of dermatoscopy, the seven-point checklist, the three-point checklist and the CASH algorithm in dermatoscopic evaluation of melanocytic lesions. *J Dermatol.* 2014 Jul; 41(7):598-603. doi: 10.1111/1346-8138.12491. Epub 2014.
- [8] Catarina Barata, M. Emre Celebi, Jorge S. Marques, Jorge Rozeir, Clinically inspired analysis of dermoscopy images using a generative model, *Computer Vision and Image Understanding*, Volume 151, October 2016, Pages 124–137.
- [9] Premaladha, J.; Ravichandran, K.S. Novel Approaches for Diagnosing Melanoma Skin Lesions through Supervised and Deep Learning Algorithms. *J. Med. Syst.* 2016, 40, 96.
- [10] Codella, N.; Cai, J.; Abedini, M.; Garnavi, R.; Halpern, A.; Smith, J.R. Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images, *Machine Learning in Medical Imaging*. In Proceedings of the 6th International Workshop, MLMI 2015, Munich, Germany, 5 October 2015; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9352, pp. 118–126.
- [11] Md. Mahmudur Rahman; Nuh Alpaslan, Automated melanoma recognition in dermoscopic images based on extreme learning machine (ELM), *Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis*, 1013414 (March 3, 2017).
- [12] Sergey Demyanov, Rajib Chakravorty, Mani Abedini, Alan Halpern, Rahil Garnavi, Classification of dermoscopy patterns using deep convolutional neural networks, 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), 364-366, 2016.
- [13] Jeremy Kawahara, Aicha BenTaieb and Ghassan Hamarneh, Deep features to classify skin lesions, 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016.
- [14] Fornaciali, M.; Carvalho, M.; Bittencourt, F.V.; Avila, S.; Valle, E. Towards Automated Melanoma Screening: Proper Computer Vision & Reliable Results. *Comput. Vis. Pattern Recognit.* 2016.
- [15] H. Iyatomi, K. A. Norton, M. E. Celebi, G. Schaefer, M. Tanaka and K. Ogawa, "Classification of melanocytic skin lesions from non-melanocytic lesions," 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, 2010, pp. 5407-5410.
- [16] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun, Dermatologist-level classification of skin cancer with deep neural networks, *Nature* 542, 115–118.
- [17] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." *Neural networks* 61 (2015): 85-117.
- [18] Argenziano G, Soyer HP, De Giorgi V, Piccolo D, Carli P, Delfino M et al. *Interactive atlas of dermoscopy CD*. Milan: EDRA Medical Publishing and New Media, 2000.
- [19] University of Auckland, New Zealand. Dermatologic image, database. <http://dermnetnz.org/doctors/dermoscopy-course/>. [Access date: 15/3/2016].
- [20] Matt Berseth, ISIC 2017: Skin Lesion Analysis towards Melanoma Detection, *Computer Vision and Pattern Recognition*, <http://isdis.net/isic-project/>, [Access date: 1/1/2017].

# Feature Selection and Extraction Framework for DNA Methylation in Cancer

Abeer A. Raweh

Dept. of Computer Science,  
Faculty of Computers and  
Information, Cairo University  
Cairo, Egypt

Mohammad Nassef

Dept. of Computer Science,  
Faculty of Computers and  
Information, Cairo University  
Cairo, Egypt

Amr Badr

Dept. of Computer Science,  
Faculty of Computers and  
Information, Cairo University  
Cairo, Egypt

**Abstract**—Feature selection methods for cancer classification are aimed to overcome the high dimensionality of the biomedical data which is a challenging task. Most of the feature selection methods based on DNA methylation are time consuming during testing phase to identify the best pertinent features subset that are relevant to accurate prediction. However, the hybridization between feature selection and extraction methods will bring a method that is far fast than only feature selection method. This paper proposes a framework based on both novel feature selection methods that employ statistical variation, standard deviation and entropy, along with extraction methods to predict cancer using three new features, namely, Hypomethylation, Midmethylation and Hypermethylation. These new features represent the average methylation density of the corresponding three regions. The three features are extracted from the selected features based on the analysis of the methylation behavior. The effectiveness of the proposed framework is evaluated by the breast cancer classification accuracy. The results give 98.85% accuracy using only three features out of 485,577 features. This result proves the capability of the proposed approach for breast cancer diagnosis and confirms that feature selection and extraction methods are critical for practical implementation.

**Keywords**—DNA methylation, feature selection; feature extraction; cancer classification; epigenetics; biomarkers; hypomethylation; hypermethylation; methylation

## I. INTRODUCTION

Cancer is a leading cause of death worldwide, it begins when some cells in a part of the body start to grow out of control. Despite the presence of more than one type of cancer that differ in the way of growing cells and spreading, the development of all these kinds is driven by “genetic alterations” and “epigenetic changes” of the DNA genome [1]. Recent research increases evidences that the epigenetic modifications play a critical role in human cancer. These modifications are heritable changes in a cellular phenotype that are independent of alterations in the DNA sequence [2], [3]. Many studies of epigenetic aberrations in tumors prove that the biology of DNA methylation is the most potential epigenetic marker for cancer detection in spite of many other epigenetic alterations in the mammalian genome such as post-translational modifications of histones, chromatin remodeling and microRNAs patterns [4]. Actually, DNA methylation acts as a gene-silencing mechanism to turn off specific genes due to its significant effects on gene expressions and the architecture of the nucleus of the cell [5]. Chemically, DNA

methylation is a relatively stable chemical modification resulting from the addition of a methyl (CH<sub>3</sub>) group at the carbon 5 position of the cytosine or guanine nucleotides in the context of 5'-CG-3' (CpG dinucleotide) by DNA methyltransferase (DNMT) enzymes [6]. Not all CpG sites in the genome are methylated; CpG islands “regions that are containing a high frequency of CpG dinucleotides” are usually not methylated in normal cells [7]. Throughout the genome, there are two types of cancer-associated DNA methylation based on the methylation level called hypermethylation and hypomethylation. Hypermethylation “the methylation exceeds normal methylation level” of tumor suppressor gene affecting the gene expression and proteins involved in cancer manifestation. On the other hand, hypomethylation “the methylation beneath normal methylation level” has been observed frequently in solid tumors [8].

Due to the huge number of probes in the DNA, the importance of providing researchers and scientists with novel, accurate and robust computational tools for studying the whole genome for the cancer predication is widely increasing. Most of the probes of the mammalian tumors genome are irrelevant classification factors and may have bad effect by introducing noises and hence decreasing predication accuracy [9]. Ideally, a good dimensionality reduction method should eliminate these irrelevant probes while at the same time retain all the highly discriminative probes. Therefore, using feature selection and extraction techniques in cancer predication becomes essential to identify the informative probes that underlie the pathogenesis of tumor cell proliferation. Thus, many recent researches applied feature selection and extraction techniques to extract useful information and diagnosis the tumor [10]-[15].

In this paper, we propose a framework based on feature selection and extraction methods, to rid of irrelevant information and improve cancer classification accuracy based on DNA methylation data. First, a novel feature selection based on statistical variation and standard deviation is utilized for identifying the small set of discriminative methylated DNA probes, afterwards, the average methylation density of three regions (hypomethylation, midmethylation and hypermethylation) is calculated as new extracted features to predict cancer.

The reminder of this paper is organized as follows. Section II elaborates on previous work, Section III presents

the attempted dataset and proposed framework, Section IV discusses our experimental results and the last Section V contains concluding remarks and demonstrates future work.

## II. RELATED WORKS

To increase the accuracy and handle the dramatically increasing tumor feature data and information, a number of researchers have turned to feature selection and extraction techniques for predicting cancer. Feature selection (FS) is one of the important steps in classification modeling of cancer based on DNA methylation [16], it could be used for eliminating unnecessary information to reduce the high dimensionality of the data. Whereas feature extraction also called data transformation, is the process of transforming the feature data into a quantified data type instead of recognizing new patterns to represent the data.

In the past decade, many feature selection and extraction methods have been proposed, resulting in great improvements of classification. Li *et al.* [10] proposed a gene extraction method by using two standard feature extraction methods, namely, the T-test method and kernel partial least squares (KPLS) in tandem. Zheng *et al.* [11] developed a hybrid of K-means and support vector machine (K-SVM) algorithms to diagnosis breast cancer disease. Kopriva *et al.* [12] proposed a general feature extraction method for cancer prediction based on the linear transformation constructed by tensor decomposition. A novel method using wavelet analysis, genetic algorithm, and Bayes classifier proposed by Liu *et al.* [13] was applied to detect the prognostic biomarkers of survival in colorectal cancer patients. Fontes *et al.* [14] applied feature extraction techniques such as *F-score*, *p-value rank* and *wrapper approaches* in order to identify which probes presented higher significance in breast cancer prediction. D.L. Tong [15] proposed an innovative hybridized model based on genetic algorithms (GAs) and artificial neural networks (ANNs), to extract the highly differentially expressed genes for specific cancer pathology. Anuradha *et al.* [17] gave a comparative study to identify the best feature extraction technique to classify Oral cancers. Zhuang *et al.* [16] performed another good comparison study of feature selection and classification methods in DNA using the Illumina Infinium platform. Cai *et al.* [18] used Ensemble-based feature extraction methods to capture the unbiased, informative as well as compact molecular signatures followed by SVM trained with Incremental Feature Selection (IFS) strategy to predict subtypes of lung cancer. A novel multiclass feature selection and classification system proposed by Sebastian *et al.* [19] for data merged from different molecular biomedical techniques demonstrated that the feature selection step is crucial in high dimension data classification problems. Furthermore, Baur *et al.* [20] developed a feature selection algorithm based on sequential forward selection to compute gene centric DNA methylation using probe level DNA methylation data. Valavanis *et al.* [8] used semantics information included in the Gene Ontology (GO) tree by graph-theoretic methodology in order to select cancer epigenetic biomarkers.

## III. PROPOSED FRAMEWORK

### A. Dataset

In this study, we conducted experiments on a dataset of large collection of cancer methylomes obtained from The Cancer Genome Atlas (TCGA) using the Human Infinium 450k assay for 4034 cancer and normal tissue samples. The dataset was downloaded from Max Planck Institute for Informatics (MPI) with a software tool for large-scale analysis that yields detailed hypertext reports and interpretation of the DNA methylation data “RnBeads” [21]. As listed in Table 1, the dataset contains several types of cancer: blood, breast, intestinal, brain and other types of cancer. The degree of DNA methylation that extracted from the regions: 31195 promoters, 31033 genes, 485577 probes and 26662 CpG Islands quantified numerically as values.

### B. Proposed Framework

The proposed framework is made for detecting cancer based on methylated DNA probes, there are three main steps to be followed in this framework. These steps are feature selection, feature extraction and classification. Fig. 1 shows the architecture of the proposed framework.

### C. Feature Selection Methods

Feature selection methods in cancer classification issues are aimed at identifying the minimal-sized subset of markers that are relevant to accurate prediction. To achieve this target, we propose two novel feature selection methods. The first one uses statistical variation in terms of standard deviation in order to select the most informative probes which distinguish normal tissue from cancer. This method measures the differences of probe methylation in all samples compared with the dispersion of this probe methylation in each class (Normal, Cancer) separately. Thus, the discriminative value (*DV*) according to the proposed feature selection for each probe (*X*) based on DNA methylation as an input is defined as:

$$DV1(X) = \frac{\frac{\sum(x-\bar{x})^2}{n-1}}{\sqrt{\frac{\sum(x^+-\bar{x}^+)^2}{n^+-1}} + \sqrt{\frac{\sum(x^--\bar{x}^-)^2}{n^- -1}}} \quad (1)$$

TABLE I. CANCER TYPES IN THE ATTEMPTED DATASET

Cancer Type	No. of Normal Samples	No. of Tumor Samples
Breast invasive carcinoma	98	573
Colon adenocarcinoma	38	253
Glioblastoma multiforme	1	125
Head and Neck squamous cell carcinoma	50	373
Kidney renal clear cell carcinoma	160	283
Acute Myeloid Leukemia	0	194
Lung adenocarcinoma	32	409
Lung squamous cell carcinoma	42	360
Rectum adenocarcinoma	7	96
Thyroid carcinoma	56	435
Uterine Corpus Endometrioid Carcinoma	46	393

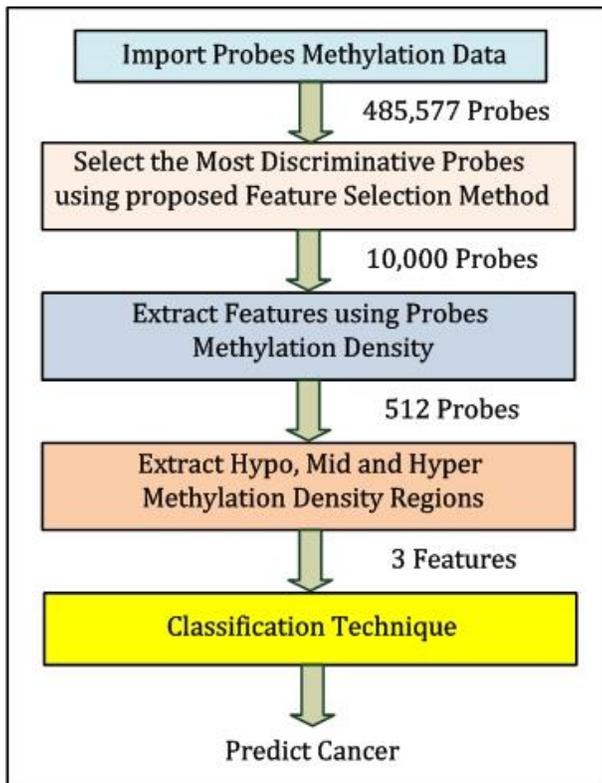


Fig. 1. Architecture of the proposed framework.

Where:

- $\bar{x}$  is the methylation average of entire dataset samples for probe  $(X)$ .
- $\bar{x}^+$ ,  $\bar{x}^-$  are the methylation average of the cancer and normal samples respectively for probe  $(X)$ .
- $x$  is the methylation of entire dataset samples for probe  $(X)$ .
- $x^+$ ,  $x^-$  are the methylation of the cancer and normal samples respectively for probe  $(X)$ .
- $n$  is the number of all samples.
- $n^+$ ,  $n^-$  are the number of cancer and normal samples respectively.

The second feature selection method is proposed to find the more variational features with less amount of uncertainty involved in its values (less disorder features). The key measure in information theory for measuring uncertainty is the “entropy” which is defined by Claude E. Shannon [22], [23] and considered as a measure to rank features. Regard to this, the above formula  $DVI(X)$  with entropy is defined as:

$$DV2(X) = \frac{DV1(X)}{H(Y|X)} \quad (2)$$

Where:

$H(Y|X)$  is the entropy for two variables  $X$  and  $Y$  that measures the uncertainty of  $Y$  when  $X$  is known.

$$H(Y|X) = - \sum p(x) \sum p(y|x) \log_2(p(y|x)) \quad (3)$$

Where:

$Y$  denotes all available classes (Normal and Cancer).  
 $X$  is the methylation of gene promoter.

$p(x)$  is the probability of interval  $x$   
 $p(y|x)$  is the probability of class  $y$  given interval  $x$ .

From 485,577 probes, 10,000 probes are selected using the proposed feature selection methods.

#### D. Feature Extraction Method

The most discriminative probes (i.e. 10,000 probes) are selected using the proposed feature selection  $DVI(X)$ . Then these features are extracted using feature extraction methods. Feature extraction is the process which involves for clarifying and detecting the methylation patterns or methylation behavior in the selected probes. As a first step, we use kernel density estimator method [24]; which infers population probability density function of the selected probes; as a feature extraction method, in order to extract 512 features for each sample from the selected 10,000 probes. The kernel density estimate of  $f$  at the point  $x$  is given by

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (4)$$

Where  $K$  denotes to so-called Gaussian kernel function that integrates to one and has mean zero. It defined as:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u^2\right) \quad (5)$$

And  $h$  denotes to a smoothing parameter  $>0$  called the bandwidth. The optimal bandwidth that gives better results can be obtained by

$$h_{opt} = \frac{0.9 X \partial}{\sqrt[5]{N}} \quad (6)$$

Where,  $\partial = \min\left(\partial, \frac{IQR}{1.34}\right)$  and  $IQR$  is the interquartile range that measures the difference between the 75<sup>th</sup> percentile ( $Q3$ ) and the 25<sup>th</sup> percentile ( $Q1$ ):  $IQR = Q3 - Q1$ .

In the second step, for each sample we extract three features from 512 features of kernel density method that have been obtained. The extracted three features are belonging to average methylation density of three regions: Hypomethylation, Middle-methylation (Midmethylation) and Hypermethylation region.

#### E. Classification

To evaluate the ability of the proposed framework for cancer classification based on methylated probes, the following classifiers: Naïve Bayes, Random Forest, Hoeffding Tree, SVM and Simple Logistic were used. The accuracy, F-Measure, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) of each classifier were used as a metrics for evaluation. 250 samples from breast tissue were used as training data and 348 samples were used as testing data. Furthermore, different approaches were used to study classifier’s ability in cancer prediction, where the first experiment used the methylation density of whole probes (485,577 probes), the second experiment used methylation density of most discriminative probes chosen by  $DVI(X)$  (10,000 probes) and the last experiment used three features only “average methylation density of three regions (Hypo, Mid, Hyper methylation)”. The next section shows the testing accuracy, F-Measure, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) of each machine learning

technique. Through these experiments, the reader can observe the ability of classifier in cancer prediction using only the extracted three features.

IV. RESULTS ANALYSIS AND DISCUSSION

Firstly, this section compares the proposed feature selection methods,  $DVI(X)$  and  $DV2(X)$ , with the existing feature selection methods such as: F-Score, Chi-Squared, Information Gain, and Symmetrical Uncertainty (SU) to evaluate their ability to select the most discriminative probes for cancer classification. To ensure a fair comparison, we conduct the experiments on breast tissue which contains the maximum number of samples in the dataset as illustrated in Table 1. For the breast tissue dataset, 250 samples were used as training data whereas 348 samples were used as testing data. Tables 2 to 4 reports the testing accuracies, F-Measure, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) of some machine learning techniques such as: Naïve Bayes, Random Forest, Hoeffding Tree, SVM and Simple Logistic for 31 selected probes. The results show that the proposed methods,  $DVI(X)$  and  $DV2(X)$ , always outperform the existing feature selection methods in terms of the predication accuracy.

TABLE II. PREDICTION ACCURACY OF DIFFERENT CLASSIFIERS BASED ON DIFFERENT FEATURE SELECTION METHODS

Classification Techniques FS Methods	Naïve Bayes	Random Forest	Hoeffding Tree	SVM	Simple Logistic
Proposed Method $DVI(X)$	98.85%	99.14%	98.85%	98.85%	98.56%
Proposed Method $DV2(X)$	99.43%	99.14%	99.43%	99.43%	98.28%
F-Score	98.28%	98.56%	98.28%	97.7%	97.7%
Chi- Squared	98.28%	98.85%	98.28%	98.56%	97.13%
Information Gain	97.99%	98.28%	97.99%	98.85%	96.84%
SU	98.85%	98.85%	98.85%	98.28%	96.55%

TABLE III. F-MEASURE OF DIFFERENT CLASSIFIERS BASED ON DIFFERENT FEATURE SELECTION METHODS

Classification Techniques FS Methods	Naïve Bayes	Random Forest	Hoeffding Tree	SVM	Simple Logistic
Proposed Method $DVI(X)$	95.9%	96.8%	95.9%	95.9%	94.8%
Proposed Method $DV2(X)$	97.9%	96.9%	97.9%	97.9%	93.3%
F-Score	93.6%	94.8%	93.6%	91.7%	91.7%
Chi- Squared	94%	95.9%	94%	94.9%	90.4%
Information Gain	93.1%	93.5%	93.1%	95.8%	88.4%
SU	95.9%	95.7%	95.9%	94%	88.7%

TABLE IV. MEAN ABSOLUTE ERROR (MAE) AND ROOT MEAN SQUARED ERROR (RMSE) OF DIFFERENT CLASSIFIERS BASED ON DIFFERENT FEATURE SELECTION METHODS

Classification Techniques FS Methods		Naïve Bayes	Random Forest	Hoeffding Tree	SVM	Simple Logistic
Proposed Method $DVI(X)$	MAE	0.01	0.04	0.01	0.01	0.02
	RMSE	0.1	0.1	0.1	0.1	0.09
Proposed Method $DV2(X)$	MAE	0.005	0.03	0.005	0.005	0.02
	RMSE	0.07	0.1	0.07	0.07	0.12
F-Score	MAE	0.01	0.03	0.01	0.02	0.06
	RMSE	0.13	0.11	0.13	0.15	0.14
Chi-Squared	MAE	0.01	0.03	0.01	0.01	0.1
	RMSE	0.13	0.1	0.13	0.11	0.15
Information Gain	MAE	0.01	0.05	0.01	0.01	0.04
	RMSE	0.13	0.13	0.13	0.1	0.14
SU	MAE	0.01	0.03	0.01	0.01	0.19
	RMSE	0.1	0.1	0.1	0.13	0.22

Furthermore, to demonstrate the ability of the proposed framework for cancer classification based on methylated probes, the following Tables 5 to 7 reports the testing accuracy, F-Measure, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) of different machine learning techniques. These tables compares the results of three approaches: the first one when using the whole probes density, the second one when using the density of 10,000 Probes choosing by  $DVI(X)$ , and the third one when using the three extracted features (average density of Hypo, Mid and Hyper regions). The results prove the capability of the proposed approach in cancer prediction using only three extracted features.

In addition, this section makes an analysis and comparison of the behavior of the valuable data in probe regions “DNA methylation” in breast tissue samples (normal and cancer). Fig. 2 shows the average methylation of 98 normal samples and 500 cancer samples in the whole probes “485577 probes”.

TABLE V. COMPARISON OF ACCURACY OBTAINED BY DIFFERENT CLASSIFIERS BASED ON DIFFERENT APPROACHES

Classifier Approach	Naïve Bayes	Random Forest	Hoeffding Tree	SVM	Simple Logistic
Whole Probes Density	80.17%	87.36%	79.89%	83.05%	82.76%
Density of 10,000 Probes choosing by $DVI(X)$	98.56%	97.70%	98.56%	97.70%	97.70%
Average density of Hypo, Mid and Hyper regions	98.28%	98.56%	98.28%	98.85%	98.28%

TABLE VI. F-MEASURE OBTAINED BY DIFFERENT CLASSIFIERS BASED ON DIFFERENT APPROACHES

Classifier Approach	Naïve Bayes	Random Forest	Hoeffding Tree	SVM	Simple Logistic
	Whole Probes Density	47.3%	42.1%	47%	27.2%
Density of 10,000 Probes choosing by $DVI(X)$	94.6%	91.5%	94.6%	91.5%	91.1%
Average density of Hypo, Mid and Hyper regions	93.5%	94.6%	93.5%	95.7%	93.3%

TABLE VII. MEAN ABSOLUTE ERROR (MAE) AND ROOT MEAN SQUARED ERROR (RMSE) OBTAINED BY DIFFERENT CLASSIFIERS BASED ON DIFFERENT APPROACHES

Classifier Approach	Whole Probes Density		Density of 10,000 Probes choosing by $DVI(X)$		Average density of Hypo, Mid and Hyper regions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Naïve Bayes	0.19	0.44	0.01	0.11	0.01	0.12
Random Forest	0.17	0.29	0.03	0.13	0.03	0.13
Hoeffding Tree	0.2	0.44	0.01	0.11	0.01	0.12
SVM	0.16	0.41	0.02	0.15	0.01	0.1
Simple Logistic	0.21	0.35	0.04	0.13	0.04	0.13

It is clear that the methylation behavior can be divided into three regions: low level of methylation region “hypomethylation”, middle level of methylation region “midmethylation” and high level of methylation region “hypermethylation”. This figure demonstrates that there is a difference between methylation behavior in normal and cancer samples, where the density of methylation in normal samples are lower in cancer samples. This difference, however, is not totally clear. Moreover, as shown in Table 5, the Random Forest classifier gave 87.36% as a higher prediction accuracy using the density of whole probes approach.

For a deep dive into the difference between methylation behavior in normal and cancer samples, we concentrated on the most informative probes that are relevant to accurate cancer prediction. Fig. 3 shows the average methylation of the most discriminative probes (10,000 Probes choosing by  $DVI(X)$ ) in all normal and cancer samples. As shown in this figure, the difference is more clearly, where the density of hypomethylation and hypermethylation are lower in cancer samples. The decreasing in density of hypomathylation in the cancer sample means that, the amount of methylation is increased in these regions, and thus all the respective genes are turned from active genes to silent genes. By contrast, the decreasing density of hypermathylation in a cancer sample means decreasing amount of methylation; therefore all the respective genes in these regions are turned from silent genes to active genes. Furthermore, using the density of discriminative probes “10,000 probes” in cancer prediction improves classifier accuracy, where both Naïve Base and Hoeffding Tree classifier gave 98.56% as a higher prediction accuracy using this approach. Moreover, Fig. 4 compares the behavior of methylation in cancer cell in some other tissues such as: Colon, Kidney and Uterine.

Probes Methylation in Breast Tissue

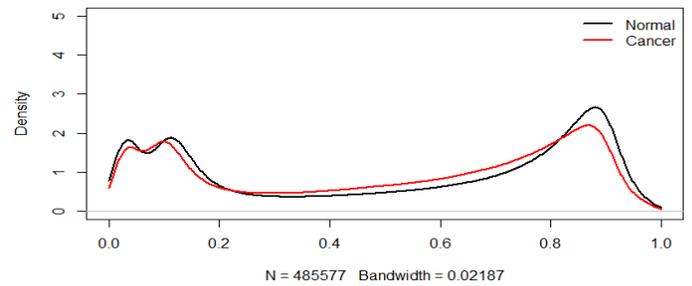


Fig. 2. Whole probes methylation in Breast tissue.

Discriminative Probes in Breast Cancer

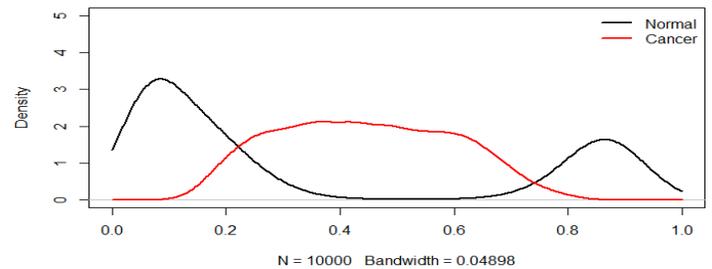


Fig. 3. Methylation of the most discriminative probes (10,000 Probes choosing by  $DVI(X)$ ).

We found that the behavior of methylation is the same in all tissues, increasing methylation of hypomethylation and decreasing methylation of hypermethylation.

As we mentioned in our experiments, we extracted three features from 512 features of kernel density estimator method. These three features belong to average methylation density of three regions: hypomethylation, midmethylation and hypermethylation region. To obtain these features, we calculated the intersection points between normal and cancer curve. As shown in Fig. 5, 0.223092 and 0.741683 are intersection points between the curves, and thus, the curves can be divided into hypomethylation, midmethylation and finally hypermethylation region. Fig. 5 shows the intersection points and these three regions, where letter *A* denotes to hypomethylation region, letter *B* denotes to midmethylation region and letter *C* denotes to hypermethylation region. In addition, as shown in Table 5, using these three features out of 485577 features “probes” in cancer prediction improves classifier accuracy (from 83.05% to 98.85%), for SVM classifier which gave a higher accuracy using this approach. These results emphasize the capability of our proposed framework in cancer classification and illustrate the importance of using feature selection and extraction for accurate cancer prediction.

To provide a better understanding of the DNA methylation mechanism that plays a major role in the development and progression of cancer, we analyze the top 31 probes that have been generated from the proposed feature selection methods ( $DVI$  and  $DV2$ ) and used in the classification experiments.

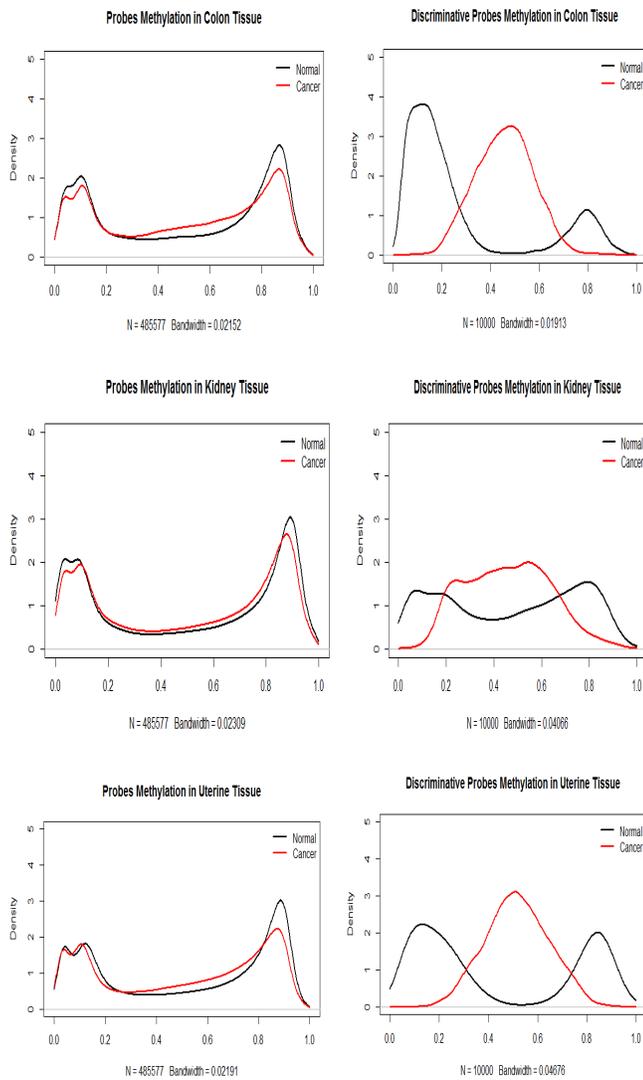


Fig. 4. Methylation behavior in Colon, Kidney and Uterine tissues.

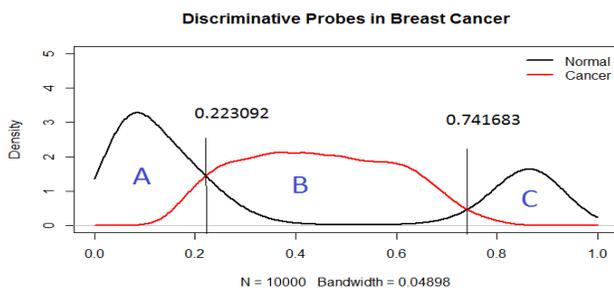


Fig. 5. Intersection points and Hypomethylation, Midmethylation, Hypermethylation regions.

Therefore, we confirm that the role of DNA methylation is to activate or silence some genes by decreasing or increasing their methylation respectively. Furthermore, we examine the ability of a new subset of probes to predict cancer, the subset contains common probes from the top 31 probes subset that have been selected by the proposed *DVI* and *DV2* methods

“intersection subset”. The accuracy values obtained by Naïve Bayes, Random Forest, Hoeffding Tree, SVM and Simple Logistic classifier using this subset are: 99.13%, 97.98%, 99.13%, 96.83% and 96.55%, respectively. These results show that cancer classification achieves lower predication accuracy than *DVI* or *DV2* or both due to missing information in intersection subset, and thus we confirm that the DNA methylation has several patterns that play significant role in human cancer. There is no single probes subset to identify these patterns and each feature selection method can provide different probes subset.

## V. CONCLUSION AND FUTURE WORK

Feature selection and extraction are of vital importance for accurate cancer classification, by skipping unnecessary information that introduce noises and decrease predication accuracy. This article proposes a framework based on novel feature selection methods along with extraction methods, to identify the informative probes that underlie the pathogenesis of tumor cell proliferation and improve cancer classification accuracy. The proposed feature selection method *DVI* uses statistical variation in terms of the standard deviation for obtaining the discriminative value while the other proposed feature selection method *DV2* uses entropy to rank features and hence obtains the more variational features with lower amount of uncertainty involved in its values. First, our framework uses *DVI* to identify the good marker probes subset, afterwards, in order to predict cancer, the average methylation density of three regions: hypomethylation, midmethylation and hypermethylation is calculated from the selected methylated probes as new features. The effectiveness of the proposed framework is evaluated by the breast cancer classification accuracy in probe regions, where the results are evidence that, our proposed framework has the ability to predict cancer using only three features out of 485577 features. As an example, SVM classifier gives 98.85% as higher prediction accuracy, and this highlights the importance of using feature selection and extraction methods in cancer classification issues based on DNA methylation.

Furthermore, observing probes subsets that have been selected from different feature selection methods confirmed that DNA methylation has several patterns and there is no single probes subset to identify these patterns. The results highlight the difference in methylation’s behavior between the normal and abnormal samples in probes regions, and this difference confirms that the role of DNA methylation in cancer is to activate or silence some genes by decreasing or increasing their methylation respectively.

A new future work is identified based on the current study. We plan to improve the formula of feature extraction method instead of the current formula “average methylation density”, to obtain higher results and improve cancer classification accuracy.

## REFERENCES

- [1] Z. Herceg and P. Hainaut, “Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis,” *Mol Oncol*, vol. 1, pp. 26–41, 2007.
- [2] M. Dawson and T. Kouzarides, “Cancer epigenetics: From mechanism to therapy,” *Cell*, vol. 150, Issue 1, pp. 12–27, 2012.

- [3] X. Ma and X. Gao, "Epigenetic Modifications and Carcinogenesis of Human Endometrial Cancer," *Austin Journal of Clinical Pathology*, vol. 3, pp. 1-9, 2014.
- [4] T. Mikeska and J. Craig, "DNA Methylation Biomarkers: Cancer and Beyond," *Genes*, vol. 5, pp. 821-864, Sep 2014.
- [5] Y. Ma, X. Wang and H. Jin, "Methylated DNA and microRNA in body fluids as biomarkers for cancer detection," *Int J Mol Sci*, vol. 14, pp. 10307-31, 2013.
- [6] M. Pouliot, Y. Labrie, C. Diorio and F. Durocher, "The Role of Methylation in Breast Cancer Susceptibility and Treatment," *Anticancer Res*, pp. 4569-74, September 2015.
- [7] K. Williams, J. Christensen, and K. Helin, "DNA methylation: TET proteins – guardians of CpG Islands?," *EMBO Rep*, 13:pp. 28-35, 2012.
- [8] I. Valavanis, E. Pilalis, P. Georgiadis, S. Kyrtopoulos and A. Chatziioannou, "Cancer Biomarkers from Genome-Scale DNA Methylation: Comparison of Evolutionary and Semantic Analysis Methods," *Microarrays (2076-3905)*, vol. 4, Issue 4, p647, December 2015.
- [9] J. Li, H. Su and H. Chen, "Optimal Search Based Gene Selection for Cancer Prognosis," in *Proc. 11th Americas Conference on Information Systems*, 2005.
- [10] S.T. Li, C. Liao, and J.T. James, "Gene Feature Extraction Using T-Test Statistics and Kernel Partial Least Squares," *Springer*, vol. 4234, pp. 11-20, 2006.
- [11] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1476-1482, 2014.
- [12] I. Kopriva, A. Jukić, and A. Cichocki, "Feature extraction for cancer prediction by tensor decomposition of 1D protein expression levels," in *Proc. IASTED Conference on Computational Bioscience CompBio*, pp. 277-283, July 2011.
- [13] Y. Liu, U. Aickelin, J. Feyereisl, and L.G. Durrant, "Biomarker CD46 Detection in Colorectal Cancer Data based on Wavelet Feature Extraction and Genetic Algorithm," *Knowledge Based Systems*, 2012.
- [14] C. M. Fontes, R. Natowicz, R. Rouzier, and A. P. Braga, "Isolation of DNA Probes and Their Correlation with Neoadjuvant Chemotherapy outcome for Breast Cancer," In: *Congresso Brasileiro de Automática*, 2011.
- [15] D.L. Tong, "Genetic Algorithm-Neural Network: Feature Extraction for Bioinformatics Data," *Bournemouth University Research Online*, 2010.
- [16] J. Zhuang, M. Widschwendter, and A. Teschendorff, "A comparison of feature selection and classification methods in DNA methylation studies using the Illumina Infinium platform," *BMC Bioinformatics*, Apr 2012.
- [17] K. Anuradha, and K. Sankaranarayanan, "Comparison of Feature Extraction Techniques to classify Oral Cancers using Image Processing," *International Journal of Application or Innovation in Engineering*, 2013.
- [18] Z. Cai, D. Xu, Q. Zhang, J. Zhang, et al, "Classification of lung cancer using ensemble-based feature selection and machine learning methods," *Mol. BioSyst*, vol. 11, pp. 791 -800, 2015.
- [19] S. Sebastian, P. Justyna, and F. Krzysztof, "Multiclass Classification Problem of Large-Scale Biomedical Meta-Data," *Procedia Technology*, vol. 22, pp. 938-945, 2016.
- [20] B. Baur and S. Bozdag, "A Feature Selection Algorithm to Compute Gene Centric Methylation from Probe Level Methylation Data," *PLOS ONE*, Edited by Jianhua Ruan, vol. 11, issue 2, Feb 2016.
- [21] <http://rnbeads.mpi-inf.mpg.de/>
- [22] V. Canedo, N. Maroño and A. Betanzos, "Feature Selection for High-Dimensional Data," *Springer*, 2015.
- [23] T. Cover, J. Thomas, "Elements of Information Theory 2nd Edition," *Wiley Series in Telecommunications and Signal Processing*, September 2006.
- [24] S.J. Sheather, "Density Estimation," *Statistical Science*, vol. 19, no. 4, pp. 588-597, 2004.

# Research Pathway towards MAC Protocol in Enhancing Network Performance in Wireless Sensor Network (WSN)

Anitha K

Research Scholar  
Visvesvaraya Technological University  
Department of Computer Science & Engineering  
RajaRajeswari College of Engineering  
Bangalore, India

Usha S

Professor & Head  
Department of Computer Science Engineering  
RajaRajeswari College of Engineering  
Bangalore, India

**Abstract**—The applications and utility of Wireless Sensor Network (WSN) have increased its pace in making an entry to the commercial market since the last five years. It has successfully established its association with Internet-of-Things (IoT) and other reconfigurable networks. However, in this advent of exponential progress in technology, WSN still suffers from elementary problems of energy efficiency, scalability, delay, and latency where Medium Access Control (MAC) protocols hold the primary responsibility. This paper reviewed the frequently used MAC protocols and studied their advantages and limitations followed by most recently carried out implementation work towards WSN performance enhancement. The paper finally outlines the unsolved problems from the existing research work and discussed briefly the research gap followed by a chalked out plan of tentative future work to address the research gap from existing review.

**Keywords**—Delay; energy issues; latency; MAC protocol; scalability; Wireless Sensor Network (WSN)

## I. INTRODUCTION

A collaboration of small sensor nodes (SN) distributed in a region to monitor the specific atmospheric parameters is considered as wireless sensor network (WSN) [1]. An SN is an extremely delicate and little electronic device which has very limited computational capability and resource availability [2]. The WSN applications are found in e.g. warehouse to habitat monitoring system. An important thing to understand in the applications of the sensor network is that there is two types of applications e.g. time critical application and mission critical application. Time critical applications are those who have constraints of time bound, but mission critical applications have multiple constraints. The majority of the real-time applications for emergency situation call for timeliness message transmission. The problem starts shooting up when it is the dynamic network. Hence, the node dissipates maximum energy to ensure sending the data packet. Unfortunately, if the node dissipates too much energy not only its network lifetime minimizes but its communication performance also degrades significantly. Also, the majority of the other problems e.g. routing problems, security problems, and load balancing problems in WSN arises from energy issue itself. Even after the decades of the research work, this problem has not been

effectively solved. In this problem, Medium Access Control or commonly known as MAC protocols are widely utilized in WSN to solve energy dissipation problems. The scheme allows a proper scheduling of a node to go to sleep and wake mode for saving energy where they are in idle state. The MAC protocols are primarily responsible for furnishing better communication channel among the sensors by sharing the access medium very precisely and fairly [3]. The protocol stack of WSN, MAC layers assist in rectifying the errors that have evolved right from the physical layer with other responsibilities too e.g. addressing, framing, error controls, solving contradiction of channel among the multiple sensors, node mobility, etc. [4]. There are MAC protocols dedicated for both single and multiple channel network structures of WSN [5]. Such schemes are essentially meant for streamlining the data transmission to a better communication system for ensuring energy efficiency and scalability [6]. It is known that inclusion of duty cycle in MAC modeling incorporates better energy modeling [7], [8]. The researches considering duty cycle was not much successful as the majority of them suffer from significant latency. This is because no data will be transmitted or received unless and until the sensors come out their sleep mode. Hence, the duty cycle is a prominent factor that directly affects the network performance of WSN. Hence, in past one decade, multiple MAC protocols have been evolved to resolve energy issue, scalability, delay, and latency. Research work on MAC protocols aged more than a decade and now it is essential to study for the objective of optimizing the system model for issues like energy optimization, time-synchronization, etc.

Therefore, this paper elaborates an extensive review of the existing research contributions, its methodologies adopted, advantages and limitations in the domain of WSN. It reviews both frequently used research work as well as recent research implementation using MAC protocols. Section II discusses the fundamental characteristics of MAC protocols in WSN followed by Taxonomies of Research Work in Section III. Section IV discusses the recently conducted research work in this direction followed by brief highlights of explored research gap in Section V. Future scope is given in Section VI followed by Conclusion in Section VII.

## II. MAC PROTOCOLS IN WSN

In WSN, the MAC protocols are responsible for defining the methods of accessing the resources present in a channel. This protocol is mainly meant for sharing resources over the same channels by a large number of sensors not impacting the network performance in the negative sense. The different functional requirements of MAC protocols are as follows [9]:

- *Reliability*: The MAC protocols ensure the error-free communication process between two nodes. It uses retransmission and acknowledgment mechanism to do so.
- *Medium Access*: The nodes are controlled to participate in the communication process at any instances of time.
- *Framing*: The MAC protocol has a precise way to represent the format of a “data frame” in achieving the specific encapsulation and appropriate operation of de-encapsulation in the communication process.
- *Error Control*: The MAC protocols in WSN should have a higher degree of error detection capability supported by codes for correcting errors.

It is essential that an efficient MAC protocol must possess the property of minimizing latency and energy consumption with maximized throughput and fairness in as a compulsory parameters in WSN. The effective set of properties that are present in the MAC protocol in WSN are as follows:

- *Collision Avoidance*: MAC protocol utilized to avoid a collision. If the node is to send data, it will check first network allocation vector. If that network allocation vector value is not set to zero, then node determines that medium is busy. Hence, proper allocations of slots precisely assist the node accordingly to its sleep and awoken states.
- *Energy Efficiency*: This is accomplished by making use of MAC protocol. During message transmission, the message is divided into the frame and sent into the network to conserve more transmittance energy.
- *Scalability and Adaptability*: MAC protocol also aims to minimize significant problems of scalability as well as adaptability. This protocol helps in fine tuning at network size, node density variation along with the topology effectiveness.
- *Channel Utilization*: As MAC protocol is all about allocation and scheduling, hence it significantly assists in optimal utilization of channel.
- *Latency*: MAC protocol maintains a superior communication among the nodes by sharing their sleep-wake schedules. This significantly attempts to minimize delay in WSN.
- *Throughput*: This is improved by making use of MAC protocol. This MAC protocol will set the amount of data transferred from sender to receiver in the unit

frame by this it will improve the throughput.

- *Fairness*: MAC protocol attempts to maintain a better balance among all the channel parameters and thereby it encourages better fairness over the WSN even in high traffic condition.

Apart from the above characteristics of MAC protocol in WSN, the significant advantageous features of it are as follows:

- Reduce energy waste caused by idle listening in sleep schedule.
- Uses time synchronization overhead, it may prevent sleep schedule announcement.
- Sleep time is high while the collision probability is less in the MAC protocol.
- In this, low latency can be attained with traffic sources and also energy consumption in the network will be minimized.
- Ensures high throughput under low contention.
- Needs less schedule maintenance.

Similarly, the inherent MAC protocols issues in WSN are as follows:

- The transmission slot value of MAC protocol is needed to be set as high as 7x time than the random access period. This results in transmission and reception duty cycle to a higher value.
- Increase idle listening is draw back present in MAC protocol this is caused by listening to all slots before sending.
- Redundancy in MAC protocol which causes waste in transmission power.
- Occurrences of overhead in Synchronization due to prolonged interval of listening period MAC protocol.

Hence, the MAC protocols are associated with both advantages and disadvantages in performance. The next section discusses the taxonomies of the research work in MAC scheme, where the frequently applied MAC protocols are discussed briefly.

## III. TAXONOMIES OF RESEARCH IN MAC SCHEMES

The MAC schemes in WSN include i.e. 1) scheduled; and 2) content-based MAC protocol. The scheduled protocols work on restricted clock synchronization demands in a WSN. The listen and transmit period is scheduled to resist the consequences of collusion, idle listening and over the hearing. They also make use of Carrier Sensed Multiple Access (CSMA) methods. The Content-based MAC (CBMAC) protocols have flexibility in time synchronization demands and are capable of automatically fine tuning itself as per the new topologies by adding up new sensors. This section will present some of the frequently known MAC protocols very briefly [10]:

A. Sensor MAC (S-MAC)

S-MAC is contention-based and is an enhanced type of conventional IEEE 802.11 standard. S-MAC uses two-time frames where a one-time frame is utilized to listen while another frame is utilized for sleeping. SMAC also uses some commonly known beacons called as SYNC e.g. Acknowledgement (ACK), Request to Send (RTS), clear To Send (CTS). However, there is energy dissipation during the listening of SN.

B. Berkley MAC (B-MAC)

This is contention-based MAC scheme in WSN and is improved ALOHA scheme with preamble sampling. It allows faster switching of the node without any chance to miss any data. B-MAC has better energy conservation feature using long preambles and unsynchronized duty cycles. It also uses a filter mechanism for enhancing the assessment of channel and reliability. The biggest advantage of B-MAC is its capability to reduce the idle listening as well as duty cycle without using any SYNC packet.

C. Timeout MAC (T-MAC)

This helps in solving the uniform sleep listen period of sleep in S-MAC. In this, threshold time is used to end the period at an idle period.

D. Wise MAC (W-MAC)

It allows the preamble to precede every data to alert the receiver. This scheme uses acknowledgments as the schedules of sleep during the data exchange for updating their neighbor node's sleep period. Finally, it adjusts its schedule by referring the same of its neighbor node.

E. Traffic Adaptive MAC Protocol (TRAMA)

The prime aim of this scheme is to save energy using TDMA scheme. This scheme considers that MAC layer can

compute the duration of transmission that data is transmitted to the application layer called as SCHEDULE\_INTERVAL. It then prioritizes the highest hop neighbor as well as evaluates the slot numbers using time as [t+SCHEDULE\_INTERVAL]. Finally, the announcements of the slots are done, and a bitmap is utilized to flag the defined receivers with their respective scheduled packets.

F. Data Gathering MAC (D-MAC)

It is a scheduled-based protocol aims to enhance the energy efficiency and minimize latency. The technique permits division of time into various small slots and uses acknowledgment for every time slots for receiving and transmitting a single data packet. One significant advantage of D-MAC is it has a smaller delay.

G. Convergent MAC (C-MAC)

This protocol is mainly meant for energy efficiency as well as latency by using excessive RTS message in the forwarding transmission. C-MAC also checks its channel twice for assessing the channel quality. The significant beneficial point of C-MAC is its capability to resist overhead due to synchronization as well as its ability to work at minimal duty cycles. C-MAC also uses any cast for performing packet forwarding mechanism to explore the forwarding node and perform converging with unsynchronized duty cycle.

Table 1 summarizes the technique used, significances and drawbacks of the frequently used MAC protocols in WSN. It can be observed that MAC protocols are mainly meant for energy efficiency enhancement of each SNs be it cluster head or member node in a sensor network. In spite of the presence of various MAC protocols, there are less number of MAC protocols which has received standardization tag among the research community. The prime reason behind this is that most of MAC protocols are highly dependent on applications posing a great challenge in evolving up as generalized standards.

TABLE. I. SUMMARY OF FREQUENTLY USED MAC PROTOCOLS

Protocol	Technique used	Advantage	Limitation
S-MAC	Virtual cluster, Uniform duty cycle	Low energy	Latency during broadcast
B-MAC	Channel assessment	Minimal overhead	Degraded performance at high traffic
T-MAC	Overhearing, duty cycle (adaptive)	Active time (adaptive)	Pre-pones sleep schedule
Wise-MAC	Lower schedule of preamble sampling	Less overhead	Not energy efficient
TRAMA	Schedule interval	Energy efficient	Random access period
D-MAC	Excessive RTS message	Minimal latency	Higher rate of aggregation

IV. EXISTING TECHNIQUES

This section intended to survey of existing recent techniques in MAC protocols. The work of Abdel et al. [11] described the Hybrid MAC Protocol (HMACP) for the real-time applications. The model contains both Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) that offers a smooth real-time communication among the randomly deployed WSN. This schedules the nodes communication which avoids collisions, interfaces optimization and gives less delay. In this, the simulation was performed for SNs and random sinks. The outcomes give the reliable, scalable and end to end delay.

The model proposed in Fig. 1 consists of four processing units such as Initialization, Synchronization & Discovery, slot or frequency attribution, and medium access. For example, Each node after the deployment turns on their radio transmitter and will wait for a message. After receiving the message by every node, the discovery step will be actuated by saving the messages of the source address in the neighbor list and also broadcast the message in a proper ID. Once the direct information is collected, every node will forward their respective neighbor's list to the base station (BS). The BS will construct a graph and is connected to graph G (V, E) representing a network, where V – set of nodes & E – Edges .

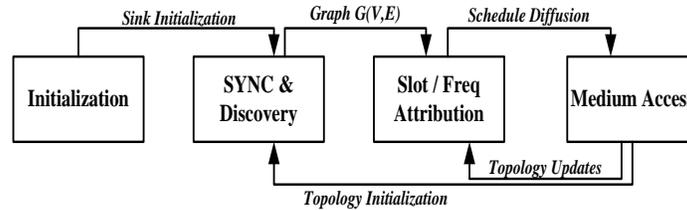


Fig. 1. Functioning process of Abdeli et al. [11] work.

The slot/frequency attribution (SFA) unit contains schedule plan construction, basing over graph (G) by the attribution of every node  $m \subseteq V$ , where m is some nodes and V is set of nodes. The BS will perform the diffusion of schedule to nodes. After the above hierarchical process nodes will send the data packets.

The sampling rate (SR) effect on the collision probability is described in Dong et al. [12] especially for hybrid MAC protocols of WSN. From the investigation of the author, it is known that the collision is the main cause of energy waste in MAC protocols under the saturate condition considering the sampling rate greater than the transmission rate. Later the investigation study towards the effect of collision for WSN was performed by considering sleeping MAC protocol. Karahan et al. [13] developed an energy efficient hybrid MAC (EEHMAC) protocol for a large-scale WSN. In this work, the techniques like TDMA and CSMA are considered. The reason for the use of TDMA is that it reduces unnecessary collisions in channels while the CSMA responses rapidly for events. With this model author achieved high throughput and energy optimization.

A study was based on traffic adaptive MAC protocol with multichannel and allocation of dynamic slots in Zhuo et al. [14] for the WSN. The contribution of this work is that it presented a hybrid iQueue-MAC (TDMA and CSMA) mechanism which adopts the variable and huge traffic. The study also discusses how the mechanism will operate in two different channel conditions like multimode and single mode. The mechanism under multiple channel condition duplicates the throughput and performs the efficient operation as compared to the Single channel operation. Vincent et al. [15] have presented the concept of versatile MAC protocol that defines mobility handling issue of WSN. The model of VMAC combines the energy saving schedule based, and CBMAC protocol for short range of communication. In the work of Zhang et al. [16] application specific and semantic specific Hybrid MAC Protocol is discussed with Quality of Service (QoS) for WSN. Author has considered SNs as sinks of different applicability. The energy optimization and delay analysis results are analyzed which indicates better performance than other MAC protocols.

The combined work carried by Haiyang et al. [17] states mechanism with MAC protocol for WSN. The study analysis of the previous researches says the WSN have the issue of low batter power that needs to sort out. The study gives the some of the important MAC protocols in solving the WSN power

issues. An integrated MAC protocol is given by Andrade et al. [18]. In this, WSN integration and radio over fiber (RoF) is performed by using MAC protocol. A very high throughput hybrid Mac protocol is presented in Jian et al. [19] for the millimeter wave WSN communication. The method significantly provides the energy efficiency in SN. This protocol consists of two networking elements like Relay master mote (RM), and Remote sensor mote (SM) and its state transitions are shown in Fig. 2.

In RM state transition once the state is initialized the RM will enter to Synchronization (Sync) that transfers a poll to build the connection among RM and SM. The SM will send an acknowledgment or ACK to the RM, and it will change its Idle state to Sleep state, which means no further synchronization is required. The scheduled data moves to the RM the state will move to Rx. Later the RM state will move to Tx State and assigns next data flow for SM to Idle. Then RM will enter the sleep state or else wait for next transition. Similarly, in next SM state transition (Right) will take place. The adaptive scheme of modulation for MAC with Fade state feedback quantization is presented by Kundu and Rajan [20]. In this one, a user will rotate the constellation without changing the transmit power that existing channel state, to meet the predefined minimum Euclidean distance at the destination. The performance and cost factors are improved than traditional methods.

The combined work for the Hybrid-MAC, Hybrid Sensor-MAC for Adhoc on multipath distance vector (AdMDC) is discussed in Kalaivaani and Rajeswari [21]. The study analysis presents that the AdMDC and H-MAC yield better performance results and energy optimization. An H-MAC protocol is presented in Hsieh et al. [22] for WSN. In this mechanism uses a cross-layer method which switches CSMA and TDMA over the network. The performance of this H-MAC is evaluated with different network density and achieved the improved energy efficiency and packet latency for WSN.

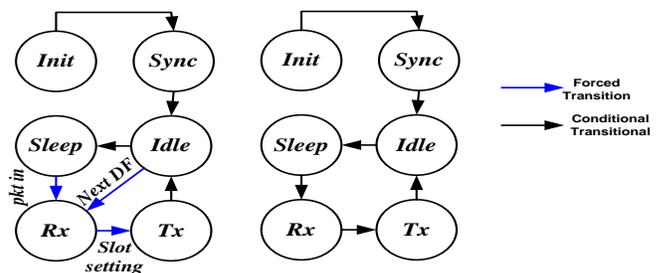


Fig. 2. State transition of RM (left) and SM (right).

The aim for energy efficiency in WSN with MaC protocol is presented in Othman et al. [23] and QoS improvement. The outcomes from simulation and analytical are compared with Q-MAC protocol, and it was found that method proposed by the author has got more reliability and also low power consumption. The research performed by Wijetunge et al. [24] described the IEEE 802.15.4 based H-MAC protocol for hybrid WSNs monitoring. The protocol is defined to get the energy efficiency, low delay, and reliability in the monitoring of WSN. The objectives functions are optimized to meet the performance goal.

Author Shen et al. [25] presented MAC protocol to achieve energy efficiency. In this, the evaluation of various performance metrics is simulated with event simulator ns-2 mainly for the energy and benchmarked. The congestion control hybrid MAC mechanism for WSN is discussed in Priya and Manohar [26]. The protocol is the combination of both CSMA and TDMA, and that results in the advancement in the energy efficiency, packet delivery ratio, and delay performance. The conceptualize diagram of [26] is represented in Fig. 3 where the congestion takes place when data is transmitted. To minimize the congestion author have calculated the queue state and analyzed it.

$$QueueState = \frac{No.ofPkt \sin Queue.2^3}{max QueueSize}$$

Later the drop tail queuing systems are adopted in WSN. The packets of the congestion will drop at queue end. Hence the node with long queue will get affected than the small queue. The evaluation of channel allocation for H.mac (Multichannel) is described in Diab et al. [27]. In this simulation is performed for channel allocation efficiency of HMC-MAC and compared it with other methods. Later the interference rate is considered as the main metric, and results give how HMC-MAC will reduce the interference rate compared with other methods. Different consequences are considered with the same channel in 3, 2 and 1-hop neighborhood, to analyze the inference rate minimization concept.

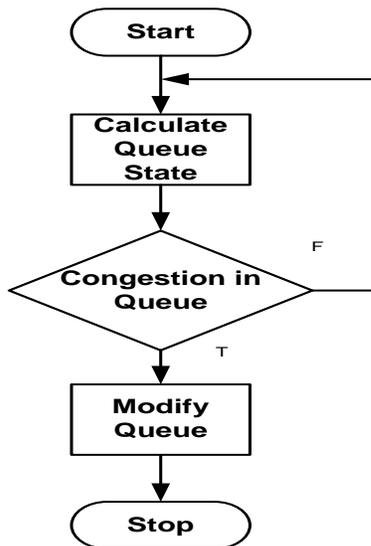


Fig. 3. Process diagram of [26].

```
1: /* Sender s sending data to Receiver r */
2: if A probing packet received from r then
3:   Send data to r.
4: else if Number of timeout > τ then
5:   Switch to T-mode and Send data to r.
6: end if
7: Switch back to R-mode.
8: /* r receiving data from s */
9: if Preamble received from s then
10:  Wait for data from s after the preamble.
11: end if
12: Go to sleep mode.
```

Fig. 4. Algorithm of [28] work.

The low powered duty cycle WSN is subjected to Asym-MAC protocol with Asymmetric links in Won et al. [28]. The experimental results of the method yield high packet reception rate, low packet transmission delay than other protocol. The algorithm adapted for the method is given in Fig. 4. Where T-mode is transmission mode, R-mode is receiving mode, τ is system throughput.

The characterization parameter in WSN is presented in Kadu and Deshpande [29]. The results exhibited from the grid topology, the average throughput achieved higher as compared to the other topologies like a chain and two random.

Author Kamal et al. [30] presented the secure/time synchronized and efficiency enhancement using the improved IHMAC. The outcomes are validated for enhanced efficiency. The performance and modeling analysis of MAP protocols with S-MAC and X-MAC are presented in Yang et al. [31]. In this Markov model together with the performance analysis gives estimated throughput, energy consumption, and delay for both synchronized and asynchronous duty-cycled MAC protocols. Akker et al. [32] described the interference effect among the heterogeneous WSN MAC protocol. The performance of TDMA, T-MAC other protocols is evaluated successfully. A dynamic de-centralized Hybrid MAC-Protocol is presented by Shah et al. [33]. In this, cognitive radios were used to solve spectrum scarcity problem for WSN applications. The outcome suggested that the mechanism is efficient, flexible and reliable for the real-time applications. Feng et al. [34] presented an approach to control overhead using S-MAC protocol.

## V. RESEARCH GAP

All the explained MAC protocols exhibit both significance and limitation. Here, the problems which were not considered in many of past research works are discussed. Such unsolved problem poses a significant impediment towards the path of research work and hence we term it as a research gap.

- *Less Applicability to the Real world:* From the previous section, it has been seen that there are various MAC protocols with potential architectural design; however, it is still questionable that how much they are applicable in a real-world scenario. Any applications over real-world will always demand the compliance towards time

and mission criticality. It is pretty significant that the data packets that are forwarded from the transmitting sensor should reach the receiver sensor before the time limit to ensure the time criticality as well the authentication of the true alarm notification. The various system calls for re-transmission which alternately schedules the data forwarding process is initiated to resist the data packet influence need to be dropped belongs to the slipped schedule. Thus, the data packet is required to be prioritized for transmission sequences. One common observation in existing technique is that it successfully maintain the timeliness, But, recently developed MAC schemes doesn't include a sequence of message transmission by the time limit. On the contrary, the algorithms are developed for minimizing the delay of data transfer. One of the unanswered questions found in all the existing techniques is that none of the techniques could explain the instantaneous sensor's response against the sudden event. For the MAC protocol to be working perfectly on real-time, it is required that they should have faster response time with lower delay (or latency), which is still unanswered.

- *Doesn't solve energy problems completely:* It is already known that the sensor node has a limited energy and battery lifetime. The limited lifetime also results in degradation of the network performance and causing intermittent (or un-stabilized links). As MAC protocols are mainly meant of energy conservation within the nodes, it is still questionable that how much of energy efficiency it can offer.
- *More disadvantages compared to advantages:* There are various MAC schemes that are using TDMA protocol in a WSN. The biggest challenges with such protocol usage are that its applicability is highly confined to specific traffic size. Till date, there is few research work which has proved optimal scalability achievement using TDMA based MAC protocols. Another flawed problem is the assumption that it makes in formulating MAC protocols. There are assumptions of the base station being covering entire sensors and compression to be existing only in higher layers. Such assumptions are completely unpractical. There is also a problem with event-driven schemes, which is large communication overhead as they mainly work on clustering principle. This degrades the performance of time synchronization that also negatively impacts on the network lifetime. More the surveillance area more is energy dissipation. A similar problem also exists for power efficient TDMA scheme. As they mainly depend on clustering, so it results in overhead. The problem becomes worst in large-scale dynamic networks. The MAC techniques that use distributed schemes to conserve energy also suffer from issues e.g. delay which is generated from the node with lower residual energy. Not only do these existing schemes also have higher dependencies towards hardware capabilities to use non-interfering communication channels. Such schemes drastically minimize the data bit rate because

of the split communication channel. Although TDMA schemes offer a better lifetime, its acceptance level is quite poor in WSN on scalability and its capability to function properly at the dynamic network.

- *Fewer studies on Hybrid MAC schemes:* A closer look at existing MAC schemes shows that almost every algorithm has advantages as well as disadvantages. Hence there are fewer attempts towards investigating about hybrid schemes. As these schemes integrate both CSMA and TDMA, such schemes can offer a greater deal of flexibility and energy efficiency. The studies [35]-[38], gives that there is a less supportability of the dynamic topology or mobility. The existing work on hybrid MAC schemes are also not found to evolve with a solution towards its inability to respond quickly to sudden event occurrence (in real-time).

Hence, a significant research gap is being explored in the area of MAC protocols in a WSN. The problems that are yet to be resolved are energy issues, scalability issues, delay and latency issues, overhead issues, and supports real time applications that will require being established with faster response time. This makes an obvious requirement to study and investigation the insights in related to these issues.

## VI. FUTURE WORK

The future work will be towards the direction of overcoming the research gap explored and briefed in the prior section. The tentative research works that can be carried out in future are briefed as follows:

- *Effective integration of TDMA and CSMA scheme:* The prime motive of this integration is to mainly achieve enhanced sustainable capacity towards increasing traffic load in WSN. To overcome the limitation of scalability and latency in the network, this integration scheme could focus on overcoming this issue. For this purpose, a specific control message could be designed with the inclusion of field for data priority. The scheme should also classify the time of communication into a uniform frames. It should be followed by an effective discovery of neighbor nodes to further assists in transmission, clustering, as well as synchronization. Better modeling could be achieved considering the state-transition of SYNC packets (i.e. RTS, CTS), etc. A robust modeling of search technique towards communication link could be designed to stabilize the energy dissipation and. The solution towards faster response time could be solved by energy modeling if the power allocation towards the sensor could be making more dynamic for precisely varying the transmittance power. A suitable balance between the crowded traffic scenario and packets to be delivered can be developed by focusing more on the packet prioritization. Such design could be suitably assessed using QoS as a performance factor.
- *Architecture to support dynamicity:* Architecture could be designed for ensuring minimal scalability, higher supportability of throughput, cutting short idle listening time. A unique discovery process using single hop could be designed which changes with the topology. A

simply carried sensing model utilized for initialization of the communication process, and we encourage using only short preamble size to address overhearing problem. To solve maximum wake up time, we plan usage of involuntary buffering mechanism that can also contribute to enhancing the network lifetime. The next contribution could be to develop a synchronization algorithm to resist complexities of communication on different types of clusters. Even an empirical modeling can be carried out for energy modeling over presented MAC scheme using multiple energy distributions of the sensor node. Such scheme would minimize the idle listening time for the entire sensor, which will let the sensor go in hibernate mode if there is no data to be delivered.

- *Decision-making model to meet peak traffic condition:* Existing studies have less computational modeling on MAC protocols towards catering up the peak traffic demands in WSN. It is essential to identify the potential feature of forwarding rate by MAC protocols. Applying the approach using hybrid MAC protocols, this feature could be further enhanced. A clustering approach will be considered where all the nodes disperse in the cluster will possess mini-slots, which is accomplished by dividing the main time slot. The aggregator node performs carrier sense to join the network till base station or its neighbor node. In case the communication channel is failed to be received the aggregator node step over to the hibernation modes/stage and is allowed out of present state using an arbitrary period. An idea is not to miss any data as well as conserve energy also. Another advantage anticipated from this approach is effective synchronization policy as well as capability to meet the demands of meeting the peak traffic condition. For better scalability, we plan to develop discrete MAC protocols using a hybrid approach for both communications inside the cluster (member node-aggregator node) and outside the cluster (aggregator node-base station).

## VII. CONCLUSION

This paper has discussed the usage of the MAC protocols and presented the most significant factors for network lifetime enhancement of the WSN. It also describes various forms of MAC schemes, along with the finding that most of the conventional MAC schemes are quite theoretically and fails to cater up to the need of optimality of the WSN performance. The paper, explains various schemes used for MAC protocol performance enhancement as well the research gap is extracted to set the future research direction. There is very less novelty in the existing approaches that causes very less improvement in network lifetime. Hence, our future work will look for continuing the investigation towards evolving up with new strategies and mathematical modeling to address such research gap.

## ACKNOWLEDGMENT

The author gratefully acknowledges the support and encouragement of management and Dr. R Balakrishna, Principal, RajaRajeswari College of Engineering, Bengaluru.

## REFERENCES

- [1] K.J. Kim, N. Wattanapongsakorn, N. Joukov, "Mobile and Wireless Technologies," Springer Technology & Engineering, pp. 252, 2016
- [2] S. Khan, A-S. K. Pathan, N. A. Alrajeh, "Wireless Sensor Networks: Current Status and Future Trends," CRC Press, pp. 546, 2016
- [3] K. J. Kim, N. Wattanapongsakorn, N. Joukov, "Mobile and Wireless Technologies," Springer, Technology & Engineering, pp. 252, 2016
- [4] M.S. Obaidat, S. Misra, "Principles of Wireless Sensor Networks," Cambridge University Press, pp. 433, 2014
- [5] I. M. M. E. Emary, S. Ramakrishnan, "Wireless Sensor Networks: From Theory to Applications," CRC Press, pp. 799, 2013
- [6] I. F. Akyildiz, M.C. Vuran, "Wireless Sensor Networks," John Wiley & Sons, pp. 520, 2010
- [7] Christophe J. Merlin, Wendi B. Heinzelman, "Duty Cycle Control for Low-Power-Listening MAC Protocols," IEEE Transactions on Mobile Computing, 2010
- [8] F. Alfayez, M. Hammoudeh, A. Abuarqoub, "A Survey on MAC Protocols for Duty-cycled Wireless Sensor Networks," International Conference on Advanced Wireless Information and Communication Technologies (AWICT 2015), Vol. 73, pp. 482-489, 2015
- [9] Joseph Kabara, M. Calle, "MAC Protocols Used by Wireless Sensor Networks and a General Method of Performance Evaluation," Computer Science, Information Systems, 2016
- [10] S. I. A. Shah, M. Ilyas, H.T. Mouftah, "Pervasive Communications Handbook," CRC Press Technology & Engineering, pp. 500, 2016
- [11] D. Abdeli, S. Zelit and S. Moussaoui, "RTH-MAC: A real time hybrid MAC protocol for WSN," Programming and Systems (ISPS), 2013 11th International Symposium on, Algiers, pp. 153-162, 2013
- [12] Q. Dong, W. Dargie and A. Schill, "Effects of sampling rate on collision probability in hybrid MAC protocols in WSN," 2010 IEEE Globecom Workshops, Miami, FL, pp. 213-218, 2010
- [13] A. Karahan, İ Ertürk, S. Atmaca and S. Çakıcı, "Energy efficient hybrid MAC protocol for large scale wireless sensor networks," 23rd Signal Processing and Communications Applications Conference (SIU), Malatya, pp. 1549-1552, 2015
- [14] S. Zhuo, Z. Wang, Y-Q. Song, Z. Wang, and L. Almeida, "A Traffic Adaptive Multi-channel MAC Protocol with Dynamic Slot Allocation for WSNs," IEEE Transactions on Mobile Computing, pp. 1536-1233, 2015
- [15] V. Ngo, A. Anpalagan and I. Woungang, "Versatile medium access control (VMAC) protocol for mobile sensor networks," 7th International Wireless Communications and Mobile Computing Conference, Istanbul, pp. 836-841, 2011
- [16] F.Zhang, F. Meng and Q.Xiong, "A novel application-semantics based hybrid QOS MAC in wireless sensor networks," Wireless Communications, Networking and Mobile Computing (WiCOM 2014), 10th International Conference on, Beijing, pp. 456-461, 2014
- [17] H. Zhang, "Classic Efficient-Energy MAC Protocols for Wireless Sensor Networks," 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), Chengdu, pp. 1-4, 2010
- [18] T. P. C. de Andrade, L. B. Oliveira, N. L. S. da Fonseca and O. C. Branquinho, "HMARS: A MAC protocol for integration of Radio-over-Fiber and Wireless Sensor Networks," IEEE Third Latin-American Conference on Communications, Belem do Para, pp. 1-6, 2011
- [19] W. Jian , C. I. Estevez, A. Chowdhury, Z. Jia, G-K. Chang, "A Hybrid MAC Protocol Design for Energy-Efficient Very-High-Throughput Millimeter Wave Wireless Sensor Communication Networks," IEEE Third Latin-American Conference on Communications, 2010
- [20] S. Kundu and B. S. Rajan, "An adaptive modulation scheme for two-user fading MAC with quantized fade state feedback," IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC), Sydney, NSW, pp. 512-518, 2012

- [21] P. T. Kalaivaani and A. Rajeswari, "An analysis of H-MAC, HSMAC and H-MAC based AOMDV for wireless sensor networks to achieve energy efficiency using spatial correlation concept," *Electronics and Communication Systems (ICECS)*, 2nd International Conference on, Coimbatore, pp. 796-801, 2015
- [22] T. H. Hsieh, K. Y. Lin and P. C. Wang, "A hybrid MAC protocol for wireless sensor networks," *Networking, Sensing and Control (ICNSC)*, 2015 IEEE 12th International Conference on, Taipei, pp. 93-98, 2015
- [23] J. Ben-Othman, L. Mokdad and B. Yahya, "An Energy Efficient Priority-Based QoS MAC Protocol for Wireless Sensor Networks," 2011 IEEE International Conference on Communications (ICC), Kyoto, pp. 1-6, 2011
- [24] S. Wijetunge, U. Gunawardana and R. Liyanapathirana, "IEEE 802.15.4 based hybrid MAC protocol for hybrid monitoring WSNs," *Local Computer Networks (LCN)*, IEEE 38th Conference on, Sydney, NSW, pp. 707-710, 2013
- [25] J. Shen, F. Yi, S. Moh and I. Chung, "Energy Efficiency of MAC Protocols in Wireless Sensor Networks," 2011 International Conference on Information Science and Applications, Jeju Island, pp. 1-10, 2011
- [26] B. Priya, S. Solai Manohar, "CH-Mac: Congestion Control Hybrid MAC for Wireless Sensor Network," 4th ICCCNT, 2013
- [27] Rana Diab, G'érard Chalhoub and Michel Misson, "Hybrid Multi-Channel MAC Protocol for Wireless Sensor Networks: Interference Rate Evaluation," *IEEE International Conference on Communications*, 2013
- [28] M. Won, T. Park, and S. H. Son, "Asym-MAC: A MAC Protocol for Low-Power Duty-Cycled Wireless Sensor Networks with Asymmetric Links," *IEEE Communications Letters*, Vol. 18, No. 5, 2014
- [29] S.D. Kadu, V.S. Deshpande, "Characterization of Throughput in Wireless Sensor Network for MAC and Routing Protocol," *International Conference on Cloud & Ubiquitous Computing & Emerging Technologies*, 2013
- [30] P. Kamal, S. Singh, Z. M. Livinsa, "Secure Time Synchronization and Efficiency Enhancement using Improved IHMAC," *IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems ICHIECS*, 2015
- [31] O. Yang, and W. B. Heinzelman, "Modeling and Performance Analysis for Duty-Cycled MAC Protocols with Applications to S-MAC and X-MAC," *IEEE Transactions on Mobile Computing*, Vol. 11, No. 6, 2012
- [32] D. V. D. Akker, B. Braem, C. Blondia, "On the Effects of Interference between heterogeneous Sensor Network MAC Protocols," *Eighth IEEE International Conference on Mobile Ad-Hoc and Sensor Systems*, 2011
- [33] M. A. Shah, Ghazanfar A. Safdar, C. Maple, "DDH-MAC: A Novel Dynamic De-Centralized Hybrid MAC Protocol for Cognitive Radio Networks," 2011
- [34] H. Feng, L. Ma, S. Leng, "A Low Overhead Wireless Sensor Networks MAC protocol," *IEEE International Conference*, 2010
- [35] B. Shrestha, K. W. Choi and E. Hossain, "A Dynamic Time Slot Allocation Scheme for Hybrid CSMA/TDMA MAC Protocol," in *IEEE Wireless Communications Letters*, Vol. 2, No. 5, pp. 535-538, 2013.
- [36] V. Nguyen, T. Z. Oo, P. Chuan and C. S. Hong, "An Efficient Time Slot Acquisition on the Hybrid TDMA/CSMA Multichannel MAC in VANETs," in *IEEE Communications Letters*, Vol. 20, No. 5, pp. 970-973, 2016.
- [37] Y. Xu, P. Xu, W. Zeng and Q. Tian, "Hybrid MAC based resource management scheme for kiosk service in 802.15.3c WPAN," 7th International Wireless Communications and Mobile Computing Conference, Istanbul, pp. 516-521, 2011
- [38] R. Zhang, L. Cai and J. Pan, "Performance Study of Hybrid MAC Using Soft Reservation for Wireless Networks," *IEEE International Conference on Communications (ICC)*, Kyoto, pp. 1-5, 2011

# Reduced-Latency and Area-Efficient Architecture for FPGA-Based Stochastic LDPC Decoders

Ghania Zerari, Abderrezak Guessoum  
“LATSI” Laboratory, Department of Electronics,  
University of Blida,  
Blida, Algeria

**Abstract**—This paper introduces a new field programmable gate array (FPGA) based stochastic low-density parity-check (LDPC) decoding process, to implement fully parallel LDPC-decoders. The proposed technique is designed to optimize the FPGA logic utilisation and to decrease the decoding latency. In order to reduce the complexity, the variable node (VN) output saturated-counter is removed and each VN internal memory is mapped only in one slice distributed RAM. Furthermore, an efficient VN initialization, using the channel input probability, is performed to improve the decoder convergence, without requiring additional resources. The Xilinx FPGA implementation shows that the proposed decoding approach reaches high performance along with reduction of logic utilisation, even for short codes. As a result, for a (200, 100) regular codes, a 57% reduction of the average decoding cycles is attained with an important bit error rate improvement, at  $E_b/N_0 = 5.5\text{dB}$ . Additionally, a significant hardware reduction is achieved.

**Keywords**—Stochastic decoding; low-density parity-check (LDPC) decoder; field programmable gate array (FPGA)

## I. INTRODUCTION

The need for increasing the throughput of modern communication systems capacity, for optical and wireless networks, requires high performance error correcting code. In 1962, Gallager presented the first version of low-density parity-check (LDPC) codes [1], offering an excellent process for repairing transmission errors, added by the channel effects. The close Shannon capacity decoding performances, of the LDPC codes [2], justify their exploitation by various digital communication standards. WiMAX (IEEE 802.16e), DVB-S2, WiFi (IEEE 802.11) and 10GBASE-T (IEEE 802.3an) standards attest this great performance.

A variety of LDPC decoding implementations have been explored to accomplish high throughput results [3]-[5]. It has been evidently shown that the higher throughput is achieved by the fully parallel decoding solutions; nevertheless they enlarged the hardware complexity. To overcome this drawback, several reduced-complexity and stochastic LDPC decoding algorithms are developed [6]-[8]. The current stochastic decoding algorithm confirmed their adaptability for fully parallel decoding approach [9]-[18]. Moreover and for additional silicon area reduction, diverse LDPC stochastic based decoding architectures and strategies are proposed.

However, an area-efficient architecture for ASIC-Based stochastic LDPC decoder can't systematically produces an efficient FPGA logic utilisation. It is straightforward that the

ASIC implementation of six bits counter requires less silicon area compared to 32 bits memory. Nevertheless an inverse result is obtained with the FPGA implementation. A Xilinx FPGA 32 bits memory implementation can be routed using only one LUT, in contrast with the counter logic utilisation. This paper introduces a new and powerful field programmable gate array (FPGA) based stochastic Low-Density Parity-Check (LDPC) decoding process, to implement fully parallel LDPC-decoders. The proposed technique is designed to optimize the FPGA logic utilization and to decrease the decoding latency in addition to improve the convergence, even for short codes. To validate the advantage of the proposed approach, an FPGA is implemented using Xilinx Virtex-6 VLX240T. The paper is organized as follows. In Section II, an overview of the LDPC stochastic decoding is provided. In Section III, the architecture of the new proposed stochastic LDPC decoding is introduced. Results of FPGA implementation and performance are presented in Section IV, and finally a conclusion is given in Section V.

## II. LDPC STOCHASTIC DECODERS

The design of an LDPC decoder is based on the  $M \times N$  parity check matrix  $H$ .  $N$  defines the number of variable nodes (VNs) while  $M$  defines the number of CNs. To encode  $k$  information bits, an  $(N, K)$  LDPC code uses  $N$  encoded bits, where  $N > K$ . LDPC decoder can be represented by a factor graph which uses  $N$  VNs and  $(N-K)$  CNs. A  $dv$  degree VN has  $(dv+1)$  ports, one of which gets the channel probability and the other  $dv$  are connected to different CNs, by bidirectional ports. In the same way, the  $dc$  degree CN has  $dc$  bidirectional ports, which are connected to different VNs, and one parity-check output port. Conventional LDPC fully parallel decoder uses fixed-point operands to represent the probabilities, exchanged between the VNs and the CNs of the factor graph. Stochastic LDPC decoders function by a bit-serial iterative process. In this architecture, the received probabilities  $P_{ch}$  from the channel are converted to Bernoulli sequences as random bits sequences. Different encoded stochastic sequences can be generated for the same probability. In a  $\{ai\}$  Bernoulli sequence of  $m$  bits, in which  $a_i \in \{0, 1\}$ , the estimated probability value is computed as:

$$P_{ch}(m) = \frac{\sum_{i=1}^m a_i}{m} \quad (1)$$

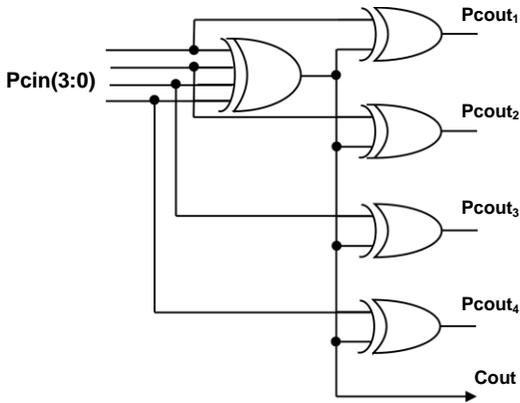


Fig. 1. Structure of a degree-4 parity-check node.

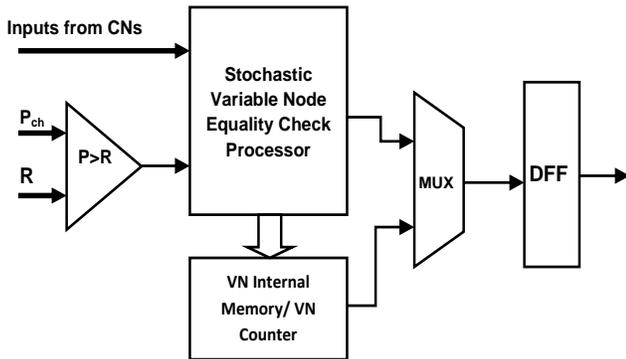


Fig. 2. Structure of recent stochastic variable node.

Let  $Cout$ ,  $Pcout_i$  and  $Pcin_i$  be the parity-check output, the CN outputs and the CN probabilities inputs respectively, where  $Pcin_i = Pr(cin_i=1)$  is the probability of each CN inputs, in which  $i \in \{1, 2, \dots, dc\}$  and  $dc$  is CN degree. The output probability  $Pcout_i$  can be computed as:

$$\left. \begin{aligned} Pcout_1 &= Pcin_2 \oplus \dots \oplus Pcin_{dc} \\ Pcout_i &= Pcin_1 \oplus \dots \oplus Pcin_{i-1} \oplus Pcin_{i+1} \oplus \dots \oplus Pcin_{dc} \\ Pcout_{dc} &= Pcin_1 \oplus \dots \oplus Pcin_{dc-1} \end{aligned} \right\} (2)$$

The parity-check output  $Cout$ , of  $dc$  degree CN, can be computed according to (3). Fig. 1 illustrates the structure of a  $dc=4$  CN used in the conventional stochastic decoder.

$$Cout = Pcin_1 \oplus \dots \oplus Pcin_{dc} \quad (3)$$

Let  $Pvin_1$  and  $Pvin_2$  be the probability of two input bits in  $dv=2$  VN. The variable node output probability  $Pvout$  can be computed as:

$$Pvout = \frac{Pvin_1 \cdot Pvin_2}{Pvin_1 \cdot Pvin_2 + (1 - Pvin_1) \cdot (1 - Pvin_2)} \quad (4)$$

If the VN inputs are same, this state is named the agreement state, one of the input bits will be transmitted to the output. When the inputs are not identical, the variable node requires an advanced method to generate the output bit. This state is named the hold state or disagreement state. One of the

advanced stochastic method bit generation (ASMBG) can be used.

$$Pvout(N) = \begin{cases} \text{Bit using (4)} & \text{if } Pa = Pb \\ \text{Bit using ASMBG} & \text{otherwise} \end{cases} \quad (5)$$

Fig. 2 shows the recent stochastic variable node principal structure. The stochastic LDPC decoding algorithm can be summarised as follows:

---

**Algorithm 1** Stochastic LDPC decoding

---

**Initialization**

1. Load the LLRs corresponding probabilities  $P_{ch}$  for each variable node (one DC) and transform  $P_{ch}$  to Bernoulli sequence  $a_i$  (each DC).

2. Initialize the variable nodes internal memories (16 to 32 DCs for 32 bits memory [10]) or the internal saturated counter (one DC [16]).

**Iterations**

3. Variable to check node: At each decoding cycle, the variable node computes there inputs bits using (5) and sends there outputs bits to the corresponding check nodes.

4. Check to variable node: At each decoding cycle, the check node computes there inputs bits using (2) and sends there outputs results to the corresponding variable nodes. Simultaneously, the check nodes send their outputs states using (3) to the syndrome checker.

5. If  $xH^T = 0$  or the maximum of DCs is reached, terminate the decoding process. Otherwise go to Step 3.

---

III. PROPOSED LDPC STOCHASTIC STRUCTURE

As mentioned in the introduction section, an efficient ASIC-based architecture algorithm can't systematically provide the best approach for an efficient FPGA implementation. In this section we present the new LDPC stochastic decoding method which aims to improve the decoder performance and to reduce the FPGA resource utilization.

It has been shown that the LDPC stochastic decoder, which there VNs use the latest output bits as code bits, provide similar BER performance to the version with saturating up/down counters as a VN output decision mechanism [17]-[18]. Furthermore, it has been demonstrated that the initialization of the first VN output bit, transmitted to the hard decision unit according to received probability channel, helps to improve the stochastic decoder convergence [15]. The proposed VN exploits the two referenced characteristics, in addition to adopt an internal memory-based approach, similar to DS and EM versions. The converted Bernoulli sequences are used as a variable node input. All output variable nodes are initialized by one bit coded probability channel, during the loading of the channel Log Likelihood Ratio (LLR) corresponding probabilities. Each VN internal memory is mapped in one FPGA LUT RAM.

The output variable node probability will be computed as:

$$P_{vout}(t) = \begin{cases} P_{st} & \text{if } t = 1 \\ Pa \text{ or } Pb & \text{if } t \neq 1 \text{ \& } Pa = Pb \\ \text{Bit from FPGA LUT RAM} & \text{otherwise} \end{cases} \quad (6)$$

$$\text{where } P_{st} = \begin{cases} 1 & \text{if } P_{ch} \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

$P_{vout}(1)$  is the first iteration VN output probability. During the disagreement state ( $t \neq 1$ ), the proposed architecture generate a new bit based on a random bit selection from the VN internal memory (IM). The IM length can be increased up to FPGA LUT RAM size. Based on (6), the hardware implementation of the new improved decoding approach does not require extra hardware complexity, for FPGA devices. Moreover, the projected technique computes the received probability without any additional decoding cycle.

The  $P_{ch}(k-1)$  signal is transmitted to the VN output DFF during the first process cycle. After the first iteration and until the last one, the multiplexer sends the variable node processor output bit to the VN output DFF. In this way and identically to the CSS process, the majority of variable nodes outputs start with a right bit and detour the random stochastic initialization. Fig. 3 represents the main structure of the proposed variable node. Similar to the EM and DS design, the decoding cycle (DC) matches to one of the iteration for the proposed LDPC stochastic decoding.

The proposed CN possess two inputs signals categories and two outputs. The inputs signals are the  $CNin_i$  signals and  $State_i$  signals, in which  $i \in \{1, 2 \dots dc\}$ . These two signals are provided by VNs outputs signals. The first outputs signals are the  $CNout_i$  signals, which are sent to VN inputs. The second output is the parity-check output state  $CNstate$ .

The new CN uses a computing process similar to the DS approach. All  $CNstate$  outputs are connected to the syndrome checker unit.

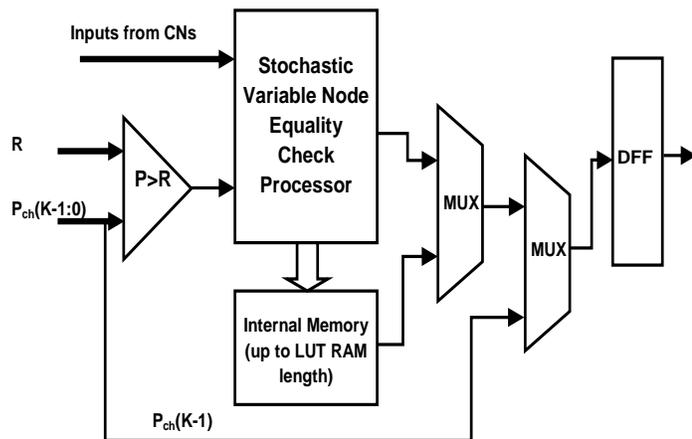


Fig. 3. The structure of the proposed stochastic variable node.

The parity-check output state  $CNstate$ , of  $dc$  degree CN, is computed in the same way of (3) and can be written as follows:

$$CNstate = \sum_{i=1}^{dc} \oplus CNin_i \quad (8)$$

Where,  $\sum \oplus$  is the bitwise XOR operation and  $dc$  is the parity-check node degree.

Each CN outputs signals  $CNout_i$  uses the  $CNin_i$  signals and  $State_i$  signals, to produce the  $CNout_i$  signals according to (9). The  $CNstate$  result given by (8) can be exploited.

$$CNout_i = \left( (CNstate) \wedge \left( \bigvee_{i=1}^{dc} state_i \right) \right) \oplus CNin_i \quad (9)$$

Where,  $\bigvee$  is the bitwise NOR operation and  $\wedge$  is the bitwise AND operation.

The new stochastic LDPC decoding algorithm can be summarised as follows:

---

**Algorithm 2** The proposed Stochastic LDPC decoding

---

**Initialization**

1. Load the LLRs corresponding probabilities  $P_{ch}$  simultaneously with initializing the variable node output (one DC) and transform  $P_{ch}$  to Bernouli sequence  $a_i$  (each DC).

**Iterations**

2. and 3. Similar to Step 3 and Step 4 of Algorithm 1.

4. If  $xH^T = 0$  or the maximum of DCs is reached, terminate the decoding process. Otherwise go to Step 2.

---

IV. IMPLEMENTATION RESULTS AND PERFORMANCE

It has been demonstrated that the enlargement of the VN internal memories size increases the LDPC stochastic decoding converges [10]-[11]. However and mainly, adding additional memory capacity implicates an extra hardware complexity and resources. The FPGA organization and implementation need special considerations. In addition to slices resources, memory can be mapped using Block RAMs or using Distributed RAMs (LUT RAM).

The main target of the proposed structure is to improve the FPGA-based LDPC decoding performance, without supplementary FPGA resources. To confirm the improvement of the new design, a medium (1024, 512) and short (200, 100) LDPC codes are implemented on Xilinx Virtex-6 VLX240T field programmable gate array (FPGA) device, with various methods.

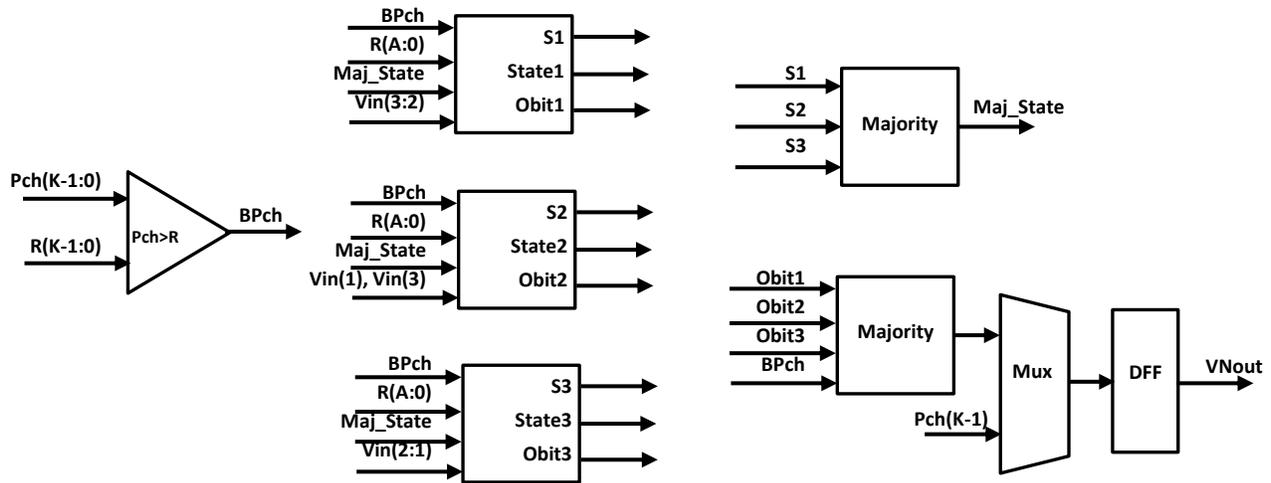


Fig. 4. The block diagram of the proposed degree-3 Variable node.

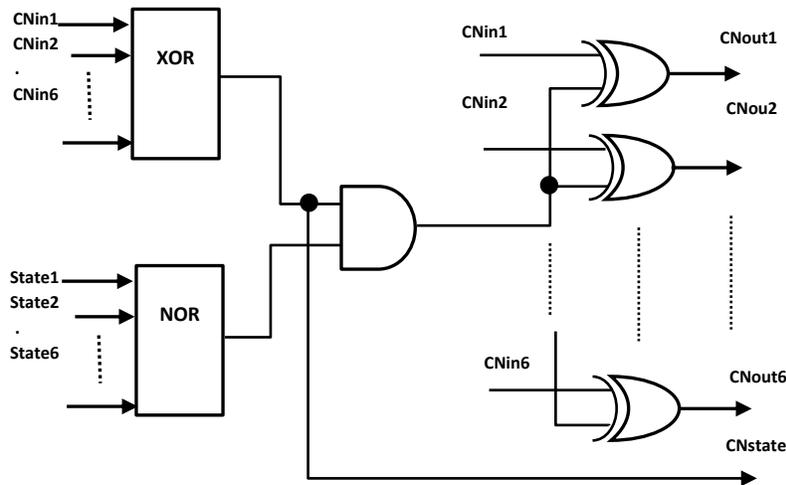


Fig. 5. The structure of the proposed degree-6 CN.

Fig. 4 presents the block diagram of the new degree-3 Variable node of the proposed (1024, 512) and (200, 100) LDPC codes. The degree-3 Variable Node is composed by 3 degree-3 sub-node. The majority state and the random address signals are connected to all sub-nodes. One of the 3 degree-3 sub-node input is connected to probability signal by a comparator. The other two sub-node inputs are connected to the check node output. The three  $S$  signals are combined to produce the majority state signal. The FPGA implementation

of VN Internal Memory is achieved by using the (LUT) Slice FPGA distributed RAMs.

The implemented CNs use similar structure to CNs adopted by the DS and the CSS decoders. Fig. 5 gives the main structure of degree-6 CN. The results of the FPGA implementation of the (200, 100) and (1024, 512) LDPC Regular Codes, with one-step initialized counter-based VN [16], DS and the proposed approach, are shown in Table 1.

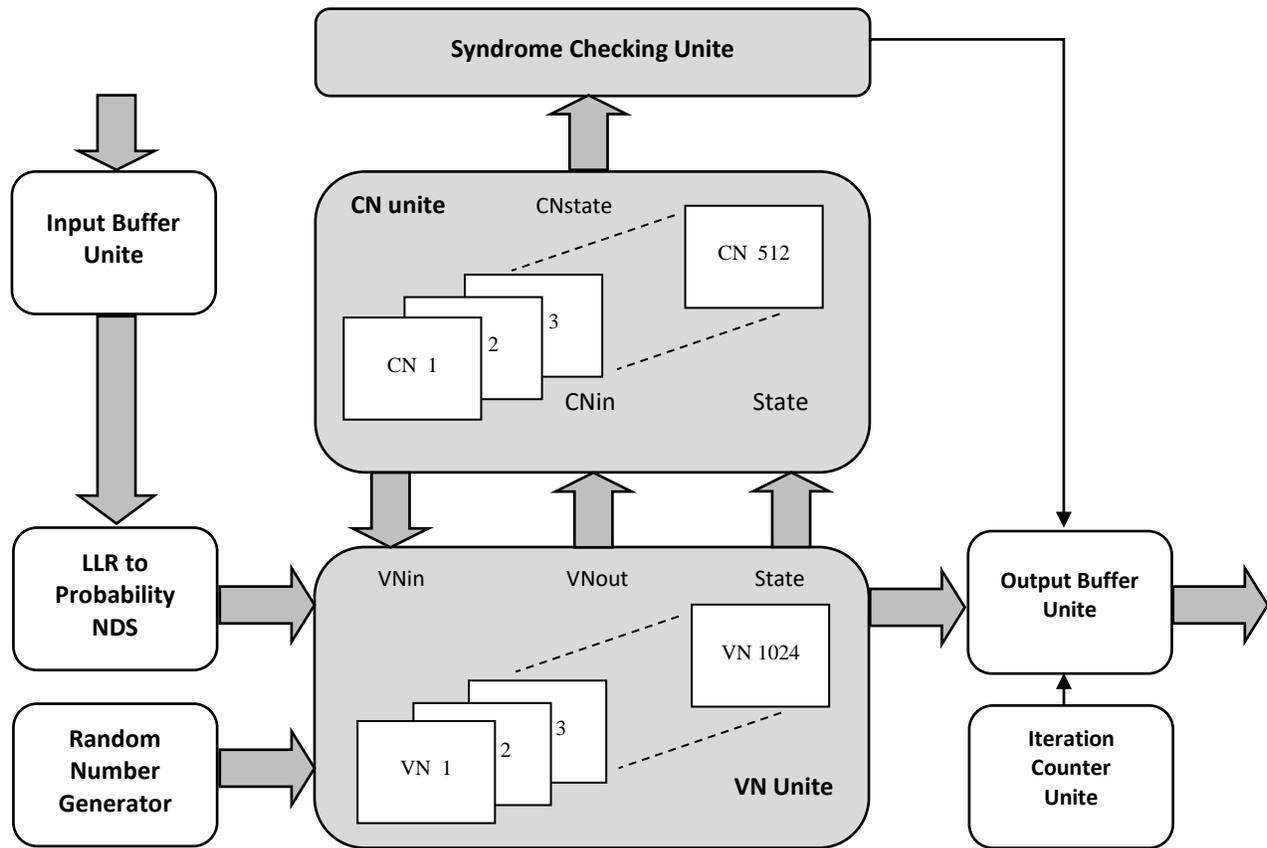


Fig. 6. The block diagram of the proposed (1024, 512) LDPC decoder.

The EM version gives an FPGA implementation result close to the DS version. The implementation of  $2 \times 1$  bits up to  $64 \times 1$  bits uses only one LUT in Virtex-6 Xilinx FPGA. Therefore, the proposed LDPC decoder version can be implemented using up to 64-bit VN internal memory without requiring additional FPGA resources. As we can see, the FPGA implementation of the one-step initialized counter based decoder need additional resources, compared to the EM and the DS versions. This disadvantage is caused by the utilisation of initializing counters instead of the VN internal memories used in DS. The additional reduction of FPGA logic utilisation seen for the new proposed decoder is principally obtained as a result of the unemployment of VN output saturated counter.

Fig. 6 presents the block diagram of the proposed (1024, 512) LDPC stochastic decoder. The main units are the variable nodes unite, the parity-check nodes unite, and the syndrome checking unite. The VN unite and CN unite exchange the stochastic information until reaching a correct code or the maximum number of iteration. The correct code is detected by the syndrome checking unite and the maximum of iterations is pointed by the iteration counter unite. The outputs of the Random Number Generator are employed with the VN comparators to generate the Bernouli sequences. Furthermore, they are directly used to drive the addresses buses of the FPGA distributed RAM, used as VN internal memory.

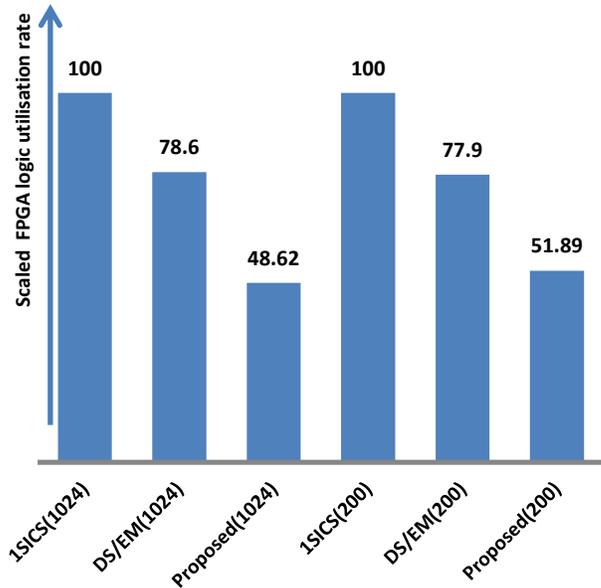


Fig. 7. FPGA logic utilisation reduction rate.

Fig. 7 shows the scaled FPGA logic utilisation rate. The proposed decoder achieves an average reduction about of 50% compared to the one-step initialized counter and an average reduction about of 35% compared to DS and EM decoders.

TABLE I. RESOURCE UTILIZATION OF STOCHASTIC LDPC DECODERS IMPLEMENTATION IN XILINX VIRTEX-6 VLX240T FPGA DEVICE

LDPC Architecture	Code	Implementation	Decoder Logic utilization (LUT)	VNs Logic utilization (LUT)	VN counter / Memory length (bits)	VNs rate utilization (%)	VNs Reduction rate compared to DS (%)	Reduction rate compared to DS (%)
Initialized Counter-based VN [16]	(1024, 512)	Stochastic fully parallel	60 947	51 200	06	84,01	- 35,13	- 27,95
	(200, 100)		11 767	10 000	06	84,98	- 35,13	- 28,36
DS [14]	(1024, 512)	Stochastic fully parallel	47 635	37 888	02	79,54	0	0
	(200, 100)		9 167	7 400	02	80,72	0	0
Proposed	(1024, 512)	Stochastic fully parallel	29 633	21 504	16	72,57	+ 43,24	+ 37,79
	(200, 100)		6 106	4 200	16	68,78	+ 43,24	+ 33,39

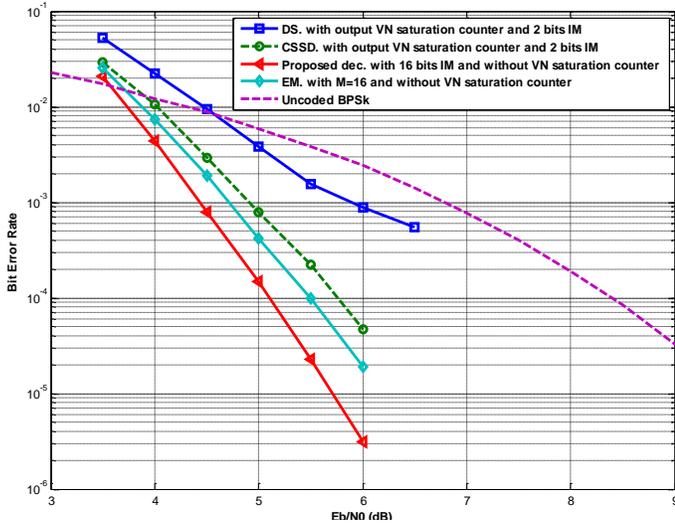


Fig. 8. BER performance of the (200,100).

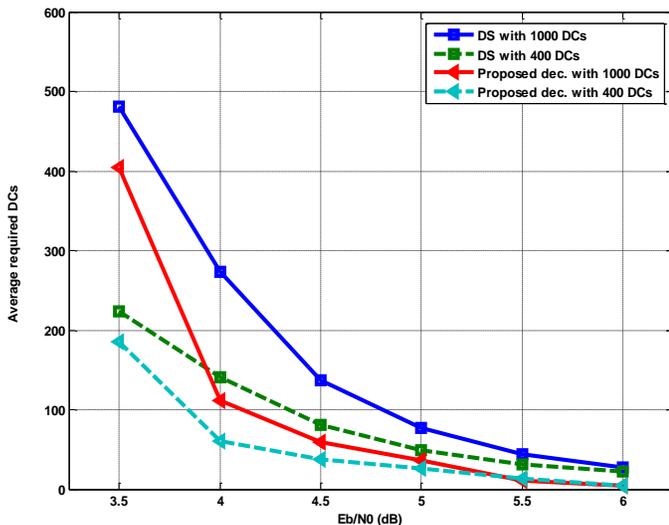


Fig. 9. Average decoding cycles.

In addition to the logic utilisation reduction, the proposed architecture offers additional performance even for short codes. Fig. 8 displays the BER performance of the (200,100) FPGA-based LDPC decoding with DS, CSS, EM and the proposed methods. The maximum decoding latency is set to 1000 DCs.

The BER of the new approach demonstrate the significant improvement. The proposed decoder with 16 bits internal memory outperforms the EM decoder by 0,5 dB at BER of  $10^{-5}$ . The average required DCs of the DS and the proposed architecture, with 1000 and 400 maximum decoding latency are presented in Fig. 9. A reduction of 85,3% and 81,1%, in the ADC compared to the DS decoder at an SNR of 6 dB, is observed for 1000 and 400 Max-DCs, respectively.

## V. CONCLUSION

In this work, we investigated the complexity and performance of FPGA-based implemented LDPC stochastic decoders. Therefore, a new fully parallel stochastic LDPC decoding approach was presented, which can outperform all state of the arte versions. The improvement was accomplished by introducing an efficient stochastic variable node.

A reduction of decoding latency and complexity, in addition of BER amelioration, are achieved even for short codes. A Xilinx Virtex-6 VLX 240T FPGA implementation results validated the advancement of the new method. An average reduction of 35% and 85% of logic utilisation and average decoding cycles respectively, compared to DS method. The BER performance of the new (200, 100) decoder exceed DS, CSS and EM versions. A gain gap of 0,5 dB, compared to EM, is observed at BER of  $10^{-5}$ .

## REFERENCES

- [1] R. G. Gallager, "Low density parity check codes," IRE Trans. Inf. Theory, vol. 8, no. , pp. 21–28, Jan. 1962.
- [2] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," Electron. Lett., vol. 32, no. 18, pp. 1645–1646, 1996.
- [3] The IEEE 802.16 Working Group [Online]. Available: <http://www.ieee802.org/16/>
- [4] The IEEE 802.3an 10GBASE-T Task Force [Online]. Available: [www.ieee802.org/3/an](http://www.ieee802.org/3/an)
- [5] IEEE Working Group for 40 Gb/s and 100 Gb/s Operation Std. [Online]. Available: <http://www.ieee802.org/3/>
- [6] The Digital Video Broadcasting Standard [Online]. Available: [www.dvb.org](http://www.dvb.org)
- [7] The IEEE 802.11n Working Group Std. [Online]. Available: <http://www.ieee802.org/11/>
- [8] V. Gaudet and A. Rapley, "Iterative decoding using stochastic computation," Electron. Lett., vol. 39, no. 3, pp. 299–301, Feb. 2003.
- [9] S. S. Tehrani, W. Gross, and S.Mannor, "Stochastic decoding of LDPC codes," IEEE Commun. Lett., vol. 10, no. 10, pp. 716–718, Oct. 2006.

- [10] S. S. Tehrani, S. Mannor, and W. Gross, "Fully parallel stochastic LDPC decoders," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5692–5703, Nov. 2008.
- [11] S. Tehrani, A. Naderi, G.-A. Kamendje, S. Mannor, and W. Gross, "Tracking forecast memories in stochastic decoders," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2009, pp. 561–564.
- [12] S. S. Tehrani, A. Naderi, G. A. Kamendje, S. Mannor, and W. Gross, "Tracking forecast memories for stochastic decoding" *J. Signal Process. Syst.* pp. 1–11, 2010 [Online]. Available: <http://dx.doi.org/10.1007/s11265-009-0441-5>
- [13] S. S. Tehrani, A. Naderi, G.-A. Kamendje, S. Hemati, S. Mannor, and W. Gross, "Majority-based tracking forecast memories for stochastic LDPC decoding," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4883–4896, Sep. 2010.
- [14] Ali Naderi, Shie Mannor, Mohamad Sawan and Warren J. Gross, "Delayed Stochastic Decoding of LDPC Codes" *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5617–5626, Nov. 2011.
- [15] Mountassar Maamoun, Rafik Bradai, Ali Naderi, Rachid Beguenane and Mohamad Sawan, "Controlled Start-Up Stochastic Decoding of LDPC Codes" in *IEEE Int. NEWCAS Conference*, Paris, France, June. 2013.
- [16] Di Wu, Yun Chen, Qichen Zhang, Yeong-luh Ueng and Xiaoyang Zeng "Strategies for Reducing Decoding Cycles in Stochastic LDPC Decoders" *IEEE Trans. Circuits and Systems II*, vol. 63, no.91, pp. 873 - 877, Sept. 2016.
- [17] Kuo-Lun Huang, Vincent Gaudet and Masoud Salehi, "Output Decisions for Stochastic LDPC Decoders" in *48th Annual Conference on Information Sciences and Systems, (CISS), Princeton, USA*, March. 2014.
- [18] Kuo-Lun Huang, "Efficient Algorithms for Stochastic Decoding of LDPC Codes" Doctor of Philosophy, Northeastern University, Boston, USA, May. 2016.

# Towards an SOA Architectural Model for AAL-PaaS

## Design and Implimentation Challenges

El murabet Amina  
Computer Engineering Department  
National school of Applied Science of Tetuan  
Tetuan, Morocco

Abtoy Anouar  
Computer Engineering Departement  
National school of Applied Science of Tetuan  
Tetuan, Morocco

Abdellah Touhafi  
Electronics-ICT Department  
Vrije Universiteit Brussel  
Brussel, Belgium

Abderahim Tahiri  
Computer Engineering Departement  
National school of Applied Science of Tetuan  
Tetuan, Morocco

**Abstract**—Ambient Assisted Living (AAL) systems main purpose is to improve the quality of life of special groups of people, including the elderly and people with physical disabilities. Driven by the critical ongoing changes in all modern, industrialized countries, there is a huge interest in IT-based equipment and services these days, to facilitate daily tasks and extend the independency time for these groups. Thence, AAL systems can benefit from the huge advances of both intelligent systems and communication technologies as promising growing research fields. The implementation of such complicated yet vital system should be established on solid bases relying on a standard architecture to satisfy and respond to the needs of heterogeneous stakeholders. This article proposes a Service Oriented Architecture model for Ambient Assisted Living Platform as a Service based on Wireless Sensors Network, it starts by presenting a classification of ambient assisted living services. Secondly, it describes some user and environmental challenges that have an impact on the service qualities. The discussion of architectural trends for AAL systems is included, and the description of challenges in designing and implementing of an effective one. Finally, this paper introduces a new vision of prototypical AAL systems architecture.

**Keywords**—Ambient Assisted Living (AAL); Ambient Assisted Living Platform as a Service (AAL-PaaS); Service Oriented Architecture (SOA); Wireless Sensors Network (WSN)

### I. INTRODUCTION

The field of AAL (Ambient Assisted Living) has taken a favorable intention between the major research fields in intelligent systems and communication technologies. AAL is taking benefits of all information technology developments, in order to denote solutions capable of improving and facilitating the life of the growing elderly population and people with physical disabilities [1]. These systems aim to help them in their daily affaires, to extend independency periods and reduce the time of needing caregivers.

The majority of the existing personal emergency response systems use emergency push buttons that can be inconvenient in critical emergencies, such as falling or unconsciousness [2]. This provokes the need of systems that do not require the involvements of the user. Therefore, there is a significant

trend to AAL systems, based on “Auto-Sensing” (Using Wireless Sensors Network and other sensing equipment) and “Auto-Acting” (Using Actuators such as alarms, phone calls, robots, etc.) which should be able to scan the local environment, obtain useful data, process this data and act according to the assembled knowledge built out of treating the collected information.

Several AAL systems proposals based their performances on different types of sensors to measure weight, blood pressure, glucose, oxygen, temperature, location, and position are available nowadays. Each system is deployed using a communication technology such as Bluetooth, USB, and Ethernet, among others [2]. In addition, the most used interfaces are developed for tablets and smartphones, although applications for health systems or set-top boxes can be found. Generally, these kinds of systems are focused on solving basic issues in services such as healthcare provision, disease management, diet and fitness, personal health records, and person location.

AAL Systems need to be affordable, considering that affording a caregiver is not an available option of all stakeholders. They also should depend on the special needs of each user because of the enormous variation of demands between elderly people and people with physical disabilities, such as the visually and hearing impaired (babies and children are also considered as a specified group of stakeholders). Having a clear classification of living assistance services is a necessity to determine the target stakeholders, and establish a model of the particular environment, to be able to draw a clear vision of the architectural model for such systems.

The heterogeneity of stakeholders gave a strong motivation for engineers to aim for a standard architecture to stand AAL systems on, so they can be easily adapted and maintained, giving the fact that AAL systems rely on various technologies and mismatched equipment.

In order to clarify an understanding of the main problem, and overall software solution and integration approaches, this paper 1) describes the classification of ambient assisted living services; 2) present some user and environmental challenges

for AAL systems; 3) briefly discusses some architectural trends; 4) proposes an SOA based model for an AAL-PaaS (Ambient Assisted Living Platform as a Service) established on WSN (Wireless Sensors Network).

## II. CATEGORIES OF AMBIENT ASSISTANCE LIVING SERVICES

To build a stable, solid and consolidated ambient assistance living system, the main concern is to determine the living assistance domains and classifying them, in order to include every assistive service that may ease daily life in all aspects. The “classification scheme” in Fig. 1 structures these domains into nine classes:

The first category is divided into “indoor” and “outdoor” living assistance. Indoor assistance services are the ones presented in a determined space: in apartments, homes, cars, hospitals, and elderly care homes. They can be built upon a well-known hardware/software installation in the specified location, thereby providing a stable environment.

Outdoor assistance services aim to support persons during activities outside their homes. It is also divided into two classes: 1) firstly at work: to allow an active and productive aging for elderly people and suitable environment for those with physical disabilities in a defined workplace and a stable environment; 2) in community: while shopping, transportation, and during other social activities. These services have to face with highly unstable environmental conditions such as special equipment and technical installations [3].

Other dimensions can be used to specify the type of service provided. According to Fig. 1, the attention is given to three types of services:

1) “Emergency treatment” presents services that aim to predict and react toward critical conditions that might result in an emergency.

	Indoor		outdoor	
		community	At work	
Emergency Treatment	Prediction Detection Prevention Action			
Autonomy Enhancement	drinking eating cleaning cooking dressing Medication scheduling	Mobility Social activities Social interaction shopping traveling banking	Collaboration Mobility Information learning	
Comfort	Safety security privacy logistics	Activity management Social inclusion Entertainment navigation	process awareness transportation Adjust workplaces	

Fig. 1. Categories of AAL services.

2) “Autonomy enhancement” services increase the independence of the assisted persons.

3) “Comfort”: These services ease the daily life but are not necessarily required. In addition, they cover all areas that do not fall into the other presented categories.

Stakeholders have different capabilities and needs which can develop over time, these needs can change and determine the categories and types of services. Therefore, this can change the presented classification.

Moreover, “emergency treatments” are considered as the main core of any AAL service portfolio, due to the increasing of emergencies coupled with the decreasing capability to deal with such circumstances.

## III. USER AND ENVIRONMENTAL CHALLENGES AND SYSTEM REQUIREMENTS FOR AAL-PAAS

### A. User and Environmental Challenges

The main goal of AAL systems is to facilitate the regular basic tasks for the stakeholders within a determined environment or space. Consequently, to help designing such systems many challenges should be taken in consideration, evaluated and treated, in order to have a consolidated and stable system. These challenges might be separated into two groups, user related and environment related ones.

Some of the main conceptual confrontations of the system are: the divergent of capabilities, needs and habits that vary from one stockholder to another, elderly people have particular demands distinct from users with physical disabilities. This diversity is also time related. Some of the individuals have an incremented demand for regular assistance due to the changing conditions in each particular situation. Moreover, users usually are not friendly to the technical problems, they do not show tolerance to the technical complications, and cannot afford maintenance every now and on. The assisted person should also be able to control the system and not vice versa. Although the system should share the user data with data centers and institutions such as hospitals, involved organizations, etc. It must preserve privacy and secure information and data of the relevant user. Hence, the system should act and react toward the user; it should also maintain his safety and conserve a highly protected space [4]. The system should also be fair in terms of space, lifetime and budget. It has to be highly productive in spite of the limited, up-to-date resources.

The user interface provokes many other challenges to the system designer. Numerous details are fundamental to be respected while the perception of the system overview, and the attention should be focused on: Human-computer interaction, where the designer is asked to take advantage of the interactive technologies, to ease behaving toward the system, the communication between the user and the system should be optimized and effective in an ergonomic way. Usability and accessibility to the system are both critical claims of all stakeholders [5]. Finally, the designer should take in consideration the information architecture and the saving politics to be followed in order to structure and organize the

data within the system itself and with other connected devices, data centers, applications, etc. to be handy for the user display.

### B. System Requirements

To help structuring a solid, reliable AAL system many requirements and major developments should be proceeding in different fields of research, among them:

- **Sensing technologies**, in AAL applications, there is a need of intelligent and innovative sensors “smart sensors” capable of collecting data, such as measuring physical and electrical quantities, miniaturized, made of low cost materials, able of taking place in anything, anywhere, anytime (home, outdoors, vehicles, public places, etc.) and qualified to perform some processing on the node level in the network.
- **Reasoning**, which is a core duty of AAL systems due to the conclusion of knowledge, resulted from processing data collected by sensors and converting it into useful information to learn from it. This should allow not only the recognition of activities, such as motions and the detection of emergencies by the use of evolutionary models; but also to predict and anticipate possible status and provide support in decision-making.
- **Event definition**, where there is a need to adapt an event driven architecture to promote the production, detection and consumption of, and reaction to events which signify a change in states. The design of AAL-PaaS should be able to transmit events among loosely coupled software components\*\* and services, this should put up with the use of cross-platform runtime environments such as Node.js.
- **Acting**, systems and services, which proactively (based on the knowledge resulted from the reasoning) act to prevent, compensate, support and provide well-being and increase the independence of senior citizens.

Finally, some primary conditions are not to be ignored and should be respected such as affordability, usability, suitability, dependability, adaptivity, extensibility, resource efficiency and heterogeneity that should be the main characteristics of such a system.

By respecting all of the quoted earlier standards, the main work can be driven toward a comprehensive and complete architectural model vision of the desirable system to implement.

### IV. ARCHITECTURAL TRENDS FOR AAL SYSTEMS

The architecture of a system gives an overview of the desirable system to be implemented; it is the central design that describes quality requirements such as cost, dependability, performance, etc. of the overall solution. It plays a pivotal role for the quality achievement. Furthermore, it comprises: software elements, the externally visible properties of those elements and the relationships among them. For the time been there is no commonly accepted architecture for AAL systems, different approaches are followed to meet the functional and quality requirements in the present and developing systems [6].

AAL systems are systems which provide assistance that has two facets: 1) an easy access for the assisted person to autonomy enhancement or comfort services, home control, social interaction, etc.; 2) the anticipatory assistance of the assisted person with proactive emergency treatment such as automatic alarms, home automation, notifications, etc. for anticipatory systems the rendered functionality as be: Awareness or Presence. The system should also be kind of closed loop controller that senses its environment and especially the persons living therein and influences the environment with its actuators.

Awareness of the system can be decomposed in three functional blocks: Sensing/perception/identification while the presence is decomposed into planning/controlling/acting. This decomposition is clear in the following Fig. 2 presentation in which the system should exchange data with the surrounded environment and the target stakeholder while facing the mentioned challenges and respecting the required principals.

A common style for ambient intelligence systems are data processing pipeline or signal [3] which made communication fluent between the elements of the system.

Several ways are conceivable to realize the conceptual decomposition of functionality in the system physically.

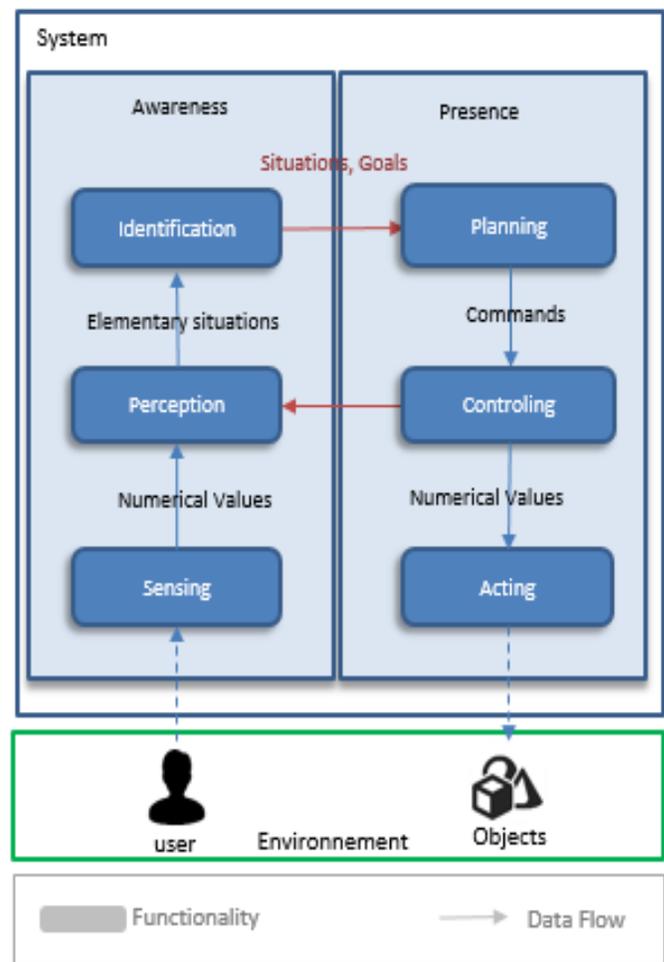


Fig. 2. Functional blocks of AAL system [3].



Fig. 3. Physical perspective of AAL System.

From a physical perspective: the topology of the AAL system will consist of tens to hundreds of different interacting nodes [using the sensors] ranging from tiny sensor nodes [with low computational power] up to powerful machines, which are interacting together.

Fig. 3 presents generally a physical overview of an AAL system with its major elements, the observation drive us to the fact that an AAL system should rely on using several nodes presenting the sensors, actuators and data servers in a direct interaction between one another and toward the user.

#### V. SOA MODEL FOR AN AAL-PAAS BASED ON WSN

An architecture model is a division of the functionality together with data flow between the pieces [5], [7]. It is an abstract representation of a system expressed primarily using software components interacting via connectors.

In this part of the paper we are trying to implement an architectural model of an AAL-PaaS system based on WSN (Wireless sensors Network) using the SOA (Service Oriented Architecture) as a promising architectural style in the AAL domain and the revolutionary technologies such as “Cloud” and “Node.js” as a successful development technology to build data-intensive applications [8].

When implementing such an AAL system, all the above-mentioned features should be taken in consideration. On one hand, sensors should be non-invasive systems they can be embedded in clothes, shoes, watches, or glasses [9], [10].

Thus, people will not mind wearing them. If sensors were visible, users could be discriminated against by other persons [4]. In addition, sensors must have wireless communication interfaces to let people move away from their homes. Smart phones, actuators, computer hardware, computer networks, software applications, should be interconnected together to collect and exchange data and provide services in an Ambient Assisted environment. The sensors and the actuators are connected with the AAL applications to send medical data to the health monitoring systems. Sensors should not be the only source of data; the system should examine the historical information of the stockholder [11] taking from other data sources such as Hospitals data servers, should also rely on existing shared applications and services through the “Cloud”.

After collecting data, the system should be able to integrate the information, analyze it, treat it and transform it into a useful knowledge, in which the system relies on to provoke actions. Service Oriented Architecture is widely regarded as the software paradigm of the next decade, especially in the field of information systems [12], [13]. A central quality of SOA is to support an easy exchange of implementation and orchestration of new functionality (which contributes positively to the modifiability and extensibility of the systems) by separating the contract (service) from the implementation (component), in this way; this architecture increases the reusability of the service and components.

Using SOA architecture means dividing the system into clear and reusable services, this approach gave us the opportunity also to use existing services and adapting them into the newly designed system. In this case, there is a necessity to assure interoperability and ease the communication between these services. In this determination, the system should obey to some standards: to guarantee communication without particular concern, to give services the possibility to evolve independently without the risk of breaking predefined communication.

In SOA architectures, the process choreographer has been chosen as a useful architecture model used for designing and implementing communication between mutually interacting software applications and services. It provides routing, transformation, mediation, etc. It presents the communication bus between the different services and processes of the proposed architectural model.

The design of AAL-PaaS should be able to transmit events among loosely coupled software components (in which each of its components makes use of, little or no knowledge of other separate components) and services, this should put up with the use of cross-platform runtime environments such as “Node.js” which has an event-driven architecture capable of asynchronous I/O [Input /Output] [14]. These design choices meant to optimize throughput and scalability in Web applications with many input/output operations, as well as for real-time Web applications.

In Fig. 4, we propose an architectural model based on the logical view of the software architectural design describing how the system is structured.

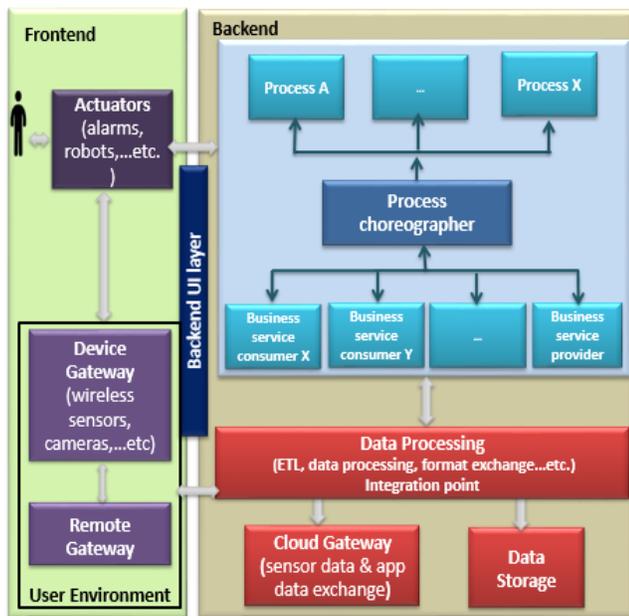


Fig. 4. AAL architectural model based on SOA.

In our approach, the implementation of the architectural model is based on SOA. The situational and environmental information about people, places and things is collected and stored into the system by means of a sensing environment installed in the indoor environment including: wireless sensors, cameras, accelerometers and movement detectors, Audio/Video perceptual components, etc. which are figured by **Device Gateway** and **Remote Gateway**. Other **Data Storages** used to gather information about historical states of illnesses and health care parameters. These data and information are processed together using **Data Processing** elements such as ETL (Extract, Transform, Load) [15] software in an integration point to gather, extract, transform and unify the collected data in order to pull knowledge from it.

Services in the architectural model are connected in order to coordinate the exchange between the **business services consumers** and **services providers**, the process choreographer to assure coherent between services and defines the appropriate processes to be launched by the defined actuator, which are equipment capable of performing an action towards the user or his environment. Their actions are stimulated based on the knowledge resulted from processing the collected data. These actuators may be Robots, alarms, etc.

## VI. CONCLUSION

In order to introduce our overview of an AAL system architecture based on SOA and EDA (Event Driven Architecture), a study of the classification of ambient assisted living services was made. From the research that has been carried out, we determined the major user and environmental challenges, to be taken into consideration while designing such systems. Furthermore, a description of the major technical requirements to build a solid, reliable system has been clarified. Finally, an illustration of our vision has been presented using a model scheme where the major layer

compositions of a firm AAL system are included. The next approach is focalized on prefunding the research in each of the architectural components and examining particular integration elements distinctively data integration and service integration components in the interest of having a harmonious, homogeneous system architecture. Future work will involve a study on the AAL data collection and integration, considering the bigness of the data, the variety of sources and the nature of the information provided.

## REFERENCES

- [1] P. Rashidi and A. Mihailidis, "A Survey on Ambient-Assisted Living Tools for Older Adults," in *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 579-590, May 2013.
- [2] J. Lloret, A. Canovas, S. Sendra, et L. Parra, "A smart communication architecture for ambient assisted living", *IEEE Communications Magazine*, vol. 53, no 1, p. 26-33, janv. 2015.
- [3] M. Becker, "Software Architecture Trends and Promising Technology for Ambient Assisted Living Systems", *Dagstuhl Seminar Proceedings*, 2008.
- [4] P. Koleva, K. Tonchev, G. Balabanov, A. Manolova, et V. Poulkov, "Challenges in designing and implementation of an effective Ambient Assisted Living system", in *2015 12th International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services (TELSIKS)*, 2015, p. 305-308.
- [5] Röcker, C. (2011). *Designing Ambient Assisted Living Applications: An Overview of State-of-the-Art Implementation Concepts*. In: X. Chunxiao (Ed.): *Modeling, Simulation and Control, Proceedings of the International Conference on Information and Digital Engineering (ICIDE 2011)*, September 16- 18, Singapore, pp. 167 - 172.
- [6] M. Klein, A. Schmidt, and R. Lauer, "Ontology-centred design of an ambient middleware for assisted living: The case of soprano," in *Proc. Annu. German Conf. Artif. Intell.*, 2007.
- [7] P. Wolf, A. Schmidt, J. P. Otte, M. Klein, S. Rollwage, B. König-ries, T. Dettborn, et A. Gabdulkhakova, *openAAL 1- the open source middleware for ambient-assisted living (AAL)*. 2010.
- [8] K. Lei, Y. Ma, and Z. Tan, "Performance Comparison and Evaluation of Web Development Technologies in PHP, Python, and Node.js," in *2014 IEEE 17th International Conference on Computational Science and Engineering*, 2014, pp. 661-668.
- [9] C. S. Raghavendra, K. M. Sivalingham, and T. Znati, *Wireless Sensor Networks*. Springer, 2006.
- [10] S. Kurt, H. U. Yildiz, M. Yigit, B. Tavli, and V. C. Gungor, "Packet Size Optimization in Wireless Sensor Networks for Smart Grid Applications," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2392-2401, Mar. 2017.
- [11] K. Häyrynen, K. Saranto, and P. Nykänen, "Definition, structure, content, use and impacts of electronic health records: A review of the research literature," *International Journal of Medical Informatics*, vol. 77, no. 5, pp. 291-304, May 2008.
- [12] H. Fouad, A. Gilliam, S. Guleyupoglu, and S. M. Russell, "Automated evaluation of service oriented architecture systems: a case study," 2017, vol. 10207, p. 102070J-102070J-8.
- [13] Dhara, Kamala Manasa; Dharmala, Madhuri; and Sharma, Chamma Krishan, "A Survey Paper on Service Oriented Architecture Approach and Modern Web Services" (2015). *All Capstone Projects*. 157
- [14] H.-K. Ra, H. J. Yoon, A. Salekin, J.-H. Lee, J. A. Stankovic, and S. H. Son, "Software Architecture for Efficiently Designing Cloud Applications Using Node.js," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services Companion*, New York, NY, USA, 2016, pp. 72-72.
- [15] M. J. Denney, D. M. Long, M. G. Armistead, J. L. Anderson, and B. N. Conway, "Validating the extract, transform, load process used to populate a large clinical research database," *International Journal of Medical Informatics*, vol. 94, pp. 271-274, Oct. 2016.

# Core Levels Algorithm for Optimization: Case of Microwave Models

Ali Haydar

Department of Computer Engineering  
Girne American University  
Mersin-10, Turkey

Kamil Dimililer

Department of Electrical and Electronics Engineering  
Girne American University  
Mersin-10, Turkey

Ezgi Deniz Ülker

Department of Computer Engineering  
European University of Lefke  
Mersin-10, Turkey

Sadık Ülker

Department of Electrical and Electronics Engineering  
European University of Lefke  
Mersin-10, Turkey

**Abstract**—Metaheuristic algorithms are investigated and used by many researchers in different areas. It is crucial to find optimal solutions for all problems under study especially for the ones which require sensitive optimization. Especially, for real case problems, solution quality and convergence speed of the algorithms are highly desired characteristics. In this paper, a new optimization algorithm called Core Levels Algorithm (COLA) based on the use of metaheuristics is proposed and analyzed. In the algorithm, two core levels are applied recursively to create new offsprings from the parent vectors which provides a desired balance on the exploration and exploitation characteristics. The algorithm's performance is first studied on some well-known benchmark functions and then compared with previously proposed efficient evolutionary algorithms. The experimental results showed that even at the early stages of optimization, obtained values are very close or exactly the same as the optimum values of the analyzed functions. Then, the performance of COLA is investigated on real case problems such as some selected microwave circuit designs. The results denoted that COLA produces stable results and provides high accuracy of optimization without high parameter dependency even for the real case problems.

**Keywords**—Metaheuristic algorithms; evolutionary algorithms; microwave circuits, optimization

## I. INTRODUCTION

Solution of optimization problems is an interesting field of study for various areas such as array antenna synthesis [1]-[3], financial analysis [4], [5], error minimization and game programming [6], [7], microwave design [8]-[10] and data mining [11], [12]. Most of the related algorithms are motivated from the nature and are aimed to find near optimal solutions of given problems [13]-[17]. The performances of proposed algorithms are usually represented with their solution quality and convergence speeds. Each algorithm has several control parameters which are needed to be well tuned depending on the optimization problem in order to achieve better performance. This can be considered as a vital step in most of the cases and affects the exploration and exploitation characteristics of the algorithms [18]-[20]. The number of control parameters and their adjustment are quite deterministic

for the performance of the algorithms. Typically, an algorithm that needs a few control parameters is assumed as a good choice in solving a given problem. However, in some cases, even the fine adjustment of control parameters is not sufficient to find the optimal points of the problems. In such problems, the exploration strategies of the algorithms may not be adequate to converge to the global optimum of the given problem. Therefore, some modifications of the algorithms are proposed to solve these specific problems [21]-[24]. Introduction of new optimizers is still an open area of research, because of the lack of an optimization algorithm that performs well in all fields. Some optimization algorithms perform well for some problems, while perform inadequate for other problems [25]. In the literature, many different optimization algorithms have been proposed to increase the solution quality for complex optimization problems with as little effort as possible.

In many microwave design problems, it is required to deal with some highly nonlinear objective functions with a large number of variables. In addition, gradient based algorithms cannot yield sufficient solutions in most of the cases since the optimization parameters in most of the problems are highly coupled with each other. Evolutionary algorithms are widely used when the analytical methods are insufficient to obtain appropriate solutions [26]-[28]. Although Genetic Algorithm (GA) is the first dominant evolutionary algorithm, which was applied on microwave and electromagnetic based problems, Differential Evolution (DE) and Particle Swarm Optimization (PSO) algorithms including their variants dominate the other evolutionary algorithms in this field [3], [29], [30].

In this paper, a new evolutionary algorithm called Core Levels Algorithm (COLA) is proposed to solve complex optimization problems including the design problems for microwave circuits. COLA uses the similar steps of evolutionary algorithms such as selection of candidates, generation of offsprings and replacement of the parents with the new offsprings which have better fitness values. In addition to that, COLA focuses on two core levels to obtain better offspring candidates by their parent vectors. A balance

is achieved between the core levels to perform iterative exploitation in near optimal regions and to have better exploration in the search region. Moreover, the second core level of the algorithm is designed in such a way that it performs an exploration scheme which is centered on the selected vector and covers the whole domain of interest in order to solve different kind of optimization problems without modifying the algorithm.

Optimization characteristics of COLA were studied with the use of benchmark functions. In this study, the correlation between the number of selected parent vectors and the optimization capability was observed. Also, a comparative analysis of COLA with DE, PSO and Harmony Search (HS) algorithms was done for multimodal functions to observe the benefits of COLA. Another goal of the paper was to verify that COLA is applicable to the real case problems. Therefore, two microwave models were selected and optimized to achieve this goal.

The rest of the paper is organized as follows: Section II introduces the main concept of COLA in detail. Section III focuses on benchmark function results, real case microwave problems and discussions. Lastly, Section IV summarizes the obtained results and concludes the paper.

## II. CORE LEVELS ALGORITHM (COLA)

The Core Levels Algorithm implements a new method to find the global optimum of a given function which is mainly based on the use of the balance between two core levels. This provides good exploration and exploitation characteristics of the algorithm. The pseudo code of COLA is shown in Fig. 1 and the detailed steps of COLA are explained in the following paragraphs:

- Initialize the population randomly using uniform distribution in the related domain of the optimization problem.

The algorithm has a few control parameters to be set which are  $np$ ; the number of elements in the population and  $k$ ; the number of elements to be combined. COLA starts by initializing randomly the solution set of the optimization problem. The initialization can be defined as follows:

$$x_{i,j} = x_{\min,j} + rand[0,1](x_{\max,j} - x_{\min,j}), \quad (1)$$

Where,  $i \in \{1, 2, \dots, np\}$  and  $j \in \{1, 2, \dots, D\}$ . Here,  $D$  represents the dimension of given problem,  $x_{\min,j}$  and  $x_{\max,j}$  represent the lower and upper bounds for the  $j$ th variable respectively and  $x_{i,j}$  is the  $j$ th component of the  $i$ th solution vector.

- Evaluate the fitness of each element in the population.
- Select  $k$  number of solutions according to their fitness values.

Using roulette wheel selection,  $k$  number of solutions are selected randomly according to their fitness values.

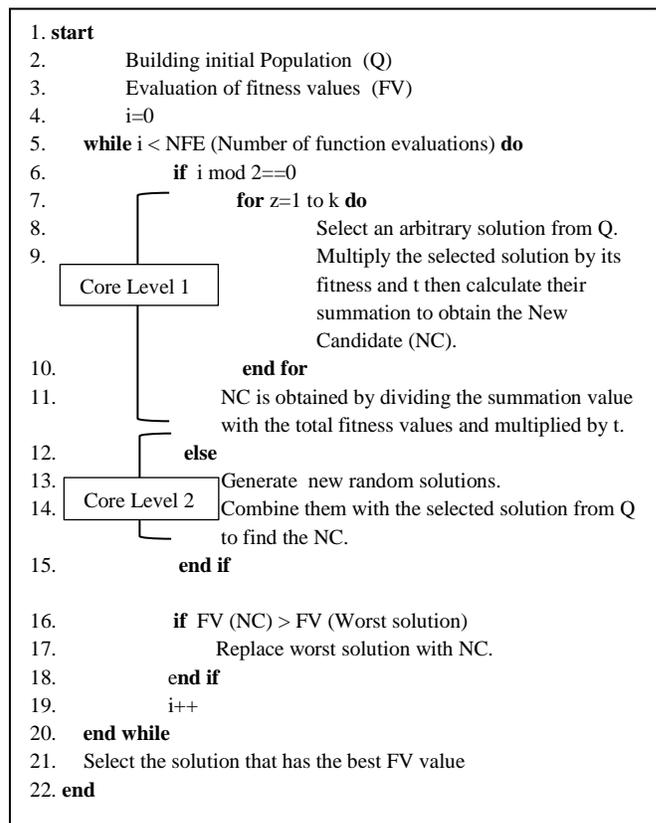


Fig. 1. Pseudo code of COLA.

- Obtain a new candidate solution from the selected solution and the randomly generated solution by using two core levels.

The new candidate solution vector is formed which is a weighted sum of the selected  $k$  number of solution vectors by core level 1. New candidate solution vector (NC) is calculated using the following expression:

$$x_j^{new} = \frac{t_1 fit_1 x_{S_1,j} + t_2 fit_2 x_{S_2,j} + \dots + t_k fit_k x_{S_k,j}}{\sum_{n=1}^k t_n fit_n}, \quad (2)$$

Where,  $x_j^{new}$  is the  $j$ th component of the new candidate vector,  $t_i \in \{1, -1\}$  and is selected as -1 or 1 with equal probability,  $fit_p$  is the fitness of the  $p$ th selected solution vector and  $x_{S_p,j}$  is the  $j$ th component of the  $p$ th selected vector. The new candidate formation using (2) is performed once in every two generations of the candidate vector.

Core level 2 is implemented for the next generation of the new candidate by using the following expression:

$$x_j^{new} = (1 - \alpha)x_{S,j} + \alpha x_{rand}, \quad (3)$$

where,  $x_j^{new}$  is the  $j$ th component of the new candidate vector,  $\alpha$  is a random number in the uniformly distributed interval

$[0, 1]$ ,  $x_{S,j}$  is the  $j$ th component of the randomly selected  $S$  vector from the population and  $x_{rand}$  is a randomly obtained number in the related domain which is calculated using (1). In the Core level 2, new vector  $x_j^{new}$  can be obtained by applying in three different methods.

- 1) Only one of the randomly selected components of  $x_S$  is updated.
- 2) The  $k$  number of components in the range of  $1 < k < D$  are updated where the number of  $k$  is selected randomly.
- 3) All of the components of  $x_S$  are updated by their linear combination with random numbers.

For each core level 2 application, only one of these methods is performed and this selection is done recursively.

- Replacement of the new candidate solution with the parent vector, if the fitness value of the new candidate is better.

The fitness of new vector is compared with the fitness of the worst parent solution vector in the population. If the fitness of new vector is better than fitness of the worst parent solution vector, then the new vector replaces with the worst parent solution vector in order to advance to the next generation. This can be expressed as follows:

$$x^{worst} = \begin{cases} x^{new} & \text{if } fitness(x^{new}) < fitness(x^{worst}) \\ x^{worst} & \text{otherwise} \end{cases} \quad (4)$$

- The steps will be repeated until the stopping criterion is met.

### III. EXPERIMENTAL RESULTS AND DISCUSSIONS

#### A. Testing with benchmark functions

In this section, COLA is applied to 12 well-known benchmark functions to demonstrate its performance. These functions are taken from literature and they have been widely used for testing of optimization problems [31]. The selected benchmark functions are shown in Table 1. Among these functions, the first seven functions are unimodal functions and the following five functions are multimodal functions. For unimodal functions, convergence rates are the distinguishing characteristics of the optimization algorithms rather than final results. However, for multimodal functions, due to the many optimum points of problems, the final result obtained by algorithm is significant. The presented experimental results are *average*, *standard deviation* and the *best value* of the functions. All values are gathered over 40 independent runs. *Average* value indicates the solution quality, *standard deviation* value specifies the stability of the algorithm for undergoing random operations and the *best value* simply expresses the closest result to the optimal solution out of 40 independent runs.

TABLE I. THE SELECTED BENCHMARK FUNCTIONS USED IN EXPERIMENTS

Function name	Function expression	Domain	$f_{min}$
Sphere	$\sum_{i=1}^n x_i^2$	[-100, 100]	0
Schwefel 2.22	$\sum_{i=1}^n  x_i  + \prod_{i=1}^n  x_i $	[-10, 10]	0
Schwefel 1.2	$\sum_{i=1}^n \left( \sum_{j=1}^i x_j \right)^2$	[-100, 100]	0
Schwefel 2.21	$\max\{ x_i , 1 \leq i \leq n\}$	[-100, 100]	0
Rosenbrock	$\sum_{i=1}^n \left[ 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \right]$	[-30, 30]	0
Step	$\sum_{i=1}^n (\lfloor x_i + 0.5 \rfloor)^2$	[-100, 100]	0
Quartic	$\sum_{i=1}^n ix_i^4 + random[0,1]$	[-1.28, 1.28]	0
Schwefel	$\left[ \sum_{i=1}^n -x_i \sin(\sqrt{ x_i }) \right]$	[-500, 500]	-12569.5
Rastrigin	$\sum_{i=1}^n \left[ \frac{x_i^2}{10} - \cos(2\pi x_i) + 10 \right]$	[-5.12, 5.12]	0
Ackley	$\left[ -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e \right]$	[-32, 32]	0
Griewank	$\left[ \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \right]$	[-600, 600]	0
Penalized	$\left[ \frac{\pi}{n} \left\{ 10 \sin^2(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 \left[ \frac{1 + 10 \sin^2(\pi y_{i+1})}{(y_{i+1})^2} + (y_n - 1)^2 \right] \right\} + \sum_{i=1}^n u(x_i, 10, 100, 4) \right]$ $y_i = 1 + \frac{1}{4}(x_i + 1)$ $u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m, & x_i > a \\ 0, & -a \leq x_i \leq a \\ k(-x_i - a)^m, & x_i < -a \end{cases}$	[-50, 50]	0

Population size is fixed to 100, dimension is set to 30 and the number of function evaluations is set to 10000 for all benchmark functions. The algorithm continues until the stopping condition is met.

The results obtained for the listed functions above are given in Table 2. In this table, it is aimed to observe the performance of COLA for different  $k$  parameter values in the set {3, 4, 5} and a random selection of  $k$  from the same set for each iteration.

TABLE II. THE DESCRIPTIVE VALUES OF BENCHMARK FUNCTIONS BY SELECTED NUMBER OF  $k$  PARAMETER AMONG THE PARENT VECTORS

10000 function evaluations		$k = 3$	$k = 4$	$k = 5$	$k = \text{rand}[3-5]$
Sphere	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Schwefel 2.22	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Schwefel 1.2	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Schwefel 2.21	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Rosenbrock	Avg	0.002908	0.000315	0.002061	1.92E-4
	Best	2.92E-27	2.92E-27	2.92E-27	2.92E-27
	Stdev	0.007363	0.000647	0.000228	0.000227
Step	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Quartic	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Schwefel	Avg	-12569.5	-12569.5	-12569.5	-12569.5
	Best	-12569.5	-12569.5	-12569.5	-12569.5
	Stdev	0	0	0	0
Rastrigin	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Ackley	Avg	-1.4E-16	-1.4E-16	-1.4E-16	-1.4E-16
	Best	-1.4E-16	-1.4E-16	-1.4E-16	-1.4E-16
	Stdev	0	0	0	0
Griewank	Avg	0	0	0	0
	Best	0	0	0	0
	Stdev	0	0	0	0
Penalized	Avg	3.67E-6	2.58E-6	2.18E-6	1.74E-6
	Best	1.36E-22	1.36E-22	1.36E-22	1.36E-22
	Stdev	3.2E-06	1.38E-6	1.09E-6	8.43E-7

It can be observed from the results in Table 2, except *Rosenbrock*, *Ackley* and *Penalized* functions, all other functions are converged to their optimal values regardless of selection of  $k$ . Moreover, when the results of *Rosenbrock* and *Penalized* functions are investigated, instead of selecting  $k$  parameter as a fixed number, the selection of it randomly in given interval for each iteration decreases the *average* and *standard deviation* values. This is the indication of better solution quality and faster convergence speed of the algorithm under the random selection of  $k$  parameter in the given interval. For these two functions, the *average* and *standard deviation* are shown for different values of  $k$  in Fig. 2 and 3, respectively.

The figures illustrate that selecting parameter  $k$  randomly in a given interval for each iteration presents better results. It is obvious that COLA performs effectively to reach to the optimum points of *Rosenbrock* and *Schwefel* functions. Since the control parameter  $k$  is randomized in the given interval, COLA performs efficiently by using only a single control parameter which is  $np$ ; the number of elements in the population.

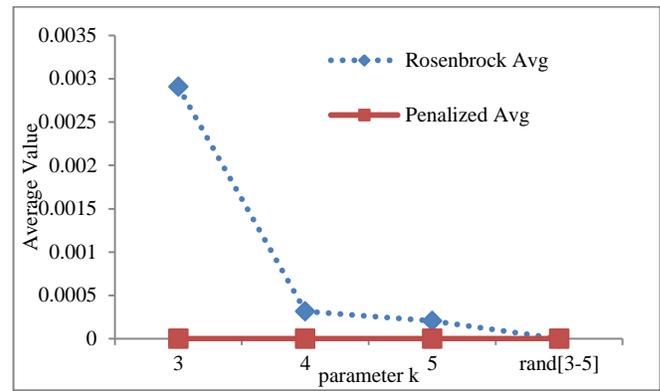


Fig. 2. Average function values obtained for *Rosenbrock* and *Penalized* functions for different values of  $k$  parameter.

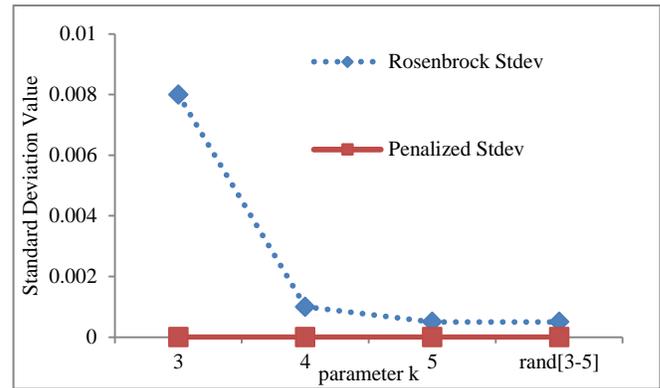


Fig. 3. Standard deviation of function values obtained for *Rosenbrock* and *Penalized* functions for different values of  $k$  parameter.

In order to verify the advantage of COLA over some effective evolutionary algorithms which are Differential Evolution, Particle Swarm Optimization and Harmony Search algorithms, it is compared for the multimodal benchmark functions under the same conditions. These comparative results are presented in Table 3. The results for DE, PSO and HS algorithms are taken from the previously reported work [32]. Since the global optimum point is located in many local optimum points, it is quite challenging to reach the global optimum of given multimodal functions. It is also known that the solution quality of the last value achieved by an algorithm is a distinguishing characteristic especially for multimodal functions. Therefore, a comparison analysis is performed only for the multimodal functions. It can be seen from Table 3 that COLA optimizes the multimodal functions with a good convergence rate, while the others are away from the global points of the functions even after 10000 number of function evaluations.

### B. Microwave Models

Importance of optimization and computer design has been realized for many years. One of the earliest papers on the area of optimization methods for microwave circuits was Bandler and Macdonald's work [33], [34]. A classical paper on the analysis part of microwave circuits in computer aided design was introduced by Monaco and Tiberio [35].

TABLE III. COMPARATIVE VALUES OF ALGORITHMS FOR MULTIMODAL FUNCTIONS.

Function name		COLA	DE	PSO	HS
Schwefel	Avg	-12569.5	-7485.74	-8531.08	-12554.6
	Best	-12569.5	-8128.58	-10353.9	-12566.1
	Stdev	0	270.62	949.247	28.8299
Rastrigin	Avg	0	137.899	46.3655	18.1256
	Best	0	126.013	19.0816	12.7443
	Stdev	0	6.6812	17.416	3.41769
Ackley	Avg	-1.4E-16	16.7758	2.86698	1.09805
	Best	-1.4E-16	15.1746	0.65150	0.56317
	Stdev	0	0.75719	1.01934	0.29879
Griewank	Avg	0	1.53190	0.39352	1.04977
	Best	0	1.29684	0.05316	0.56317
	Stdev	0	0.19544	0.31861	0.0222
Penalized	Avg	1.74E-6	5.08107	4.36306	0.29210
	Best	1.36E-22	2.79334	0.37862	0.04269
	Stdev	8.43E-7	2.13136	2.94708	0.24432

In this paper, different methods used in the analysis programs of linear circuits in frequency domain were described. Also, determination of sensitivity and convenience of using one or the other method in relation to the number of parameters and different analysis methods were explained.

Different methodology which used the combination of experimental design and computer-aided design was demonstrated in [36]. In 2002, computer-aided design summary of works to date is included as a survey paper [37]. It is also indicated that there are three essential reasons for simulation of radio frequency and microwave circuits; to understand the physics of a complex system of interacting elements, to test new concepts and to optimize the designs. Over the years many papers for the computer aided design or optimization of microwave circuits can be found. Recently a new technique for rapid multi-objective optimization of the compact microwave passive components was presented [38], [39].

### C. Microwave Tapered Matching Circuit Design

Microwave matching networks are important in the design of many different types of microwave circuitry. Only with proper matching such a circuit can attain maximum power transfer and eliminate the reflection. In microwave matching circuit design, especially when one needs to match real load impedances, one of the most useful network is a tapered microwave matching network which can be considered as a series of cascaded quarter wavelength transmission lines. The design for tapered lines is usually done by using computer algorithms for continuous sections [40], [41]. For this kind of structure design optimization using nature-inspired metaheuristic methods, namely, particle swarm optimization, was done [10]. An example circuit is shown in Fig. 4.

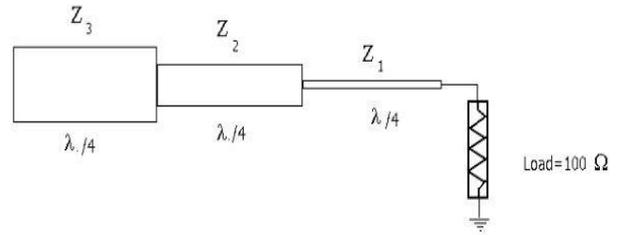


Fig. 4. Microwave Tapered matching network.

As can be seen from Fig. 4, the tapered line in this model of study consists of a series of  $\lambda/4$  transmission lines. This type of configuration can be used to achieve a match for real loads. In the model,  $100\Omega$  load is matched to a  $50\Omega$  line using a series of cascaded three transmission lines of a quarter wavelength long. Starting from the load higher impedance transmission lines exist and as moving along the tapered line, lower impedance transmission lines are obtained. The objective function of this circuit is given as below:

$$f(Z_1, Z_2, Z_3) = \frac{1}{100} \left[ \frac{Z_1 \cdot Z_3}{Z_2} \right]^2 - 50, \tag{5}$$

Where, the values of parameters, which are the characteristic impedances,  $Z_1$ ,  $Z_2$  and  $Z_3$  must be found also with the condition that  $Z_3 < Z_2 < Z_1$  and the values are restricted to be in the range [0-100]. In order to observe the performance of COLA on this microwave circuit, 40 independent runs were performed. For each run, 1000 function evaluations were executed to obtain values of  $Z_1$ ,  $Z_2$  and  $Z_3$  that provide optimal value of function (5). From the performed 40 independent trials, a sample of five impedance values ( $\Omega$ ) are shown in Table 4.

The reflection coefficient obtained by using first and the second trial values are plotted in Fig. 5 for the design frequency of 5 GHz to demonstrate the impedance values. As it can be seen in Fig. 5, reflection coefficient values are around  $-35$  dB and below at 5 GHz, which indicates a good matching. This indicates that the values obtained by COLA are all correct yielding proper designs at the end.

TABLE IV. IMPEDENCE VALUES FOR FIVE INDEPENDENT TRIALS

	$Z_1$ ( $\Omega$ )	$Z_2$ ( $\Omega$ )	$Z_3$ ( $\Omega$ )
Trial 1	81.809594	50.364347	43.531533
Trial 2	93.715554	63.237186	47.714004
Trial 3	82.929957	61.988993	52.855251
Trial 4	82.588533	66.591554	57.014379
Trial 5	92.639060	57.197776	43.658620

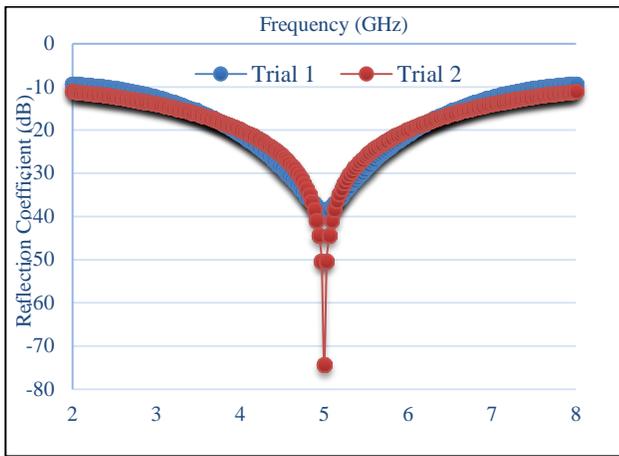


Fig. 5. Reflection coefficient values for two different trial sets.

Another aspect that needs to be analyzed is the sensitivity analysis of the obtained results using the algorithm. Simple 5% error is introduced to the obtained values to see how the design is affected. In this case it is observed that there is not a change in design operating frequency since the lengths are not different. The reflection value is worse, however still lower than -25 dB which yields a reasonable match since the reflection coefficient value is very low.

#### D. Microwave Amplifier Design

Amplification is necessary for most of the electronic circuits and for microwave circuit systems. Nowadays, with the development of transistor technology, all microwave amplifiers use transistor devices which are more reliable and rugged. The main advantage of using transistor devices is that they can easily be integrated into monolithic circuits. Design of amplifiers in general requires the matching network design for input and output parts of the network. If the work is done by hand, first the stability of the transistor is checked and drawing the Gain Circles and selecting optimum points, one can perform the design operation using Smith Chart. This process can be performed by using metaheuristic algorithms especially the nature-inspired metaheuristic algorithms. Similar works, using metaheuristic algorithms to solve amplifier design problems, were performed by the following researchers for the given specific problems [9], [42]-[46]. Simple two-port microwave network which produces amplification with a proper design is shown in Fig. 6.

In the Fig. 6, there are two matching network designs that should be done simultaneously to achieve the desired gain. The overall design also needs a compromise in gain and in return loss at the same time. The design in this case is the design of two impedance matching networks to achieve the desired gain goal. In other words, the design requires finding the proper lengths of transmission lines  $d_1$ ,  $d_2$ ,  $l_1$  and  $l_2$  at the central operation frequency. The power function that needs to be optimized is given by the following expression:

$$G_T = \frac{|S_{21}|^2 (1 - |\Gamma_S|^2) (1 - |\Gamma_L|^2)}{|(1 - S_{11}\Gamma_S)(1 - S_{22}\Gamma_L) - (S_{12}S_{21}\Gamma_S\Gamma_L)|^2} - 16. \quad (6)$$

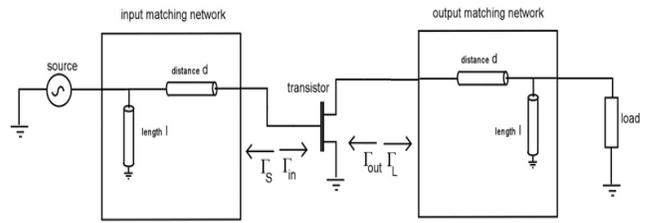


Fig. 6. Microwave amplifier design illustration.

This expression depends on transistor s-parameters:  $S_{11}$ ,  $S_{12}$ ,  $S_{21}$  and  $S_{22}$ . The reflection parameters are  $\Gamma_S$  and  $\Gamma_L$  for source and load respectively. Expressions for  $\Gamma_S$  and  $\Gamma_L$  and  $Z_1$ ,  $Z_2$ ,  $Z_3$  and  $Z_4$  were derived in [9] and are as follows:

$$\Gamma_S = \frac{Z_2 - 1}{Z_2 + 1}, \quad (7)$$

$$\Gamma_L = \frac{Z_4 - 1}{Z_4 + 1}, \quad (8)$$

$$Z_2 = \frac{\text{Re}(Z_1) + j(Z_0 \tan(d_1) + \text{Im}(Z_1))}{Z_0 - \text{Im}(Z_1) \tan(d_1) + j \text{Re}(Z_1) \tan(d_1)}, \quad (9)$$

$$Z_1 = \frac{jZ_0 \tan(l_1)}{1 + j \tan(l_1)}, \quad (10)$$

$$Z_4 = \frac{\text{Re}(Z_3) + j(Z_0 \tan(d_2) + \text{Im}(Z_3))}{Z_0 - \text{Im}(Z_3) \tan(d_2) + j \text{Re}(Z_3) \tan(d_2)}, \quad (11)$$

$$Z_3 = \frac{jZ_0 \tan(l_2)}{1 + j \tan(l_2)}, \quad (12)$$

Where,  $Z_0$  is the characteristic impedance of the transmission line. In our design, transistor FHX35X, manufactured by Fujitsu Cooperation was used. With the chosen s-parameters, the design was centered at frequency of 10 GHz. The design was optimized to get a gain of 16, which in decibels is 12 dB. The characteristics were gathered over 40 independent runs. Each run had 1000 function evaluations. The values for  $d_1$ ,  $d_2$ ,  $l_1$  and  $l_2$  were restricted to be range  $[0, 2\pi]$ . These values for  $d_1$ ,  $d_2$ ,  $l_1$  and  $l_2$  were obtained in terms of radians and five of them were tabulated in Table 5. In addition, a microwave simulator was used to obtain characteristics. Fig. 7 demonstrates results of Trial 1 for gain and reflection values.

The plot shows  $S_{11}$  and  $S_{22}$  which are the reflections in the ports 1 and 2, respectively and  $S_{21}$  which is the transmission from port 1 to 2. In this case, gain value is  $S_{21}$  value. In a design, especially at the design frequency  $S_{11}$  and  $S_{22}$  values should be minimized and if possible  $S_{11}$  and  $S_{22}$  values should be kept below 0 dB at all times to avoid oscillations. It is seen that this is roughly happening all throughout the band of observation from 7 GHz to 13 GHz.

TABLE V. A SAMPLE SOLUTION SET FOR DIFFERENT NUMBER OF TRIALS

	$d_1$	$d_2$	$l_1$	$l_2$
Trial 1	4.279915	4.826180	5.600057	5.307381
Trial 2	1.077591	3.406553	2.602834	4.395524
Trial 3	0.905713	1.858479	2.642391	2.087756
Trial 4	0.069484	4.511318	0.641231	2.218899
Trial 5	3.332904	3.381557	3.622599	1.301090

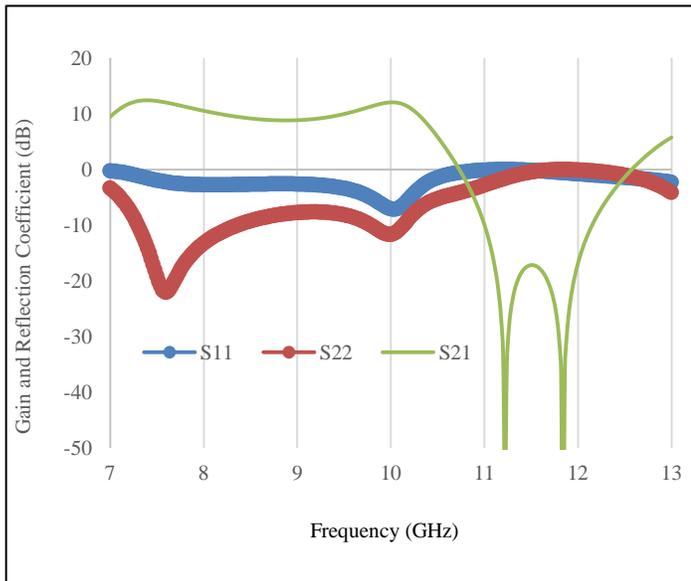


Fig. 7. Gain and reflection coefficient values of Trial 1 for the amplifier design.

At the design frequency  $S_{11}$  value is -8 dB and  $S_{22}$  value is -11 dB. The gain value is also at 12 dB. A sensitivity test is also performed to see if the values produced by the algorithm are sensitive to errors. 3% error is introduced to all of the length values, the overall gain dropped to 10.77 dB; however the simulation indicated that the transistor is stable. When a 5% error is introduced to all of the length values, the maximum gain dropped from 12 dB to 9.82 dB which is not a very desirable feature. However this kind of result is expected, since the gain performance in the model relies on correct transmission line lengths and introducing a 5% error in all lengths is actually a significant change in the design. These results overall indicate that COLA performed efficiently and it can be a good candidate of optimization algorithm for designing a stable amplifier.

#### IV. CONCLUSIONS

In this paper, a new optimization algorithm COLA is proposed to find the global optimum points of given problems by providing good solution quality with robust solutions to the random operations in the algorithm. Two core levels are applied simultaneously to reach these goals by providing a balance between exploration and exploitation characteristics.

The algorithm is tested with different characteristics of benchmark functions, compared with powerful evolutionary algorithms and then applied to two real microwave circuit

design problems. The results obtained for benchmark functions indicate that COLA provides better solution quality than the analyzed algorithms and its convergence speed is fairly good even for the first stages of optimization such as 10000 function evaluations. The results for microwave circuits obtained by COLA are verified by microwave simulators and it is seen that it produces accurate results. It is studied that the control parameter  $k$  is randomized in given interval produces better results. Therefore, COLA uses only one control parameter  $np$  which is not necessary to be well tuned for the problems studied in this paper. The number of population  $np$  is fixed to 100 for all the problems in this study. In other words, it can be considered that the algorithm is almost parameter free which can be used for any problem without tuning any control parameter. This advantage of the algorithm makes it very practical to be used for any real case problems.

According to these results, it can be concluded that the solution quality of the algorithm is better than the analyzed algorithms and also is quite robust even though there is no parameter to be tuned. The algorithm COLA can be suggested as a candidate for optimization problems including real case problems from different fields. As future work, a new research can be done to compare COLA with recent hybrid optimization algorithms.

#### REFERENCES

- [1] Y. Chen, S. Yang and Z. Nie, The Application of a modified differential evolution strategy to some array pattern synthesis problems, IEEE Trans. Antennas and Propagation, 56, pp. 1919-1927, 2008.
- [2] B. Choudhury, S. Manickam and R.M. Jha, Particle Swarm Optimization for Multiband Metamaterial Fractal Antenna, J. Optim. 2013, pp. 1-8, 2013.
- [3] E.D. Ulker, A. Haydar and K. Dimililer, Application of Hybrid Optimization Algorithm in the Synthesis of Linear Antenna Array, Math. Prob. Eng. 2014, pp. 1-7, 2014.
- [4] P. Wen-Tsao, A new Fruitfly Optimization Algorithm: Taking a Financial Distress model as an Example, Knowledge-Based Systems, 26, pp. 69-74, 2012.
- [5] E. Hadavandi, H. Shavandi, A. Ghanbari, Integration of Genetic Fuzzy Systems and Artificial Neural Networks for Stock Price Forecasting, Knowledge-based Systems. 23, pp. 800-808, 2010.
- [6] R. Köker, Genetic Algorithm Approach to a Neural-Network based inverse kinematics solution of robotic manipulators based on error minimization, Spec. Sect. on New Trends in Ambient Intelligence and Bio-inspired Syst. 222, pp. 528-543, 2013.
- [7] Z.W. Geem, Harmony Search Algorithm for Solving Sudoku, Knowledge-based Intelligent Information and Engineering Systems, Lect. Notes Comput. Sci. 4692, pp. 371-378, 2007.
- [8] E.D. Ulker and S. Ulker, Application of Particle Swarm Optimization to Microwave Tapered Microstrip Lines, Int. J. Comput. Sci. Eng (CSEIJ) 4, pp.59-64, 2014.
- [9] E. Deniz, S.Ulker, Clonal selection algorithm application to simple microwave matching network, Microw. Opt. Technol. Lett. 53, pp. 991-993, 2011.
- [10] S. Ulker, Broadband Microwave Amplifier Design Using Particle Swarm Optimization, J. Comput. 6, pp. 2272-2276, 2011.
- [11] A.A. Freitas, Data mining and knowledge discovery with evolutionary algorithms. Springer Science and Business Media, 2013.
- [12] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, C.A.C Coello, A survey of multiobjective evolutionary algorithms for data mining: Part I, IEEE Trans. on Evol. Comput.. 18, pp. 4-19, 2014.
- [13] Z.W. Geem, J.H. Kim and G.V. Loganathan, Harmony Search Optimization: Application to Pipe Network Design, J. Model. Optim. 22, pp. 60-68, 2002.

- [14] J. Kennedy, R. Eberhart, Particle Swarm Optimization, IEEE International Conference on Neural Networks. pp.1942-1948, 1995.
- [15] R. Storn, K. Price, Differential Evolution: A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, J. Glob. Optim. 11, pp.341-359, 1997.
- [16] L.N. De Castro, F.J. Von Zuben, Learning and optimization using the clonal selection principle, IEEE Trans. Evol. Comput. 6, pp. 239-251, 2002.
- [17] J.H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, Michigan, 1975.
- [18] Y. Shi and R. Eberhart, Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization, Proc. Congr. Evol. Comput., pp. 84-88, 2000.
- [19] A.S.D. Dymond, A.P. Engelbrecht and P.S. Heyns, The sensitivity of single objective optimization algorithm control parameter values under different computational constraint, IEEE Congress Evol. Comput. pp. 1412-1419, 2011.
- [20] S.K. Smit and A. Eiben, Comparing Parameter Tuning Methods for Evolutionary Algorithms, IEEE Congress on Evol. Comput., pp. 399-406, 2009.
- [21] M. Mahdavi, M. Fesanghary and E. Damangir, An Improved Harmony Search Algorithm for Solving Optimization Problems, Appl. Math. Comput. pp. 1567-1579, 2007.
- [22] Ghosh, S. Das, A. Chowdhury and R. Giri, An Improved Differential Evolution Algorithm with Fitness-based Adaptation of the Control Parameters, Infor. Sci., pp. 3749-3765, 2011.
- [23] M.M. Ali, P. Kaelo, Improved particle swarm optimization for global optimization, App. Math. Comput. 196, pp. 578-593, 2008.
- [24] Lihu, S. Holban, A Study on the Minimal Number of Particles for a Simplified Particle Swarm Optimization Algorithm, 6th IEEE International Conference on Applied Computational Intelligence and Informatics. pp. 299-303, 2011.
- [25] D.H. Wolpert and W.G. Macready, No Free Lunch Theorem for Optimization, IEEE Trans. Evol. Comput. 1, pp. 67-82, 1997.
- [26] Hoorfar, Evolutionary Programming in Electromagnetic Optimization: A Review, IEEE Trans. Antennas and Propagation. 55, pp. 523-537, 2007.
- [27] L. Zhang L, P.H. Aaen, T. Dhaene, A. Sahu and V. Devabhaktuni, An Introduction to Evolutionary Optimization for Microwave Engineering, Wiley Encyclopedia of Electrical and Electronics Engineering, 1999.
- [28] S.K. Goudos, K. Siakavara, T. Samaras, E.E. Vafiadis and J.N. Sahalos, Self Adaptive Differential Evolution Applied to Real-valued Antenna and Microwave Design Problems, IEEE Trans. Antennas and Propagation, 59, pp.1286-1298, 2011.
- [29] C. Lin, A. Qing and Q. Feng, Synthesis of Unequally Spaced Antenna Arrays by using Differential Evolution, IEEE Trans. Antennas and Propagation. 58, pp.2553-2561, 2010.
- [30] S.K. Goudos, C. Kalialakis and R. Mittra R, Evolutionary Algorithms Applied to Antennas and Propagation: A Review of State of the Art, International J. Antennas and Propagation, 1, pp.1-12, 2016.
- [31] X. Yao, Y. Liu and G. Lin G, Evolutionary Programming Made Faster, IEEE Trans. Evol. Comput. 3, pp. 82-102, 1999.
- [32] E.D. Ulker and A. Haydar, A Hybrid Algorithm based on Differential Evolution, Particle Swarm Optimization and Harmony Search Algorithms, IEEE International Conference on Computer Science and Information Systems, pp. 417-420, 2013.
- [33] J.W. Bandler, Optimization Methods for Computer Aided-Design, IEEE Trans. Microw. Theory Tech. 17, pp. 533-552, 1969.
- [34] J.W. Bandler, P.A. MacDonald, Optimization of microwave networks by razor search, IEEE Trans. Microw. Theory Tech. 17, pp.552-562, 1969.
- [35] V.A. Monaco and P. Tiberio, Computer-Aided Analysis of Microwave Circuits, IEEE Trans. Microw. Theory Tech. 22, pp. 249-263, 1974.
- [36] J. Carrol, K. Chang, Statistical computer-aided design for microwave circuits, IEEE Trans. Microw. Theory Tech. 44, pp.24-32, 1996.
- [37] M. Steer, Computer-Aided Design of RF and Microwave Circuits and Systems, IEEE Trans. Microw. Theory Tech. 50, pp. 996-1005, 2002.
- [38] S. Koziel, A. Bekasiewicz, Rapid Simulation-Driven Multiobjective Design Optimization of Decomposable Compact Microwave Passives, IEEE Trans. Microw. Theory Tech. 64, pp. 2454-2461, 2016.
- [39] S. Koziel, A. Bekasiewicz, Rapid Microwave Design Optimization in Frequency Domain using Adaptive Response Scaling, IEEE Trans. Microw. Theory Tech. 64, pp. 2749-2757, 2016.
- [40] C. Marcu, A.M. Niknejad, A 60 GHz High-Q Tapered Transmission Line Resonator in 90 nm CMOS, IEEE Microwave International Symposium Digest, pp. 775-778, 2006.
- [41] D.L.L De Villiers, P.W. Van Der Walt and P. Meyer, Design of conical transmission line power combiners using tapered line matching sections, IEEE Trans. Microw. Theory Tech. 56, pp. 1478-1484, 2008.
- [42] Sallem, B. Benhala, M. Kotti, M. Fakhfakh, A. Ahaitouf A and M. Loulou, Application of swarm intelligence techniques to the design of analog circuits: evaluation and comparison, Analog Integrated Circuits Signal Process. 75, pp. 499-516, 2008.
- [43] S. Ulker, Design of Low Noise Amplifiers Using Particle Swarm Optimization, Int. J. Artif. Intell. Appl. 3, pp. 99-106, 2012.
- [44] O.J. Ushie, M. Abbod, E.C. Ashigwuike and S. Lawan, Constrained Nonlinear Optimization of Unity Gain Operational Amplifier Filters Using PSO, GA and Nelder-Mead, Int. J. Intell. Control Syst. 20, pp. 26-34, 2015.
- [45] F. Gunes, S. Demirel, P. Mahouti, A simple and efficient honey bee mating optimization approach to performance characterization of a microwave transistor for the maximum power delivery and required noise, Internat. J. Numer. Model. Electron. Networks, Devices, Fields, 29, pp. 4-20, 2015.
- [46] F. Gunes, S. Demirel, S. Nesil, Design Optimization of LNAs and Reflect array Antennas Using the Full-Wave Simulation-Based Artificial Intelligence Models with the Novel Metaheuristic Algorithms, Simulation-Driven Model Optim. 153, pp. 261-298, 2016.

# A Proposed Adaptive Scheme for Arabic Part-of-Speech Tagging

Mohammad Fasha

Department of Computer Science  
King Abdulla II School for Information Technology  
The University of Jordan, Amman

**Abstract**—This paper presents an Arabic-compliant part-of-speech (POS) tagging scheme based on using atomic tag markers that are grouped together using brackets. This scheme promotes the speedy production of annotations while preserving the richness of resultant annotations. The proposed scheme is comprised of two main elements, a new tokenization approach and a custom tool that enables the semi-automatic implementation of this scheme. The proposed model can serve in many scenarios where the user is in a need for better Arabic support and more control over the Part-of-Speech tagging process. This scheme was used to annotate sample narratives and it demonstrated capability and adaptability while addressing the various distinguishing features of Arabic language including its unique declension system. It also sets new baselines that are prospect for further exploration by future efforts.

**Keywords**—Arabic natural language processing (ANLP); part-of-speech (POS) tagging; part-of-speech tokenization scheme; morpho-syntactic tagging; Arabic declension system

## I. INTRODUCTION

Part-of-Speech (POS) tagging is the process of classifying and labeling words in a sentence according to their grammatical categories, i.e., verbs, nouns, particles, ... etc. [1]. It is considered as an important step in many Natural Language Processing (NLP) implementations [2] as it deliver a layer of abstraction over the vast variances of the lexical, syntactic and semantic content of natural language. This generalization process renders that vast amount of knowledge into controllable artifacts that are valuable for many related implementations.

In contrast to other languages, Arabic has several distinguishing and challenging features, more importantly, its rich morphology and highly inflectional nature. A single Arabic word can bear more meaning than it's English counterparts [3]. Therefore and more often, information is either lost or misrepresented using the conventional Part-of-Speech tagging schemes. Moreover, there is a noticeable shortage in terms of standards related to Arabic Part-of-Speech tagging schemes, whether for the used tagsets or for the tokenization process [4], [5].

To assist in mitigating some of these challenges, we propose a new Part-of-Speech tagging scheme that can provide rich annotations while being simpler and less demanding than the detailed parsing of corpora, which is cumbersome and time consuming [6]. The scheme we are proposing is based on using tagsets of atomic tag markers that

can be aggregated and grouped together using brackets. Having such arrangements, users are provided with fundamental baselines that enable them to seamlessly commence with a rich morpho-syntactic annotation process for Arabic text.

The contributions of this work includes the definition of a declension system (نظام الاعراب) complaint morpho-syntactic tagging scheme that promotes simplicity, clarity and agility of the produced annotations as well as the tagging process itself. Further, to the best of our knowledge, this is one of the rare studies that surveys Arabic Part-of-Speech tagging schemes and discusses their pros and cons. This important subject needs further investigation due to the unique linguistic features of Arabic language, while most related work concentrates on establishing rule-based or statistical motivated Part-of-Speech taggers and morphology analyzers.

This paper is structured as follows. In Section II, we present a brief introduction about the distinguishing features of Arabic language. In Section III, we discuss the related previous work. Section IV presents some of the challenges that are related to the conventional Arabic Part-of-Speech tagging schemes. In Section V, the proposed tagging scheme is presented in more detail. Section VI presents the custom annotation tool. Section VII presents a sample narrative annotated using the proposed scheme and finally in Section VIII we present the conclusion and the suggested future work.

## II. ARABIC DISTINGUISHING FEATURES

Arabic is a Semitic language spoken by over 300 million speakers in 22 Arabic countries, it has a liturgical importance as it is the language of Quran, the Holy book for over 1.2 billion Muslims around the world [7].

In contrast to many other languages i.e. Indo-European languages, Arabic has many distinguishing features. These features are related to its rich morphology, highly inflectional nature, subject dropping, free words order, short vowels omission, large lexicon and vocabulary and many others [8], [9]. Accordingly, it is quite often challenging to identify the correct Part-of-Speech of a given word under a certain context.

The rich morphology of Arabic can be related to its template nature where new words are derived from root ones by applying a set of fixed patterns. In addition, Arabic has a concatenate nature where words (nouns and verbs) are inflected to indicate different senses. For example, Arabic

nouns can be inflected to indicate number (singular, dual, plural), gender (masculine, feminine), definiteness (definite, indefinite) and case (nominative, accusative, genitive) as well as possession. Similarly, Arabic verbs are inflected to indicate aspect (perfective, imperfective, imperative), voice (active, passive), tense (past, present, future), mood (indicative, subjunctive, jussive), subject (person, number, gender) as well as object clitics. In addition, Arabic words can be prefixed with functional morphemes (single particles or prepositions) to indicate various senses (causality, conjunctions, assertion, inquiry, association ... etc.).

To demonstrate the richness of Arabic language and the amount and diversity of information that can be conveyed in a single word, we consider the surface word (wa sa nokhberu hum, وسنخبرهم, and we shall inform them) as an example. This single word is comprised of the following constituents:

- The proclitic morpheme (wa, و, and) which indicates coordinating conjunction.
- The proclitic morpheme (sa, س, shall) which indicates a future event.
- The inflection particle (nun, ن) which indicates first voice plural speaker (us).
- The stem (khabara, خبر, tell) which is the verb itself.
- Finally, the enclitic morphemes (hum, هم, them) which is an attached pronoun that indicates a plural object of the verb.

In [10], the author provides a more detailed discussion about Arabic morphology and its distinguishing features. Nevertheless, annotating the previous sample word with a verb marker (VB) according to its grammatical category shall waste numerous information. Therefore, a viable Arabic part-of-speech tagging scheme has to possess the capacity to address Arabic distinguishing features and to accurately classify Arabic words without losing information or creating ambiguities. In order to be able to support the distinguishing features of Arabic language, the required part-of-speech tagging scheme has to be able to fully support Arabic's declension system (نظام الاعراب).

In the next section, we present a brief discussion about the related previous work and highlight their main challenges.

### III. RELATED WORK

A limited number of part-of-speech Taggers were presented for Arabic language [11]. Generally, these automated taggers can be classified under three main schemes: the statistical-based schemes, the rule-based schemes and the hybrid ones [2]. More importantly, reviewing the previous related work, we noticed an overlapping between part-of-speech tagging and morphology analyses process. For example, Stanford NLP toolkit uses the reduced Penn tagset, while others like the Buckwalter AraMorph incorporates the syntactic category of a given word within the generated morphology analyses results.

Nevertheless, in this work, we are interested in the part-of-speech annotating scheme and format that was implemented

by every one of these tools. We start our listing with an early effort that was presented by [12] who introduced a hybrid algorithm for Arabic part-of-speech tagging. That algorithm used a custom tagset comprised of (130) fixed morpho-syntactic markers that were defined based on Arabic grammar rules. Each marker identifies the grammatical category and the inflections of a given word. For example, a perfect verb in the second person masculine plural form is annotated using the (VPPI2M) marker and a singular masculine accusative definite adjective is annotated using the (NACSGMAD) marker.

An interesting tagging scheme was presented in Arabic Treebank (ATB) project [13]. That tagging scheme was based on the well-known rule-based Buckwalter Arabic Morphology Analyzer (BAMA) [14]. (BAMA) uses around (70) basic tag markers that can be combined together to form a larger number of composite tags. For example, in (BAMA), the (IV\_PASS) marker indicates imperfective passive verb, three types of information are aggregated together in that composite tag, i.e., imperfect, passive and verb. (BAMA) include tags for indicating person, voice, mood and aspect for verbs, and gender and number for their subjects. It also includes gender, number, case and state for different types of nominals [5].

Another important tagging scheme was introduced by the Prague Arabic Dependency Treebank (PADT) project that was presented in [15]. In that work, a multi-level annotation scheme for a selected corpus was implemented. The first level of annotation involved the morphology analyses of Arabic words. For that part, a morphology compliant tagset was used to construct a (15) slots structure covering the various morphological aspects of a given word i.e. gender, number, person, aspect ... etc. In PADT, a single character represented each morphology feature. A challenge in (PADT) tagging scheme was that the meaning of the same character might differ according to a specific internal structuring procedure. For example, the letter (P) on the second position is to be read as Passive Particle if it was preceded by an (N, Noun), and as a Perfect if it was preceded by a (V, Verb). This arrangement requires specialized skills and knowledge to be able to use and interpret (PADT) tagging scheme [16].

Similarly, CATiB project [17] presented an Arabic Treebanking scheme that was designed with the motivation of providing rich annotations while being simpler than other similar efforts i.e. ATB and PADT. The focus of CATiB was primarily on the speedy production of the manually annotated corpus while the inspiration was not to duplicate information that could be extracted or indicated by other means, i.e., by syntactic analysis. Consequently, CATiB introduced a succinct (POS) tagset comprised of (6) POS tags which are: NOM for nominals. PROP for proper nouns, VRB for active-voice verbs, and VRB-PASS for passive-voice verbs, PRT for particles and PNX for punctuations. Other markers were identified for the deeper level of syntactic-motivated annotations.

In [18], authors presented a functional based (POS) tagset where words are tokenized and (POS) tagged based on their grammatical functions rather than their morpho-syntactic structure. For example, the sentence (زمانها خلصت المسيرة), the march must have finished) is labeled as a modal (MD)

although the direct (POS) for the Arabic word (زمن, Time) is (NN, Noun).

A relatively recent effort was introduced by [11] who presented a systematic scheme for establishing Arabic compliant tagsets. In that work, a three level categorization of Arabic morpho-syntactic tagsets was defined. The first level was comprised of 7 tags, the inner level included 23 tags while lower level included 54 tags. Accordingly, the user of the system can use the depth of tagging that can better address his needs.

Finally, [2] and [5] presented interesting reviews on Arabic part-of-speech taggers and tagsets where the former concentrated on tagsets while the later presented a listing of the most prominent taggers along with a discussion about their challenges and limitations.

#### IV. CHALLENGES RELATED TO THE EXISTING ARABIC (POS) SCHEMES

The review process that was presented in the previous section revealed several challenges and limitations that are related to the existing tagging schemes. To further assess these schemes, we examined a number of the accessible taggers and morphology analyzer which included Stanford NLP toolkit [19], NLTK toolkit [20], AL-Khalil morphology analyzer [3], BAMA morphology analyzer [14] as well as MADAMIRA [21] and SAFAR platforms [22]. Table 1 below presents a listing of the results that were captured while examining these tools over a sample sentence. Analyzing the results from a

linguistic perspective, we concluded to the following list of observations:

a) There is no standardized or a community adopted (POS) scheme for Arabic language. Our examination revealed that different (POS) tagsets were used by different (POS) taggers; some of these tagsets were generalized while others were more detailed to better address Arabic distinguishing features. The observation was also noted by [4]. Similarly, the tokenization scheme of the tag markers is also different in each tool.

b) The accuracy of the examined (POS) taggers was questionable. For example, Stanford NLP produced numerous errors in the generated tagging such as the noun (كرته, his ball) which was annotated as (NNP) or a proper noun. Similarly, and for a different sample sentence, MADAMIRA identified the word (شعر, felt) as a noun (poetry) rather than a verb, also the verb (حضر, came) was identified as a verb inflected for third person singular masculine while the correct interpretation according to the context was a third person plural masculine. Likewise, the verb (خدعتكم, I deceived you) was identified as a verb inflected with a third person singular feminine subject while it was masculine according to the context. Moreover, the library failed to analyze some words e.g. (مسرعين, in a hurry) which were tagged as NO-ANALYSIS.

TABLE I. ANALYSIS RESULTS OF ARABIC (POS) TAGGERS AND MORPHOLOGY ANALYZERS FOR A SAMPLE SENTENCE

Arabic Sentence	وركل	كرته	فاتجيت	نحو الزجاج	فحطته
<b>English Translation</b>	and he hit/inflected as Singular Masculine Subject Inflection	his ball/ attached possession pronoun	So it went Singular Feminine Subject Inflection	towards the window/glass	so it broke it/inflected as Singular Feminine Subject + Singular Masculine Object
<b>Buckwalter Transliteration</b>	wrkl	krth	fitajahat	naHow AlzujAj	fHTmt
<b>Stanford NLP</b>	وركل/VBD	كرته/NN	فاتجيت/VBD	نحو/NN الزجاج/DTNN	فحطته/VBD
<b>Al-Khalil</b>	12 solutions, verbs and gerunds	17 solutions, verbs, nouns and gerunds	5 solutions, verbs	15 solutions for both	13 solutions, verbs and gerunds
<b>BAMA</b>	2 solutions including VERB_PERFECT and NOUN	6 solutions, NOUN	6 solutions including VERB_PERFECT With different subject inflections	4 solutions for naHowa and 5 solutions for AlzujAj	9 solutions VERB_PERFECT Different Subject and Object inflections
<b>SAFAR</b>	7 solutions, different subjects inflections	15 solutions, different subjects and objects inflections	4 solutions, different subject inflections	14 solutions for naHowa and 17 solutions for AlzujAj	12 solutions, different subjects and objects inflections
<b>MADAMIRA</b>	wa/CONJ+rakal /PV+a/PVSUFF _SUBJ:3MS	kur/NOUN+ap/NSUFF_FEM_SG +a/CASE_DEF_ACC+hu/POSS_PRON_3MS	fa/CONJ+[it~ajah/PV+at/PVSUFF_SUBJ:3FS	naHowa/PREP Al/DET+zuAj/NO UN+a/CASE_DEF _ACC	fa/CONJ+HaT~am/PV+at/PVSUFF_SUBJ:3FS+hu/PVSUFF_DO:3MS

c) Some of the examined tools were not suitable for automated (POS) processing as they generate all the possible

interpretations for a given word. This observation was noticed in BAMA and AL-Khalil morphology analyzers. Moreover, Al-Khalil does not employ any (POS) tokenization scheme,

rather, it generates all its results in plan Arabic text according to Arabic declension system, this features makes it unsuitable for any integration potentials.

d) Some of the investigated tools, i.e., SAFAR were a collection of other tools that were aggregated and compiled under a single platform. These tools were not stand-alone products by themselves and they did not introduce any original add-ins in terms of the Part-of-Speech tagging functionalities.

e) In many situations, words were tagged with an overly generalized version of tag markers where useful information was lost. This can be witnessed in Stanford (POS) tagger that employs the English Penn Treebank tagset for annotating Arabic words. That tagset lacks Arabic morphology features. Similarly, useful information is wasted as the examined tools are not fully compliant with Arabic declension system (نظام الاعراب). For example, gender information proper nouns, some adjectives and nouns were not included. Likewise, functional characters have an important role in Arabic language, yet the functional specificity for some Arabic particles was neglected such as the conditional (إذا, if).

f) The number of basic tag markers and the number of their possible combinations can reach large amounts that can complicate the tagging process. In [23], the authors identified over (2000) markers for Arabic while the combination of these markers can theoretically reach (33000) different tag combination [24].

g) Overlapping and duplications can be witnessed in some of the existing tagging schemes. Such overlapping can complicate string-based matching over the Part-of-Speech strings. For example, in the Penn Treebank tag markers presented below, we notice that the concept of feminine gender is represented using the single character (F), yet this same character appears as part of the (PVSUFF) marker in the same string.

```
VERB_PASSIVE+PVSUFF_SUBJ:3FS  
VERB_PASSIVE+PVSUFF_SUBJ:3FS  
VERB_PASSIVE+PVSUFF_SUBJ:3MP
```

The same remark can be observed for the singular number marker (S) and the plural (P) as they overlap with characters in the word (PASSIVE).

h) In addition, we can observe that the same concept might be represented using different markers within the same scheme. For example, the tags markers presented below demonstrate how the singular number was represented using (SG) in the first sample and using an (S) in the second.

```
ADJ+NSUFF_FEM_SG  
IV3FS+VERB_IMPERFECT
```

The same is true for the feminine gender markers i.e. the (F) and (FEM). Such inconsistency can create confusion during the use of the markers and weakens the scheme's standardization potentials.

i) For generating morpho-syntactic tagging, it is required that we perform a full tokenization for sentences prior to the tagging process. Such requirement might be cumbersome and time consuming and it should be useful if we can develop a simpler scheme that can replace the explicit tokenization with an implicit one as the missing information can be recovered using algorithmic measures.

j) Considering the previously discussed challenges and limitations, manual intervention is often required to fine-tune the automatically generated annotations. This intervention is required to verify and/or extend the generated annotations and to validate their accuracy and adequacy for further stages of processing, which brings us to another challenge in this respect and that is the scarcity of available and accessible annotation tools that can enable and facilitate such functions of manual intervention.

In the next section, we present our proposed (POS) tagging scheme which might assist in addressing some of the aforementioned challenges as well as setting new perspectives for further exploration in future.

## V. THE PROPOSED TAGGING SCHEME

In this section, we present the proposed part-of-speech tagging scheme including its objectives, design principles, the initial tagset, the tagging process as well as the custom tool that was prepared to enable this scheme.

### A. Objectives and Design Principles of the Proposed Scheme

The main objective of the proposed tagging scheme was to provide users with initial baselines that enable them to implement a rich morpho-syntactic declension-system compliant annotation for Arabic words in a clear, simple and agile manner. Using this scheme, users can experiment with different tag markers that are more compliant with Arabic language, and would be able to examine their influence on different Natural Language Processing (NLP) applications e.g. Information Extraction, Text Translation, Text Summarization ... etc.

The clarity, simplicity and agility of the proposed scheme were established by allowing users to commence with the annotation process without the need for the explicit tokenization of words. Rather, the tokenization is achieved using different brackets as shall be presented later. The inspiration for this arrangement was motivated by the tagging scheme that was presented in [17]. In that work, the speedy production of annotations was enabled by eliminating the annotation of information that could be extracted by other means. For instance, case markers for nominals could be identified from syntax, therefore, the Part-of-Speech annotation scheme presented in [17] did not include such markers in its tagset.

The morpho-syntactic richness of the annotations is enabled by the support of different categories of tag markers that are compliant with Arabic declension system, this includes lexical categories of words; morphology related markers, functional markers as well as declension system specific ones.

To enable the aforementioned objectives, the proposed scheme was based on the following design principles:

a) All the defined tag markers in the scheme were standalone and atomic. Each marker is self-explaining and self-contained and clearly defines a single concept e.g. gender, number, case, mood...etc. This design principle promotes the clarity of markers and ensures that no duplication or overlapping between markers can occur. For example, if a marker indicates a certain concept e.g. FEM for feminine gender, this same marker will be used for all words categories that might be inflected to indicate gender i.e. nouns, verbs, adjectives, pronouns, relative pronouns...etc. No other marker will be used for the same concept regardless of the word category. Therefore, the challenges that were stated in items g) and h) of section IV cannot occur.

b) Composite markers are established as aggregates of the basic and atomic ones. For instance, a plural noun is represented using the (NN) marker and the (PLR) marker, not with a single marker i.e. (NNS), for both concepts. This design principle preserves clarity and allows extensibility using clear composition of markers; it also facilitates string-based matching operations that can be implemented over part-of-speech annotations.

### B. Initial (POS) Tagset

The definition of a coherent Arabic-compliant tagset is out of the scope of our current work. In [11] and [25], the authors provided interesting guidelines that can assist in defining an Arabic-compliant tagset in a more systematic manner.

Nevertheless, for assessing our proposed model, we established an initial tagset to demonstrate the capability of the scheme and the diversity of markers that it can seamlessly support. This initial tagset (presented in Appendix A) classify the tag markers according to the following categories:

- Lexical markers:

This category includes the basic grammatical classification of words according to Arabic language rules. This includes the classification of nominals, verbs and particles, the three main Arabic word types along with their direct subsets.

- Morphology related markers:

This includes the markers that identify affixations and inflections related to nouns and verbs.

- Functional markers:

Functional markers include the tags that indicate the functional role of a given lexical entity. This includes senses of causality, modality, time and space relations, assertion, confirmation, negation, sequencing and conjunction coordination as well as others.

- Arabic declension system:

This category includes markers that are related to case definitions for Arabic nouns and mood definitions for Arabic verbs, as well as other features that signals specific insights

that are related to Arabic language e.g. (Kana and its sisters, كان وأخواتها).

### C. The Proposed Tokenization Scheme

A main objective of the proposed model was to better support Arabic declension system i.e. (نظام الاعراب) where the user is able to employ adequate combination of markers that can better satisfy his needs and his language proficiency.

Having an extended and diverse tag set, it was important to define an adaptive, dynamic and flexible tokenization scheme that can utilize these diverse markers in a simple, clear and agile manner.

Two types of brackets were employed to establish the proposed tokenization scheme, the round brackets or parenthesis “( )” and the braces or the curly brackets “{ }”. Using these brackets, different levels of grouping and hierarchies could be established to annotate different word categories. The parentheses are used to establish word level groupings while the curly brackets are used to create token level annotations. This arrangement combines concepts from conventional Part-of-Speech tagging, morphology analysis as well as syntactic tree parsing as a single Arabic word can encompass a multi-token paragraph according to its morphology.

To demonstrate the proposed bracketing scheme, we consider the sample surface word that was presented in Section 2 (wa sa nokhberu hum, وسنخبرهم, and we shall inform them). Using the proposed scheme, this single word is annotated as following:

- {RP+WA+CC}: The proclitic morpheme (wa, و, and) which indicates coordinating conjunction particle.
- {RP+SA+FTR}: The proclitic morpheme (sa, س, shall) which indicates a future event particle.
- {PLRL+stV}: The inflection particle (nun, ن) which indicates first voice plural speaker (us).
- {VB}: The stem (khabara, خبر, tell) which is the verb itself.
- {PRN+SFX\_OBJ+PLRL+MSC}: The enclitic morphemes (hum, هم, them) which is an attached pronoun that indicates a plural masculine object.

While the composite tag for this word is defined as following:

{(RP+WA+CC){RP+SA+FTR}{VB+PLRL+stV}{PRN+SFX+OBJ+PLRL+MSC)}

### D. Advantages of the Proposed Scheme

To demonstrate the advantages of the proposed tagging scheme over other available schemes, we performed several examinations for annotation sample words using Stanford (POS) tagger, MADAMIRA morphology analyzer and the proposed scheme.

TABLE II. COMPARING THE PROPOSED SCHEME AGAINST OTHER SCHEMES

	Annotation Scenario	Sentence Sample	Stanford (POS) Tagger	MADAMIRA Morphology Analyzer Scheme	Proposed Scheme
1	Composite words	شاهدته She saw him	VBD	{bw:\$Ahad/PV+tu/PVSUFF_SUBJ:1S+hu/PV SUFF_DO:3MS}	{{VBD+SNG+FEM+rdV} {SFX+OBJ+PRN+SNG+MSC}}
2	Kana and its sister كان وأخواتها	كانت السماء ماطرة The sky was raining	VBD	{bw:kAn/PV+at/PVSUFF_SUBJ:3FS}	{{VBD+KANA+SNG+FEM+rdV}}
3	ENNA and its sister إنَّ وأخواتها	إنها تمطر بغزارة It is raining heavily	VBP	{bw:<in~a/FUNC_WORD+hA/PRON_3FS}	{{IN+ENNA}{PRN+SNG+FEM+rdV}}
4	إنَّ: gloss: if/whether	إن تدرس تنجح If you study you succeed	IN	{bw:<in/FUNC_WORD}	{{IN+CND}}
5	Active Participle اسم الفاعل	هي ذاهبه I am going	JJ	{bw:*Ahib/ADJ+ap/NSUFF_FEM_SG}	{AP+SNG+FEM}
6	Passive Participle اسم المفعول	هو مظلوم He is oppressed	NNP	{bw:maZoluwm/ADJ}	{PP+SNG+MSC}
7	Relative Pronouns الأسماء الموصولة	الطفل الذي يبكي The baby that is crying	WP	{bw:Ai~a*iy/REL_PRON}	{{RPRN+SNG+MSC}}
8	Demonstrative Pronouns ضمائر الاشارة	هذا كتابي This is my book	DT	{bw:h*A/DEM_PRON_MS}	{{PRN+SNG+MSC+NR}}
		ذلك كتابي That is my book	DT	{bw:h*A/DEM_PRON_MS}	{{PRN+SNG+MSC+FR}}
9	Pronouns	هي تلعب بالكرة She is playing with the ball	PRN	{bw:hiya/PRON_3FS}	{{PRN+SNG+FEM+rdV}}
		انت تلعبين بالكرة You are playing with the ball	PRN	{bw:hiya/PRON_2FS}	{{PRN+SNG+FEM+ndV}}
10	Distinguish Prepositions	ذهنا الى المدرسة We went to school	IN	{bw:<ilaY/PREP}	{RP+ELA}
		جلسنا على المقعد We sat on the chair	IN	{bw:EalaY/PREP}	{RP+ALA}
11	Gender and Number Markers for Nouns	شاهدت السماء	DTNN	{bw:Ai/DET+samA'/NOUN+u/CASE_DEF_NOM}	{DT+NN+SNG+FEM+CSN}
12	Adverbs of manner	ركض الولد سريعا The boy ran quickly	JJ	{bw:sariyE/ADV+AF/CASE_INDEF_ACC}	{{RB+MNR}}
13	Interrogative Nouns	كم How much	WRB	{bw:kam/INTERROG_PART}	{{WP+QTY}}
		متى When	WRB	{bw:mataY/INTERROG_PART}	{{WP+TIM}}
		كيف How	WRB	{bw:kayofa/INTERROG_PART}	{{WP+MNR}}
		أين Where	WRB	{bw:>ayona/INTERROG_PART}	{{WP+LOC}}
		لمن Whose	WP\$	{bw:li/PREP+man/INTERROG_PART}	{{WP+POSS}}

Stanford tagger produces basic syntactic based tag markers for Arabic, while MADAMIRA provides a more extended version of markers that includes syntactic word classifications as well as the morphology analysis related ones. Table 2 below presents a listing of the gathered results.

As demonstrated in the table, the proposed scheme can deliver the same set of capabilities that are provided by the other models only it has the following additional advantages:

- The format of the proposed tagging scheme falls between the briefed Stanford format and the extended format of MADAMIRA. Nevertheless, the proposed scheme provides all the information that is delivered by those two schemes in a simplified manner that includes

the syntactic word type classification as well as the morphology related ones.

- The use of brackets eliminates and substitutes the explicit tokenization of composite words. As demonstrated in the first sample, that composite word is comprised of two parts, the perfect verb and the attached pronoun. Curly brackets surround each of these two word parts and parenthesis surrounds the whole string. While in the other schemes, the aggregation is achieved by attaching characters together without any separators or using separators such as the underscore marker “\_”, the plus sign “+”, the colons “:”, as well as other approaches e.g. PV+PVSUFF\_SUBJ:3MS.

- The proposed scheme does not use single-character markers as they can create ambiguities and overlaps. Rather, multi-character atomic tag markers are used to establish a self-explaining set of annotations.
- Also, unlike [12], [14], [16], [17], no aggregate markers are used in the proposed scheme, rather, all aggregations are established using the plus sign “+” character which is inserted between the atomic markers. Reference [26] presents an interesting listing for tokenization alternatives that are used by a number of different schemes. While in the previous efforts, different approaches were employed to achieve the same objective where a combination of the tokenization process, part-of-speech tagging and morphology analysis are all combined causing overlapping and ambiguity.
- Finally, the proposed scheme enables the introduction of different categories and types of tagsets and tag markers, whether they are related to basic syntactic and grammatical markers, functional markers, morphology related and semantic markers or any other type that might be needed for a specific objective. The expendability while maintaining clarity and simplicity is a powerful feature that maximizes the benefits of the proposed scheme. This can be observed in many samples in the previous table where explicit markers are used for different Arabic linguistic features e.g. active participle, passive participle, KANA and its sisters ... etc. Using such explicit markers can facilitate later efforts such as information extraction since these explicit markers can signal the existence of specific types of information.

## VI. THE CUSTOM ANNOTATION TOOL

To enable the proposed scheme, a Java based custom tool was prepared. We refer to this custom tool as the Bracket

Based Arabic Annotation (B2A2) tool as it employs brackets to establish morpho-syntactic compliant part-of-speech annotations for Arabic language.

Fig. 1 below presents a screenshot of the (B2A2) tool that demonstrates the tagging hierarchies (left) and the available tag markers (right). To commence with a new tagging process, a newline-terminated text file is uploaded into this tool where it will be initially bootstrap annotated using Stanford (POS) tagger. Later, the user uses the custom tool to review the initial annotations and modify/extend them accordingly. As demonstrated in the figure, the tool is delivered with an initial tagset where markers are classified into a number of categories e.g. base or lexical tags, functional tags, Arabic specific ... etc. These tags and tagsets can be easily modified and configured by the user who can introduce new tagsets or tag markers or modify the existing ones according to his needs. The modification for these markers can be introduced into the designated (tag\_def) database table i.e. SQL Server database. The structure of the tag definition table is described in Table 3 next. The user can modify the markers themselves as well as their categorization. The custom tool dynamically incorporates any modifications on the markers or their categories during its initialization process. This dynamicity in marker definition as well as their utilization by the user allows users to use different formats for annotating the same word.

The variance in annotations is related to the defined tag markers, the required depth of coverage and richness of the annotation process as well as the user’s linguistic proficiency.

Fig. 2 below demonstrates a screenshot of the (B2A2) tool, which clarifies how different annotations can be implemented for the same word according to the user’s defined annotation guidelines.



Fig. 1. The custom annotation tool.

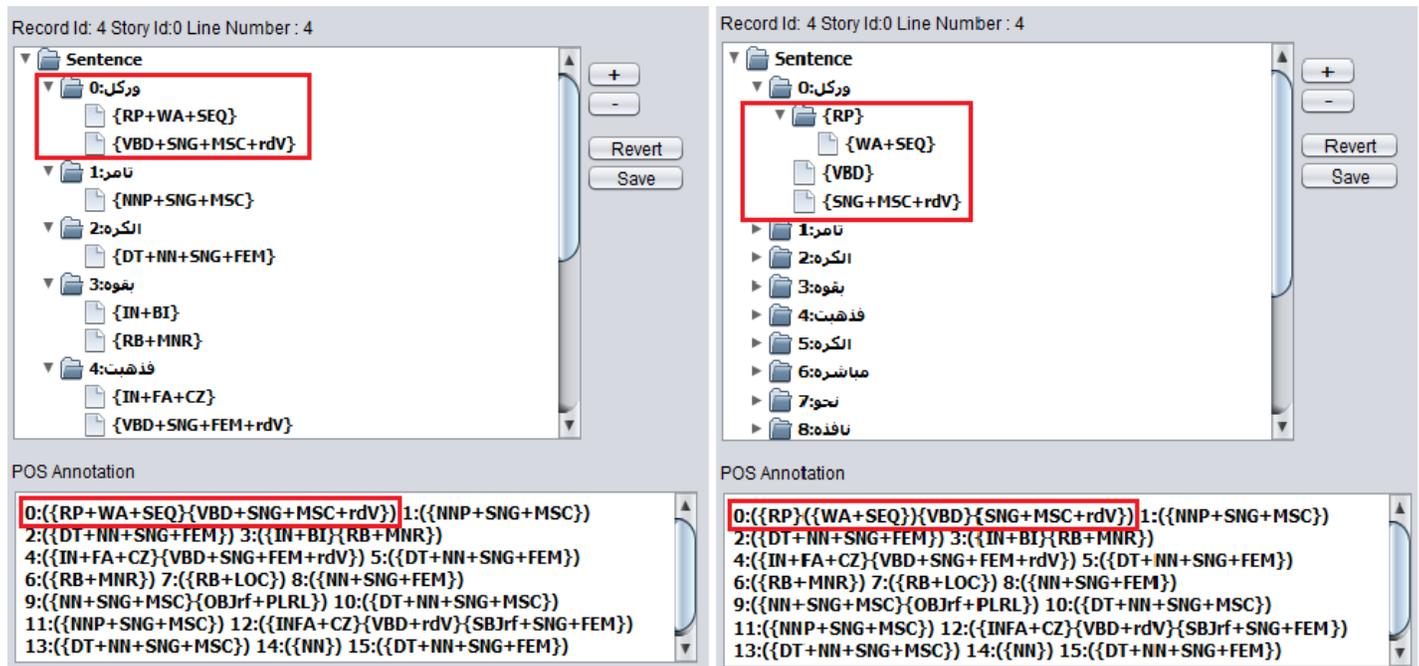


Fig. 2. Words and their constituents can be annotated different according to the user's definitions and requirements.

TABLE III. TAGS DEFINITION TABLE

Column Id	Explanation	Example
id	Unique identifier for the tag	
tag_order	The order or the precedence of the tag within a composite annotation	10
tag	The tag marker – acronym	RB
short_name	The short name for the marker	Adverb
english_description	English description for the marker	An adverb is a word that changes or qualifies the meaning of a verb.
arabic_description	Arabic description for the marker	حال أو ظرف مكان أو زمان
tag_category	The category where this marker belongs i.e. appears in the custom tool	Lexical Markers

## VII. ANNOTATING A SAMPLE NARRATIVE

To assess the proposed scheme in action, we used the (B2A2) tool to annotate a sample narrative comprised of a few sentences. As discussed in the previous section, (B2A2) scheme provides different alternatives for annotating text in terms of the tag markers that can be used as well as their arrangement and grouping using brackets. In this respect, the following guidelines were defined and enforced during the annotation process:

- Verbs annotations were extended with number, gender and person markers.
- Verb prefixes were distinguished using custom particles tagging i.e. (واو WAW, فاء FA, سين SA...etc.).
- Nominals tagging was extended using number and gender markers.

- Noun and adjectives prefixes were distinguished using custom particles tagging i.e. (واو WAW, كاف KA, باء BI ... etc.).
- A more precise tag set was used to annotate propositions i.e. (في FEE, من MIN, الى ELA ... etc.).
- Prefix particles, propositions and affixes were separated and grouped using dedicated brackets.
- Arabic (KANA, كان واخواتها, Was) was annotated using a custom tag (VBD+KANA) so that it can be better identified for future purposes.
- Occasionally during the annotation process, Arabic declension system was used in order to determine the correct grammatical analyses of some words and phrases so that ambiguous interpretations are resolved.

TABLE IV. A SAMPLE STORY ANNOTATED USING THE CUSTOM SCHEME

Line #	Text and (POS) Annotation	
1	اهدت:0 ليلى:1 شقيقها:2 تامر:3 كره:4 جديده:5	
	0:({VBD+SNG+FEM+rdV}) 1:({NNP+SNG+FEM})	
	2:({NN+SNG+MSC}) {POSS+SNG+FEM})	
	3:({NNP+SNG+ MSC})	
	3:({NN+SNG+ FEM})	
2	وكان:0 تامر:1 سعيدا:2 جدا:3 بها:4	
	0:({WA+RP+CC}) {KANA+VBD+SNG+MSC+rdV})	
	1:({NNP+SNG+MSC})	
	2:({JJ+SNG+MSC})	
3	ركض:0 تامر:1 نحو:2 الحديقه:3 ليلعب:4 بكرته:5 الجديده:6	
	0:({VBD+SNG+MSC+rdV}) 1:({NNP+SNG+MSC})	
	2:({RB+LOC})	
	3:({DT+NN+SNG+FEM})	
	4:({IN+LI+CZ}) {VBP+SNG+MSC+rdV})	
	5:({IN+BI}) {NN+SNG+FEM}) {POSS+SNG+MSC})	
4	وركل:0 كرهته:1 بقوه:2 فذهبت:3 الكره:4 مباشره:5 نحو:6 نافذه:7 جارهم:8 السيد:9 عادل:10 فتحطم:11 زجاج:12 النافذه:13	
	0:({RP+WA+CC}) {VBD+SNG+MSC+rdV})	
	1:({DT+NN+SNG+FEM}) {POSS+SNG+MSC})	
	2:({IN+BI}) {RB+MNR})	
	3:({IN+FA+CZ}) {VBD+SNG+FEM+rdV})	
	4:({DT+NN+SNG+FEM}) 5:({RB+MNR})	
	6:({RB+LOC}) 7:({NN+SNG+FEM})	
	8:({NN+SNG+MSC}) {POSS+PLRL+MSC})	
	9:({DT+NN+SNG+MSC}) 10:({NNP+SNG+MSC})	
	11:({IN+FA+CZ}) {VBD+SNG+FEM+rdV})	
	12:({DT+NN+SNG+FEM}) 13:({NN})	
	14:({DT+NN+SNG+FEM})	
	5	راى:0 تامر:1 السيد:2 عادل:3 قائدا:4 الى:5 منزله:6
		0:({VBD+SNG+MSC+rdV})
1:({NNP+SNG+MSC+rdV}) 2:({DT+NN+SNG+MSC})		
3:({NNP+SNG+MSC}) 4:({AP+SNG+MSC})		
5:({IN+ELA})		
6:({NN+SNG+MSC}) {POSS+SNG+MSC})		
6	فخاف:0 واختبا:1 خلف:2 شجره:3	
	0:({IN+FA+CZ}) {VBD+SNG+MSC+rdV})	
	1:({RP+WA+CC}) {VBD+SNG+MSC+rdV})	
7	راى:0 السيد:1 عادل:2 تامر:3 مختبئا:4 وابتسم:5 وقال:6 لا:7 تختبئ:8 يا:9 صغيري:10	
	0:({VBD+SNG+MSC+rdV}) 1:({DT+NN+SNG+MSC})	
	2:({NNP+SNG+MSC}) 3:({NNP+SNG+MSC})	
	4:({AP+SNG+MSC})	
	5:({RP+WA+CC}) {VBD+SNG+MSC+rdV})	
	6:({RP+WA+CC}) {VBD+SNG+MSC+rdV})	
	7:({RP+DMND+NEG}) 8:({VBP+SNG+MSC+ndV})	
	9:({RP+YAA})	
	10:({JJ+SNG+MSC}) {POSS+SNG+MSC})	
	8	اخبرني:0 الحقيقه:1 ولا:2 تخف:3
0:({VMP+SNG+MSC}) {PRN+SNG+MSC})		
1:({DT+NN+SNG+FEM})		
2:({RP+WA+CC}) {RP+DMND+NEG})		
9	خرج:0 تامر:1 من:2 مخياه:3 واخبر:4 السيد:5 عادل:6 الحقيقه:7 وهي:8 ان:9 كرهته:10 تسببت:11 في:12 تحطيم:13 زجاج:14 النافذه:15	
	0:({VBD+SNG+MSC+rdV}) 1:({NNP+SNG+MSC})	

10	وتاسف:0 ووعده:1 الا:2 يكرر:3 هذا:4 الفعل:5 مره:6 اخرى:7	
	0:({RP+WA+CC}) {VBD+SNG+MSC})	
	1:({NNP+SNG+MSC})	
	2:({RP+WA+CC}) {VBD+SNG+MSC+rdV}) {SFX_OBJ+SNG+MSC}) 3:({RP+CNF}) {RP+NEG})	
	4:({VBP+SNG+MSC+rdV}) 5:({RPRN+SNG+MSC})	
	6:({NN+SNG+FEM}) 7:({NN})	
	11	تبسم:0 السيد:1 عادل:2 وقال:3 يا:4 بنى:5 لقد:6 احسنت:7 صنعا:8 بقولك:9 الحقيقه:10 وتاسفك:11
		0:({VBP+SNG+MSC+rdV}) 1:({DT+NN+SNG+MSC})
		2:({NNP+SNG+MSC})
		3:({RP+WA+CC}) {VBD+SNG+MSC+rdV})
		4:({RP+YAA})
		5:({NN+SNG+MSC}) {POSS+SNG+MSC})
		6:({RP+CNFRM}) 7:({VBD+SNG+MSC+ndV})
8:({VN}) 9:({IN+BI}) {VBG+SNG+MSC+ndV})		
10:({DT+NN+SNG+FEM})		
11:({RP+WA+CC}) {VN+SNG+MSC}) {POSS+SNG+MSC})		
12		والان:0 خذ:1 الكره:2 وارجو:3 ان:4 لا:5 يتكرر:6 مثل:7 هذا:8 الفعل:9 مجدا:10
	0:({RP+WA+CC}) {RB+TIM})	
	1:({VMP+SNG+MSC+ndV})	
	2:({DT+NN+SNG+FEM})	
	3:({RP+WA+CC}) {VBP+SNG+stV}) 4:({IN+CNFRM})	
	5:({RP+NEG}) 6:({VBP+SNG+MSC}) 7:({NN})	
	8:({DT+SNG+MSC+NR}) 9:({DT+VN+SNG+MSC})	
	10:({JJ})	

The result of annotation the sample narrative is presented in Table 4 above. For example, the noun (شقيقها, shaqequha, her brother) was annotated using two segments, the first one belongs to the noun part along with its inflection, and the second is related to the attached pronoun suffix. The first part is annotated using {NN+SNG+MSC} tag group while the second part is annotated using the {POSS+SNG+FEM} tag group. As presented, each part is identified using a pair of curly brackets while the whole word (multi-token word) is grouped using a pair of parenthesis.

The annotation process demonstrated the efficiency of the proposed tagging scheme in representing the required syntactic and morphological information in simple yet rich manner. Further, the (B2A2) tool provided an enabling framework that accelerated the process of revising the automatically generated Part-of-Speech tagging and facilitated extending it using the proposed tagging scheme.

The proposed framework (the proposed Part-of-Speech tagging scheme and the B2A2 tool) can serve in numerous scenarios where the user is in a need to annotate a given corpus using a rich morpho-syntactic annotation while that labeled corpus can be used later for different Natural Language Processing (NLP) implementations e.g. Information Extraction from text.

## VIII. CONCLUSION AND FUTURE WORK

This paper presented a proposed scheme for Arabic-compliant part-of-speech tagging (POST).

Acknowledging the complexity and the richness of Arabic language, along with the shortages in the related standardizations, efforts and resources, the proposed (POST) scheme presented new perspectives that might assist in enhancing Arabic-based part-of-speech tagging process as well as opening doors for new perspectives and insights to regular such efforts.

The theme of the proposed model is relatively simple and straightforward yet powerful and capable in representing different types of information specific to Arabic language and its declension system. This scheme is based on: 1) using well-defined atomic part-of-speech markers; and 2) grouping these markers using two types of brackets, the curly brackets for sub-word level and the parenthesis for the word level of groupings.

A custom tool that is bootstrapped using Stanford (POS) tagger enabled the initial version of the proposed (POST) scheme. This tool is freely available online and it can assist users to commence with a rich Part-of-Speech tagging process in a controllable and seamless manner.

The next work we intend to implement is to examine the benefits that can be achieved by using the proposed scheme in information extraction implementations. In addition, we intend to investigate the bootstrapping of the enabling tool using a morphology aware part-of-speech tagging library, e.g., MADAMIRA.

### REFERENCES

- [1] S. Alqrainy, "A Morphological - Syntactical Analysis," 2008.
- [2] R. A. Abumalloh, H. M. Al-Sarhan, O. Bin Ibrahim, and W. Abu-Ulbeh, "Arabic Part-of-Speech Tagging," *J. Soft Comput. Decis. Support Syst.*, vol. 3, no. 2, pp. 45–52, 2016.
- [3] A. Boudlal, A. Lakhouaja, A. Mazroui, A. Meziane, M. Ould Abdallahi Ould Bebah, and M. Shoul, "Alkhalil Morpho SYS1: A Morphosyntactic Analysis System for Arabic Texts," *Int. Arab Conf. Inf. Technol.*, pp. 1–6, 2010.
- [4] S. Alqrainy, A. Ayes, and H. Almuaidi, "Automated Tagging System And Tagset Design For Arabic Text," vol. 1, no. 2, pp. 55–62, 2010.
- [5] A. M. S. Alosaimy and E. S. Atwell, "A review of morphosyntactic analysers and tag-sets for Arabic corpus linguistics," *Corpus Linguist.* 2015, pp. 16–19, 2015.
- [6] N. Habash, R. Faraj, and R. Roth, "Syntactic Annotation in the Columbia Arabic Treebank," *Proc. MEDAR Int. Conf. Arab. Lang. Resour. Tools*, Cairo, Egypt, pp. 125–132, 2009.
- [7] F. Al-shargi and O. Rambow, "DIWAN : A Dialectal Word Annotation Tool for Arabic," pp. 49–58, 2015.
- [8] M. Attia, "An Ambiguity-Controlled Morphological Analyzer for Modern Standard Arabic Modelling Finite State Networks," *Challenges Arab. NLP/MT Conf. ....*, pp. 48–67, 2006.
- [9] M. Sawalha, E. Atwell, and M. a. M. Abushariah, "SALMA Standard Arabic Language Morphological Analysis," *Proceedings ICCSPA International Conference on Communications, Signal Processing, and their Applications*. Sharjah, UAE, 2013. [Online]. Available: <http://www.comp.leeds.ac.uk/eric/sawalha13iccsa.pdf>. [Accessed: 23-Dec-2015].
- [10] N. Y. Habash, "Introduction to Arabic Natural Language Processing," *Synth. Lect. Hum. Lang. Technol.*, vol. 3, no. 1, pp. 1–187, 2010.
- [11] Y. O. M. Elhadj, A. Abdelali, R. Bouziane, and A. H. Ammar, "Revisiting Arabic Part of Speech Tagsets," *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA*, vol. 2014, pp. 793–802, 2014.
- [12] S. Khoja, "APT : Arabic Part-Of-speech Tagger," *Proc. Student Work. NAACL*, pp. 20–25, 2001.
- [13] M. Maamouri and A. Bies, "Developing an Arabic treebank: methods, guidelines, procedures, and tools," *Proc. Work. Comput. Approaches to Arab. Script-based Lang.*, pp. 2–9, 2004.
- [14] T. Buckwalter, "Arabic Morphological Analyser Version 1.0," *Linguist. Data Consort. numéro LDC2002L49*, 2002.
- [15] J. Hajič, O. Smrž, P. Zemánek, J. Šnidauf, and E. Beška, "Prague Arabic Dependency Treebank Development in Data and Tools," *Proc. {{NEMLAR}} Int. Conf. {{A}}rabic Lang. Resour. Tools*, pp. 110–117, 2004.
- [16] O. Smrž, J. Šnidauf, and P. Zemánek, "Prague Dependency Treebank for Arabic: Multi-Level Annotation for Arabic Corpus," *Adv. Math. (N. Y.)*, vol. 217, no. 6, pp. 2401–2442, 2008.
- [17] N. Habash and R. M. Roth, "CATiB: the Columbia Arabic Treebank," '09 *Proc. ACL-IJCNLP 2009 Conf.*, no. August, pp. 221–224, 2009.
- [18] R. Al-sabbagh and R. Girju, "Supervised POS Tagger for Written Arabic Social Networking Corpora," *Proc. KONVENS 2012 (Main track oral Present. Vienna, Sept. 19, vol. 2012, pp. 39–52, 2012.*
- [19] C. D. Manning, J. Bauer, J. Finkel, S. J. Bethard, M. Surdeanu, and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," *Proc. 52nd Annu. Meet. Assoc. Comput. Linguist. Syst. Demonstr.*, pp. 55–60, 2014.
- [20] S. Bird, S. Bird, and E. Loper, "NLTK : The natural language toolkit NLTK : The Natural Language Toolkit," *Proc. ACL-02 Work. Eff. tools Methodol. Teach. Nat. Lang. Process. Comput. Linguist.* 1, no. March, pp. 63–70, 2016.
- [21] A. Pasha, M. Al-badrashiny, M. Diab, A. El Kholly, R. Eskander, N. Habash, M. Pooleery, O. Rambow, and R. M. Roth, "MADAMIRA : A Fast , Comprehensive Tool for Morphological Analysis and Disambiguation of Arabic," *Proc. 9th Lang. Resour. Eval. Conf.*, pp. 1094–1101, 2014.
- [22] Y. Souteh and K. Bouzoubaa, "SAFAR platform and its morphological layer SAFAR platform and its morphological layer," no. December 2011, 2015.
- [23] A. Bies and M. Maamouri, "Penn Arabic Treebank Guidelines," *Draft January*, vol. 28, no. December, p. 2003, 2003.
- [24] N. Habash, O. Rambow, and R. Roth, "MADA+TOKAN: A Toolkit for Arabic Tokenization, Diacritization, Morphological Disambiguation, POS Tagging, Stemming and Lemmatization," *Proc. Second Int. Conf. Arab. Lang. Resour. Tools*, no. November 2015, pp. 102–109, 2009.
- [25] G. Leech and A. Wilson, "EAGLES recommendations for the morphosyntactic annotation of corpora," *Version of March*, 1996.
- [26] N. Habash, O. Rambow, and R. Roth, "MADA+TOKAN Manual," 2010.

### Appendix A. Initial Part-of-Speech Tagset

Lexical Markers			Functional Markers – Semantic Driven		
NN	Noun	اسم	LOC	Location	دلالة مكانية
JJ	Adjective	نعت - صفة	TIM	Time	دلالة زمنية
RB	Adverb	حال أو ظرف مكان أو زمان	CZ	Cause	دلالة سبب

RP	Particle	حرف	EFCT	Effect	دلالة النتيجة
IN	Preposition	حرف جر	SEQ	Sequence	دلالة تتابع
PRN	Pronoun	ضمير	BGN	Begin of	دلالة بداية وقت
DT	Determiner	اسم إشارة	END	End of	دلالة نهاية وقت
VBP	Verb Present	فعل مضارع	CND	Condition	دلالة شرط
VBD	Verb Past	فعل ماضي	CNF	Confirmation	دلالة تأكيد
NNP	Proper Name	اسم عاقل	ASRT	Assertion	دلالة اخبار
FW	Foreign Word	كلمة اجنبية	CC	Conjunction	دلالة عطف
VN	Verbal Noun	مصدر	INTR	Interrogative	دلالة استفهام
PP	Passive Participle	اسم مفعول	QTY	Quantity	دلالة كميات
AP	Active Participle	اسم فاعل	NEG	Negation	دلالة نفي
VMP	Imperative	فعل أمر	EXP	Explanation	دلالة تفسير
RPRN	Relative Pronoun	اسم موصول	DMN	Demand	دلالة طلب
WP	Wh-pronoun	اسم استفهام	PRD	Predicate	خبر
<b>Number Markers</b>			PRD	Predicate	WHY
SNG	Single	مفرد	WHN	When	استفسار متى
DUAL	Dual	مثنى	HOW	How	استفسار كيف
PLRL	Plural	جمع	WHO	Who	استفسار من
			SWR	Swearing	دلالة قسم
<b>Gender Markers</b>			SWR	Swearing	MNR
MSC	Masculine	مذكر	DGR	Degree	درجة الفعل
FEM	Feminine	مؤنث	NR	Near	دلالة القرب
			FR	Far	دلالة البعد
<b>Voice Markers</b>			<b>Arabic Declension System Specifics</b>		
stV	First Voice	First Voice Verb	YAA	YAA	يا النداء
ndV	Second Voice	Second Voice Indicator	KANA	KANA	كان واخواتها
rdV	Third Voice	Third Voice Indicator	INNA	INNA	ان واخواتها
<b>Active-Passive Markers</b>			ZRFL	Locative Adverb	ظرف مكان
PSV	Absent Person	صيغة الغائب	ZRFZ	Temporal Adverb	ظرف زمان
ATV	Present Person	صيغة الحاضر	CSA	Accusative Case Ending	علامة النصب
<b>Suffix Markers</b>			CSN	Nominative Case Ending	علامة الرفع
SFX	Attached Pronoun	ضمير متصل	CSG	Genitive Case Ending	علامة الكسر
POSS	Possession	مؤشر على الملكية	CSNU	Nunation Case	علامة التنوين
OBJ	Object Reference	مؤشر على المفعول به	AAN	AAN	عن
<b>Prefix Markers – Functional Particles</b>			ALA	ALA	على
BI	BI	باء	FEE	FEE	في
LI	LI	لام	MEN	MEN	من
FA	FA	فاء	HATTA	HATTA	حتى
SA	SA	سين	ELA	ELA	الى
WA	Waw	واو	SBJ	Subject Reference	مؤشر على الفاعل

# The Performance of the Bond Graph Approach for Diagnosing Electrical Systems

Dhia Mzoughi, Abderrahmene Sallami, Abdelkader Mami

Laboratoire d'Application de l'Efficacité Énergétique et Énergies Renouvelables (LAPER),  
Faculté des Sciences de Tunis,  
Université de Tunis El Manar,  
Campus Universitaire Farhat Hached,  
B.P. No. 94, Rommana, 1068 Tunis, Tunisia

**Abstract**—The increasing complexity of automated industrial systems, the constraints of competitiveness in terms of cost of production and facility security have mobilized in the last years a large community of researchers to improve the monitoring and the diagnosis of this type of processes. This work proposes a reliable and efficient method for the diagnosis of an electrical system. The improvement of the reliability of the systems depends essentially on the algorithms of fault detection and isolation. The developed method is based on the use of analytical redundancy relations allowing the detection and isolation of faults which occur in the various elements of the system using a structural and causal analysis. In this context, the bond graph appears as an interesting approach since it models physical systems element by element which facilitates the detection and location of faults. The simulation of the system is performed through 20-sim software dedicated to the bond graph applications.

**Keywords**—Bond graph; faults detection and isolation; electrical system; analytical redundancy relations

## I. INTRODUCTION

The modelling tool used in this work is the bond graph approach defined by Paynter [1]; it is a graphical representation language of physical systems, based on the modelling of the energy phenomena occurring within these systems. This energy approach allows to highlight the analogies that exist between the different fields of physics (mechanics, electricity, hydraulics, thermodynamics, acoustics, etc.) and to represent in a homogeneous form the multidisciplinary physical systems [2], [3]. In this way, the utility of the bond graph tool for the supervision of industrial systems will be presented.

The paper is structured as follows. Initially, an overview on the bond graph approach is performed. Then, a description of the bond graph representation of a monitoring system is given highlighting the difference between the quantitative approach and the qualitative approach and briefly recalling the method of ARRs generation. Afterwards, the principle of analysis of residue sensitivity using bond graph is presented. Then, the diagnosis by residues generation and the robust diagnosis of an RLC circuit are presented. Finally, the last part is devoted to conclusion.

## II. BOND GRAPH FOR MODELLING

Modelling based on bond graph relies mainly on the concept of generalized effort and flux variables allowing the representation of energy exchanges and balances between the different elements of a system [4]. In this approach, an energy exchange between two elements is represented by a half-arrow link indicating the direction of the transfer. These half-arrows are called bonds; each is labelled by an effort variable  $e$  and a flux variable  $f$ . The product of these two variables corresponds to the power carried by the bond. The advantage of this modelling is that the choice of the effort  $e$  and the flux  $f$  depends only on the physical domain of the system to be represented (Fig. 1).

This description is made in terms of interconnected components by links through the ports at their disposal. The components are classified by the number of ports they have available, these are multi-ports or n-ports as described in [5]. There are three types of bond graph used each in a particular stage of the design process [6]-[8]:

- Word bond graphs where the components represent subsystems described by black boxes, this level allows a first decomposition of the system to have a global view of the energy exchanges implemented.
- The acausal bond graphs where the components are elementary components are indivisible and whose behaviour is known (resistance, inductance, capacitor, etc.), this level is used at an advanced stage of the design process, where the components can be assimilated to perfect elementary components.
- The causal bond graphs enabling to establish the equations of the system.

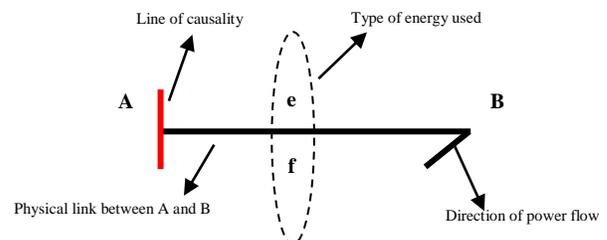


Fig. 1. Representation of a physical system by bond graph.

### III. SUPERVISION OF INDUSTRIAL SYSTEMS USING BOND GRAPH APPROACH

The purpose behind the bond graph representation is to use a single tool for modelling, the generation of analytical redundancy relations (ARRs), structural analysis, system monitoring and diagnostics. A bond graph-based supervision system [9]-[12] can be represented as shown in Fig. 2.

There are essentially two parts: one concerns the transfer of power and energy (constituted of the process and the actuators), while the second represents the signals (the information system, i.e. the sensors and the control system).

The bond graph model represents the energy part of the system. The process is usually modelled by the common bond graph elements (R, C, I, and the junctions). The actuators (pump, heat source, etc.) are modelled by sources of effort and /or flow. The sources can be either simple ( $Se, Sf$ ) or modulated ( $MSe, MSf$ ) (i.e. controlled by an external signal provided by a controller or operator).

The sensors and the control system form the information system. In the first system (energy), the power exchanged is represented by a half-arrow (a power link) evaluated in the effort and flow variables. In the second system (information system), the power exchanged is negligible, it is then represented by an information link (arrow) which is the same used in conventional block diagrams.

The monitoring algorithms (fault detection and isolation FDI) receive information from sensors (effort and flow sensors  $De$  and  $Df$ ) and then deliver alarms to the supervisory system. Information about the faulty elements state is transmitted to the maintenance service.

In what follows, the different approaches of FDI [13], [14] by bond graph are presented. There are two main bond graph approaches to process monitoring: the quantitative approach and the qualitative approach.

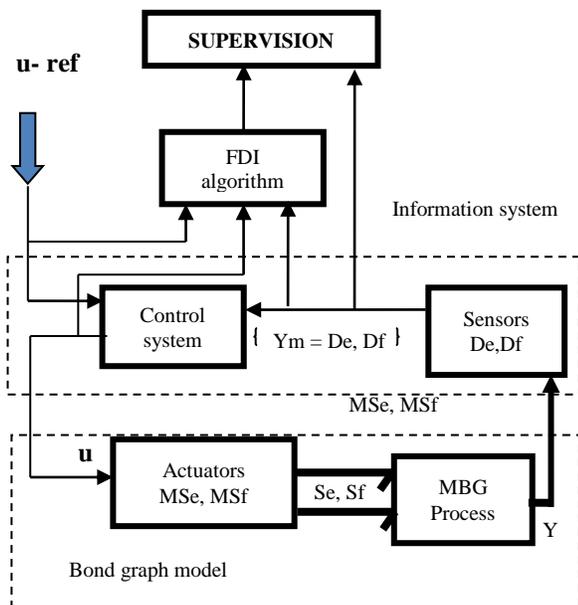


Fig. 2. Bond graph representation of a monitoring system.

#### A. Qualitative Approach

This approach does not require a very precise model. In contrast to the conventional knowledge representations used to describe the structure of the system and its state through various tools (block diagram, differential equations, etc.); the qualitative bond graphs explicitly describe only the location of the components of the system and their interconnections.

#### B. Quantitative Approach

Contrary to the qualitative approach, the quantitative approach is based on physical laws and therefore requires a deep knowledge of the structure of the system and the numerical values of the parameters [15]-[17]. The advantage of this approach is the simplicity of understanding the ARRs since they correspond to relations and variables which are displayed by the bond graph, image of the physical process. These relations are deduced directly from the graphical representation, they can be generated in symbolic form and therefore suitable for computer implementation.

#### C. Generation of ARRs

The method to generate ARRs from linear mono energy bond graph model by following the causal paths is studied in [18]. At the junction structure level (junctions 0, 1, TF and GY), several relations between different flows and efforts can be established. From the algebraic sum of flows on a junction 0 for example and by expressing the variables contained therein as a function of the known variables, it leads to ARRs. The aim is to study all the causal paths relative to the junction considered up to the sources and sensors. The method is interesting because it generates as many relations as junctions.

### IV. ANALYSIS OF RESIDUE SENSITIVITY USING BOND GRAPH MODEL

The sensitivity analysis of residues has been developed in recent years. Indeed, methods are proposed to evaluate these residues. When residues are assumed to be normally distributed around a known average, the statistical methods for generating normal operating thresholds are well suited. In the case where the uncertainties do not occur at the same frequency as the faults, the filtering methods are well adapted while the actuator and sensor faults are determined using the parity space. Unfortunately, these residues generation methods are not effective since they neglect the parametric inter-correlation (the thresholds are often overstated and may diverge).

The bond graph approach provides an efficient solution to the problem of parametric dependencies since the generation by Bond Graph-Linear Fractional Transformations (BG-LFT) automatically separates residues and adaptive thresholds [19]-[22]. In this work, the BG-LFT model will be used to generate residues and adaptive thresholds for normal operation.

#### A. Generation of Performance Indices

To improve diagnostic performance [23], it is necessary to determine the performance indices [24] (sensitivity index and fault detectability index).

1) *Sensitivity index*: The parametric standardized sensitivity index explains the evaluation of the energy provided to the residue by the uncertainty on each parameter

by comparing it with the total energy provided by all the uncertainties.

$$SI_{ai} = \frac{|a_i|}{d} \frac{\partial d}{\partial |a_i|} = \frac{|w_i|}{d} \quad (1)$$

with  $a_i$  is the uncertainty on the  $i^{\text{th}}$  parameter,  $i \in \{R, C, I, TF, GY\}$ ,  $w_i$  the  $i^{\text{th}}$  modulated input corresponding to the uncertainty on the  $i^{\text{th}}$  parameter.

2) *Detectability index*: It represents the difference between the effort (or flux) provided by the defects in absolute value and that provided by the set of uncertainties in absolute value.

- Junction 1:

$$DI = |Y_i| |e_{in}| + |Y_s| - d \quad (2)$$

- Junction 0:

$$DI = |Y_i| |f_{in}| + |Y_s| - d \quad (3)$$

Then the conditions of faults detectability will be as follows:

- Undetectable fault:

$$DI \leq 0 \quad (4)$$

- Detectable fault:

$$DI > 0 \quad (5)$$

## V. BOND GRAPH MODEL OF ELECTRICAL SYSTEM

The diagram of an electrical circuit RLC and its bond graph model are given in Fig. 3. We will detect and locate faults at the effort sensor  $De$ .

### A. Diagnosis By Residues Generation Using Bond Graph

In static mode, the bond graph model is linear, the establishment of the structural equations at the junctions of the bond graph model of Fig. 3 gives us:

- For the junction 1, we find as structural equation:

$$\begin{cases} e_1 = Se = U \\ e_2 = Rf_2 = RC \frac{dDe}{dt} \\ e_3 = L \frac{df_3}{dt} = LC \frac{dDe}{dt} \\ e_4 = e_5 = De \end{cases} \quad (6)$$

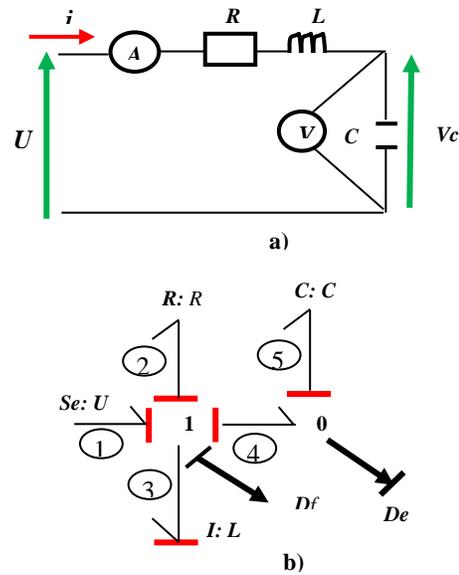


Fig. 3. a) Schematic of an RLC circuit, b) Bond graph model of the RLC circuit.

And using structural equations at the junctions, we get the following equation of the residue  $r_1$ :

$$\begin{cases} r_1 = e_1 - e_2 - e_3 - e_4 \\ r_1 = U - LC \frac{dDe}{dt} - RC \frac{dDe}{dt} - De \end{cases} \quad (7)$$

- For the 0 junction, we find as structural equation:

$$\begin{cases} f_4 = f_2 = \frac{e_2}{R} = \frac{e_1 - e_3 - e_4}{R} = \frac{1}{R} (U - LC \frac{dDe}{dt} - De) \\ f_5 = C \frac{dDe}{dt} \end{cases} \quad (8)$$

From these relations, we can deduce the equation of the residue  $r_2$  independent of the unknown variables of the system:

$$\begin{cases} r_2 = f_4 - f_5 \\ r_2 = \frac{1}{R} (U - LC \frac{dDe}{dt} - De) - C \frac{dDe}{dt} \end{cases} \quad (9)$$

Fig. 4 shows the evolution of residues  $r_1$  and  $r_2$ . The curves show that, in the case of normal operation, the average values of the residues are almost zero.

### B. Robust Diagnosis Using Bond Graph

To check if there's causal conflict or not, the integral model of the RLC circuit should be determined then deduce therefrom the derivative model to determine the residues. Fig. 5 shows the BG-LFT model of the RLC circuit in integral causality.

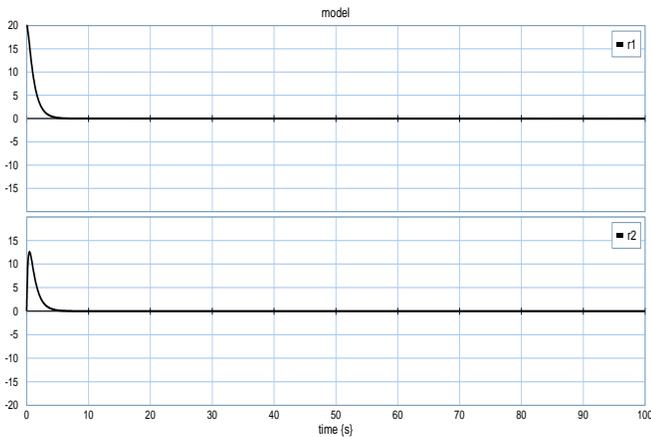


Fig. 4. Evolution curves of residues  $r_1$  and  $r_2$ .

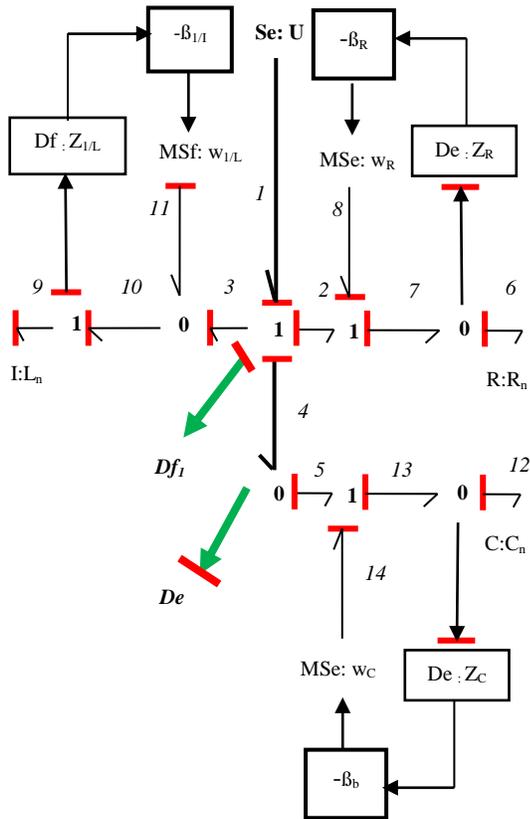


Fig. 5. RLC circuit BG-LFT model in integral causality.

Fig. 6 shows the BG-LFT model of the RLC circuit in derivative causality.

From the BG-LFT model of Fig. 6, we can determine the  $ARR_s$  equations.

1) Junction 1:

$$e_2 : SSf \rightarrow \Psi_{R_n}(f_6, e_6) \rightarrow e_2 = R_n \cdot SSf \quad (10)$$

$$e_3 : SSf \rightarrow \Psi_{L_n}(f_9, e_9) \rightarrow e_3 = L_n \cdot SSf \quad (11)$$

$$e_4 : SSf \rightarrow \Psi_{C_n}(f_{12}, e_{12}) \rightarrow e_4 = (1/C_n) \cdot SSf \quad (12)$$

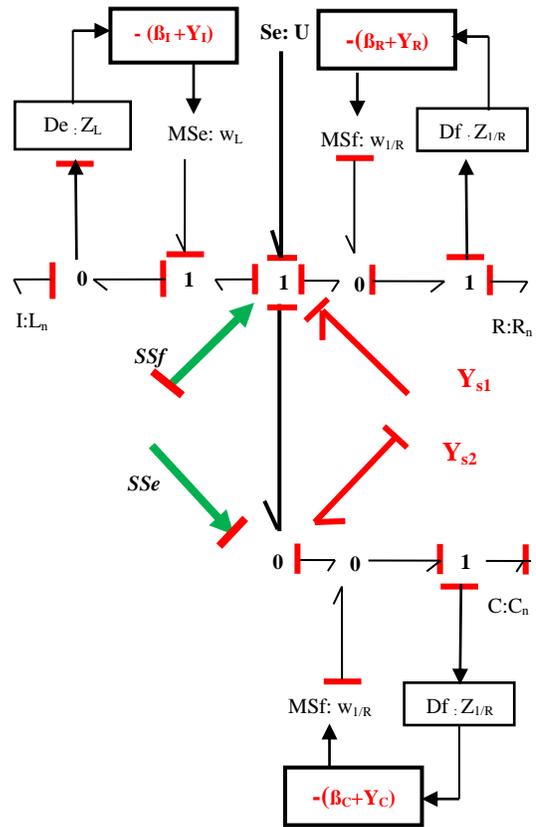


Fig. 6. RLC circuit BG-LFT model in derivative causality.

In case where there is no fault, the  $ARR_I$  equation can be written as:

$$\begin{cases} ARR_1 = r_{1n} + d_1 \\ ARR_1 = U - R_n SSf - L_n \frac{dSSf}{dt} \\ \frac{1}{C_n} \int SSf + w_R + w_L + w_C = 0 \\ r_{1n} = U - R_n SSf_1 - L_n \frac{dSSf_1}{dt} - \frac{1}{C_n} \int SSf \\ d_1 = |w_{1/R}| + |w_L| + |w_C| \end{cases} \quad (13)$$

In the event of a fault, the  $ARR_I$  equation can be written as:

$$\begin{cases} ARR_{11} = r_{11n} + d_{11} \\ ARR_{11} = U - R_n SSf - L_n \frac{dSSf}{dt} \\ \frac{1}{C_n} \int SSf + w_R + w_L + w_C = 0 \\ r_{11n} = U - R_n SSf_1 - L_n \frac{dSSf_1}{dt} - \frac{1}{C_n} \int SSf \\ d_{11} = |w_{1/R}| + |w_L| + |w_C| + |Y_{s1}| + |Y_{R e_{Rn}}| + |Y_{L e_{L1}}| + |Y_C f_{Cn}| \\ d_{11} = d_1 + |Y_{s1}| + |Y_{R e_{Rn}}| + |Y_{L e_{L1}}| + |Y_C f_{Cn}| \end{cases} \quad (14)$$

In this way, the fault detectability index of residue  $ARR_{1l}$  is obtained:

$$DI_1 = d_{11} = d_1 + Y_{s1} + |Y_R e_{Rn}| + |Y_L e_{Ln}| + |Y_C f_{Cn}| \quad (15)$$

- The detectable rate  $Y_R$  of a fault in the element  $R$ :

It is assumed that  $Y_L = Y_C = Y_{s1} = 0$ :

If  $DI_1 > 0$  then

$$Y_R = \frac{|d_1|}{e_{Rn}} \quad (16)$$

- The detectable rate  $Y_L$  of a fault in the element  $L$ :

It is assumed that  $Y_R = Y_C = Y_{s1} = 0$ :

If  $DI_1 > 0$  then

$$Y_L = \frac{|d_1|}{e_{Ln}} \quad (17)$$

- The detectable rate  $Y_C$  of a fault in the element  $C$ :

It is assumed that  $Y_R = Y_L = Y_{s1} = 0$ :

If  $DI_1 > 0$  then

$$Y_C = \frac{|d_1|}{e_{Cn}} \quad (18)$$

- The detectable value  $Y_s$  of the structural fault:

It is assumed that  $Y_R = Y_L = Y_C = 0$ :

If  $DI_1 > 0$  then

$$Y_s > |d_1| \quad (19)$$

2) Junction 0:

$$f_4 : SSf \rightarrow \Psi_{Rn}(f_6, e_6) \rightarrow f_6 = (1/L_n) \cdot SSf \quad (20)$$

$$f_5 : SSe \rightarrow \Psi_{Cn}(f_{12}, e_{12}) \rightarrow f_{12} = C_n \cdot SSf \quad (21)$$

The  $ARR_2$  equation, in case where there is no fault, can be written as:

$$\begin{cases} ARR_2 = r_{2n} + d_2 \\ ARR_2 = \frac{1}{L_n} \frac{dSSf}{dt} - C_n \frac{dSSe}{dt} + w_L + w_C = 0 \\ r_{2n} = \frac{1}{L_n} \frac{dSSf}{dt} - C_n \frac{dSSe}{dt} \\ d_2 = |w_{1/L}| + |w_C| \\ d_2 = \left| \delta_L \frac{1}{L_n} \frac{dSSf}{dt} \right| + \left| \delta_C C_n \frac{dSSe}{dt} \right| + |Y_L e_{Ln}| + |Y_C f_{Cn}| \end{cases} \quad (22)$$

In the case of a fault, the  $ARR_2$  equation can be written as:

$$\begin{cases} ARR_{21} = r_{21n} + d_{21} \\ ARR_{21} = \frac{1}{L_n} \frac{dSSf}{dt} - C_n \frac{dSSe}{dt} + w_L + w_C = 0 \\ r_{21n} = \frac{1}{L_n} \frac{dSSf}{dt} - C_n \frac{dSSe}{dt} \\ d_{21} = |w_{1/L}| + |w_C| + Y_{s2} + |Y_L e_{Ln}| + |Y_C f_{CLn}| \\ d_{21} = d_2 + Y_{s2} + |Y_L e_{Ln}| + |Y_C f_{CLn}| \end{cases} \quad (23)$$

In this way, the fault detectability index of residue  $ARR_{2l}$  is obtained:

$$DI_2 = d_{21} = d_2 + Y_{s2} + |Y_L e_{Ln}| + |Y_C f_{Cn}| \quad (24)$$

- The detectable rate  $Y_L$  of a fault in the element  $L$ :

It is assumed that  $Y_C = Y_{s2} = 0$ :

If  $DI_2 > 0$  then

$$Y_L = \frac{|d_2|}{e_{Ln}} \quad (25)$$

- The detectable rate  $Y_C$  of a fault in the element  $C$ :

It is assumed that  $Y_L = Y_{s2} = 0$ :

If  $DI_2 > 0$  then

$$Y_C = \frac{|d_2|}{f_{Cn}} \quad (26)$$

- The detectable value  $Y_s$  of the structural fault:

It is assumed that  $Y_L = Y_C = 0$ :

If  $DI_2 > 0$  then

$$Y_s > |d_2| \quad (27)$$

With the bond graph approach one can clearly see the residue equations separately compared to their uncertainties as well as to faults that can occur.

## VI. CONCLUSION

In this manuscript, we have shown how to use a bond graph approach for modelling, detection and isolation of fault and simulation of an electrical system. A procedure for the automatic generation of robust residuals and adaptive thresholds for normal operation has been developed and implemented using appropriate software tools. The performance of the diagnosis is controlled by an analysis of the sensitivity of these residues allowing to define indices of sensitivity to parametric uncertainties and indices of detectability of faults. By its physical nature, the bond graph approach allows to estimate the detectable values of physical faults. The choice of the bond graph approach for diagnosis is due to its energetic and multi-physical aspect as well as its structural analysis. The integration of the bond graph model of the electrical system with the observer model will be the object of a future work, in order to rapidly detect faults.

## REFERENCES

- [1] H. M. Paynter, Analysis and design of engineering systems. Boston, USA: MIT Press, 1961.
- [2] D. Karnopp, D. Margolis, and R. Rosenberg, System Dynamics: A Unified Approach. New York: Wiley, 1990.
- [3] Z. Chalh, A. Frih, M. Mrabti, and M. Alfydi, "Bond graph methodology for controllability of LTV systems," *Int. J. Mod. Ident. & Cont.*, vol. 24, pp. 257–265, 2015.
- [4] M. Tagina and G. Dauphin-Tanguy, La méthodologie bond graph. Principes et applications. Centre de Publication Universitaire, 2003.
- [5] M. K. Hales and R. C. Rosenberg, "Structured modelling of mechatronic components using multiport templates," *Proceedings ASME IMECE 2000 DSC*, vol. 69-2, pp. 787-794, 2000.
- [6] N. Zaytoon, Systèmes dynamiques hybrides. Paris: HermesScience Publications, 2001.
- [7] M. Djem Ai and M. Defoort, Hybrid Dynamical Systems. Springer International Publishing, 2015.
- [8] W. Borutzky, Bond graph model-based fault diagnosis of hybrid systems. Switzerland: Springer International Publishing, 2015.
- [9] B. Ould Bouamama, G. Dauphin-Tanguy, M. Straoswcki, and F. Busson, "Bond Graph Technique as a Decision-Making Tool in Supervision Systems," *HKK Conference & Symposium in Graph Theoretic & Entropy Methods in Engineering*, University of Waterloo, June 13-15, 1999, pp. 91-97.
- [10] M. Bayart, B. Ould Bouamama, and B. Conard, "FDI of smart actuators using bond graphs and externalmodel," *IFAC Proceedings*, vol. 35(1), pp. 391-396, 15th Triennial World Congress, Barcelona, Spain, 2002.
- [11] B. Ould Bouamama, K. Medjaher, M. Bayart, A. K. Samantaray, and B. Conard, "Fault detection and isolation of smart actuators using Bond graph and external models," *Cont. Eng. Pract.*, vol. 13 (2), pp. 159-175, 2005.
- [12] B. Ould Bouamama and G. Dauphin-Tanguy, Modélisation Bond Graph Element de base pour l'énergetique. Technique de l'ingénieur, BE 8 280, 2005.
- [13] S. Benmoussa, B. Ould Bouamama, and R. Merzouki, "Bond graph approach for plant fault detection and isolation: Application to intelligent autonomous vehicle," *IEEE Trans. on Auto. Scien. & Eng.*, vol. 11(2), pp. 585-593, 2014.
- [14] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. on Indus. Elect.*, vol. 62(6), pp. 3757-3767, 2015.
- [15] P. M. Frank, "Advanced fault detection and isolation schemes using nonlinear and robust observers," *10th World Congress on Automatic Control, IFAC, RFA*, vol. 20(5), pp. 63-68, July 1987.
- [16] P. M. Frank, "Fault Diagnosis on the basis of Dynamic Process Models," *Proceedings IMACS, 12th World Congress on Scientific Computation*, Paris, France, July18-22, 1998, pp. 414-419.
- [17] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge -based redundancy- a survey and some new results," *Automatica*, vol. 26(3), pp. 459- 474, 1990.
- [18] M. Tagina, Application de la modélisation bond graph à la surveillance des systèmes complexes. PhD Thesis, Lille University, France, 1995.
- [19] D. Henry, A. Zolghari, "Norm-based design of robust FDI schemes for uncertain systems under feedback control: comparison of two approaches," *Cont. eng. pract.*, vol. 14(9), pp. 1081-1097, 2006.
- [20] Z. Han, W. Li, and S. L. Shah, "Fault detection and isolation in the presence of process uncertainties," *Cont. eng. pract.*, vol. 13(5), pp. 587-599, 2005.
- [21] M. A. Djeziri, B. Ould Bouamama, and R. Merzouki, "Modeling and Robust FDI of Steam Generator Using Bond Graph Aproch," *Journal of process control*, vol. 19(1), pp. 149-162, 2009.
- [22] Y. Touati, M. A. Mellal, and D. Benazzouz, "Multi-thresholds for fault isolation in the presence of uncertainties," *ISA trans.*, vol. 62, pp. 299-311, 2016.
- [23] N. Chatti, B. Ould Bouamama, A. L. Gehin, and R. Merzouki, "Signed bond graph for multiple faults diagnosis," *Eng. App. of Art. Intel.*, vol. 36, pp. 134-147, 2014.
- [24] M. A. Djeziri, R. Merzouki, B. Ould Bouamama, and G. Dauphin-Tanguy, "Bond Graph Model Based For Robust Fault Diagnosis," *In American Control Conference ACC'07, IEEE*, New York City, USA, 2007, pp. 3017-3022.

# Application of the Tabu Search Algorithm to Cryptography

Zakaria KADDOURI<sup>1</sup>

Laboratory of computer science research  
Department of Computer Science  
Mohammed V University Abu Dhabi  
Abu Dhabi, United Arab Emirates

Fouzia OMARY<sup>2</sup>

Laboratory of computer science research  
Department of Computer Science  
Mohammed V University – Faculty of Sciences Rabat  
Rabat, Morocco

**Abstract**—Tabu search is a powerful algorithm that has been applied with great success to many difficult combinatorial problems. In this paper, we have designed and implemented a symmetrical encryption algorithm whose internal structure is mainly based on Tabu search algorithm. This heuristic performs multiple searches among different solutions and stores the best solutions in an adaptive memory. First of all, we coded the encryption problem by simulating a scheduling problem. Next, we have used an appropriate coding for our problem. Then we used the suitable evaluation function. Through the symmetric key generated by our algorithm, we have illustrated the principle of encryption and decryption. The experimentations of our approach are given at the end of this paper, from which we examined our new system strengths and the elements that could be improved.

**Keywords**—Symmetric encryption; heuristic; Tabu search; algorithm; scheduling problem; combinatorial problems

## I. INTRODUCTION

Since the invention of writing, *men* expressed the need to transmit information securely making it unintelligible to anyone outside the exchange. Indeed the messages cannot be understood by the enemy, even if they are intercepted. At the current time, information in all its form circulates in digital format throughout the world in a split second on the networks. This information is exchanged every day from one point to another either by telephone, cable, optical fiber or satellite. It is likely to be read, deleted or falsified. Cryptography is a science in full operation and meets the needs of today's information security [1]. Metaheuristics are a family of optimization algorithms that are designed to solve general classes of mathematical problems by combining research procedures to quickly find the best solution [2]. Tabu search is an example of such a heuristic. It starts from an initial solution and attempts to improve it by transforming it iteratively. At each iteration, the neighborhood of the current solution is generated and the best solution in this neighborhood is chosen. To avoid circling, a Tabu list is defined to prohibit revisiting of the solutions already examined [3], [4].

We will apply the Tabu search in the main phase of cryptography. To start, we have transformed the problem of encrypting a message  $M$  to an optimization problem like to Evolutionary Ciphering System [5]-[7]. Then, we coded this problem in a particular way to bring us back to scheduling problems.

This design is based on the enrichment of the search space and the application of a well-defined model during the process of Tabu search. We notice that the random aspect used in our algorithm is crucial to the success of the method, especially if the number of iterations is relatively small.

In the next section of this article, we describe the general algorithm of TABU search and we present a detailed description of our encryption algorithm entitled Symmetrical Tabu Search Ciphering (STSC). Then, we will analyze the security of our approach, and we will compare it with same kind of systems SEC (Symmetrical Evolutionist-based Ciphering) and SMC (Symmetrical Memetic Ciphering).

## II. DESCRIPTION OF TABU SEARCH

### A. Definition

Tabu search was proposed by F.Glover in 1986. The algorithm is called Tabu because there is prohibition from resuming recently visited solutions [8]. Since then, the method has become very popular, thanks to its successes to solve many problems. This algorithm introduces a notion of memory in the strategy of the search space exploration [9]. Tabu search uses local or neighborhood iterative procedures to move from solution  $x$  to a solution  $x'$  (in the vicinity of  $x$ ) until the stopping conditions are met [10].

### B. Principle of Tabu Search

The principle of Tabu search is based on a method of moving on the space of the solutions, while continually seeking to improve the current best solution and by storing in memory the list of previous moves [11], thus guiding the research outside the previously traveled zones.

The basic idea is inspired by the research techniques used in artificial intelligence. That is to keep the track of the past path of the research process in one or more memories and to use this information in order to orient future development. In practice, we will not memorize all the displacement (very costly in memory), but we will prevent only the access to some solutions during a certain number of iterations.

The neighborhood of a solution is defined by an elementary transformation (movement) permitting the switch from a solution to another solution nearby with a slight modification of the structure of the solution.

The Tabu search is based on:

- The use of flexible memory structures (short, medium and long term) allowing the full exploration of the evaluation criteria and the search history.
- A control mechanism based on alternating between the conditions that restrict (restriction Tabu) and those that liberate (aspiration criterion) the search process.
- The incorporation of the strategies of intensification and diversification of the search:
  - Intensification strategy uses the medium term memory, and serves to strengthen the search in the regions of the best solutions found recently.
  - Diversification strategy uses the long term memory, and serves to search in new regions.

C. General Algorithm of Tabu Search

We present below the general algorithm of Tabu search:

- 1) Get an initial solution (initialization).
- 2) Create a list of candidates' movements.
- 3) Choose the best candidate. This choice is based on Tabu restrictions and the aspiration criteria.

This provides an alternative, which will not be registered only if it is better than the previous solution [12].

- 4) Apply the stopping criterion.
  - Continue: change the candidates of eligibility (Tabu restriction and aspiration criterion). Go to 2.
  - Stop: Go to strategies of intensification and diversification.

The flowchart of Tabu search method is shown in Fig. 1.

The general algorithm can be represented with the following pseudo-code:

Let  $NT(s)$  be all candidate solutions,  $T$  the tabu list,  $N(s)$  the neighborhood of solutions and  $s^*$  the current optimal solution:

$NT(s) = \{s' \in N(s) \text{ such as } s' \notin T \text{ or } f(s') < f(s^*)\}$   
Process Tabu\_method (initial solution  $s$ )

```

Put  $T \leftarrow \emptyset$  and  $s^* \leftarrow s$ ;
Repeat
    Choose  $s'$  that minimizes  $f(s')$  in  $N_T(s)$ 
    If  $f(s') < f(s^*)$  then put  $s^* \leftarrow s'$ 
    Put  $s \leftarrow s'$  and update  $T$ 
Until the termination criterion is satisfied
End
    
```

III. DESCRIPTION OF OUR ALGORITHM

A. Problem formalization

We denote by  $M$ , the binary encoding of the message  $M_0$ :

We represent the message to be encrypted by the lists which are the elements of a partition of the set  $\{1, 2, \dots, m\}$ . The lists are composed by the different positions of each binary block. Let  $B_1, B_2, \dots, B_m$  the different blocks of  $M$ .

Let  $L_i (1 \leq i \leq m)$  a list containing the different positions of the block  $B_i$  and  $card(L_i)$  the number of occurrences of  $B_i$ .

Note. This breakdown only takes place for larger size messages.

We note that  $L_i \cap L_j = \emptyset$  if  $i \neq j, \forall i, j \in \{1, 2, \dots, m\}$ .

The message  $M$  may be represented by the following vector:

(B1,L1)	(B2,L2)	...	(Bm,Lm)
---------	---------	-----	---------

Our algorithm seeks to create a maximum disorder in the positions of the blocks. For this, we iteratively change the distribution of lists  $L_i (1 \leq i \leq m)$  on the different blocks of  $B$  (without changing the content of the lists) so that the difference between the cardinal of the new list assigned to each block  $B_i$  and the cardinal of the original list  $L_i$  is maximal [13]. Therefore, we are faced with a problem of optimization and we can use the Tabu search method, including that used in scheduling problems. The latter has several versions, the most used is the one described below:

Definition of variables

- $i$ : the current solution
- $i'$ : the next solution achieved (neighbor solution)
- $N(i)$ : the space of neighboring solutions at  $i$  (the set of  $i'$ )
- $m$ : movement from  $i$  to  $i'$
- $Best\_Sol$ : the global optimal solution that minimizes the objective function  $f(i)$ .
- $i^*$ : the current optimal solution  $f(i^*)$
- $T$ : list of Tabu movements. There can be multiple lists simultaneously. The elements of the list are  $t(i, m)$ .
- $a(i, m)$ : criterion for aspiration. Determines when it is advantageous to undertake  $m$ , despite its status Tabu.

B. Skeleton of the Algorithm

To represent this, we have:

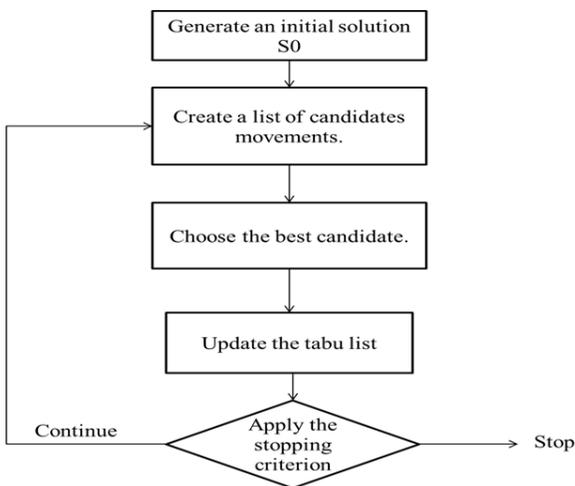


Fig. 1. Flowchart of Tabu search method.

- Tabu List: contains prohibited movements.
- Movement (to go from one solution to another) is to swap randomly the positions of the two lists of the current solution
- The function  $f$  to minimize: eliminates solutions for which only a minority of list values have changed over initial solution.

**Step 1:** Choose an initial solution  $i$  in  $S$  (the set of solutions) that we call *Original-Sol*

A solution is a vector  $v$  of size  $m$ . The content of  $v$  is the list  $L_i$  ( $1 \leq i \leq m$ ) of position blocks.  $L_j$  is the  $j^{th}$  list which contains the new positions that will take the block  $B_j$ .

We apply a random permutation algorithm on the initial solution.

$i^* = i$

$k = 0$

$T = 0$

**Step 2:**  $k = k + 1$  and generate a neighborhood of solutions in  $N(i, k)$ :

The neighborhood of the solutions will be generated by the application of permutations on the positions of the lists. Precisely, we apply random permutation on the positions of the current solution in order to generate neighboring solutions.

- The Tabu movements are not selected.
- An aspiration criteria  $a(i, m)$  is applicable.

**Step 3:** Choose the best solution  $i'$  from the set of neighboring solutions  $N(i, k)$

$i = i'$

Let  $i'$  be a solution of  $N(i, k)$  in which the lists are  $L'_{j1}, L'_{j2} \dots L'_{jm}$  and let  $f$  be the evaluation function on the set of solutions  $i'$

by:

$$f(i') = -\sum_{i=1}^m |card(L'_{ji}) - card(L_i)|$$

**Step 4:** If  $f(i) \leq f(i^*)$ , a better solution was found

$i^* = i$

**Step 5:** Update the list  $T$  and the criteria aspiration.

Add the best solution in Tabu list (it is Tabu for the next iterations).

**Step 6:** If a stop condition is reached, stop.

Alternatively, return to Step 2.

The flowchart of *STSC* algorithm is shown in Fig. 2.

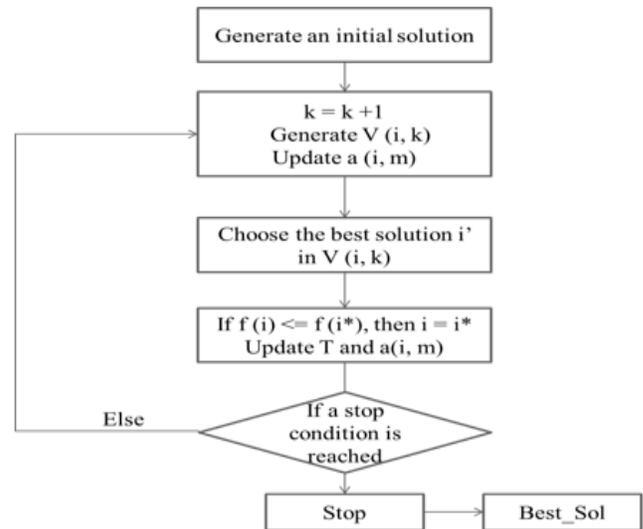


Fig. 2. Flowchart of *STSC* algorithm.

Let *Best\_Sol* (the global optimal solution) the final solution given by *STSC*. We create our encryption key (*Tabu-key*) from *Original\_Sol* and *Best\_Sol*, Then, using the *Tabu-key* generated by our algorithm we set the corresponding cipher text block by changing the distribution lists on the various characters of the message  $M$ . Then, we concatenate the encrypted blocks (obtained by different processes). Thus, we obtain the encrypted message  $M$  from the original message  $M_0$ .

### C. Decryption

Decryption must begin by looking for the reciprocal operation of the last encryption one. The message  $M$  will be broken down again into  $m$  blocks  $B_i$  that have the same size. Because of the *Tabu-key* the blocks are going to recover their lists of corresponding positions [22].

The principle of encryption and decryption can be summarized by the scheme in Fig. 3.

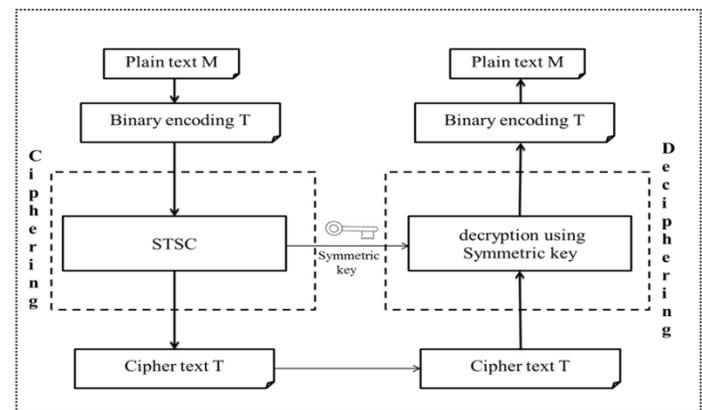


Fig. 3. Scheme of our encryption system.

IV. EXPERIMENT RESULTS

We used our algorithm to encrypt and decrypt files such as document, image, audio, etc. The experimental environment is as follows: Processor: Intel® Core™ i5-2450M (2.50 GHz), 4 Go of RAM; Operating System: Windows 7-x64; Programming language: Java.

A. Comparison of the Frequencies Analysis

Comparing the frequency analysis is a very important indicator in cryptography [14]-[16]. To illustrate the performance of the new system *STSC*, we tested our program on several messages with different sizes.

The results are shown graphically in Fig. 4.

In fact, due to the binary coding and the implementation of the system encryption *STSC*, the frequencies of the characters are no longer recognized. Therefore, cryptanalysis based on the study of the occurrence frequency cannot rely on incorrect statistics.

B. Configuration

We test our system on several texts of different sizes, and for each one of them, we try to find the best parameters to achieve the optimal solution in an ideal time. For this, we record the results on the number of iterations needed for the convergence of the system. Table 1 shows these results.

We can see that in the case where the size of the binary blocks is  $k = 5$  or  $k = 6$ , our system converges and generates the encryption key in less operations compared to the other cases.

C. Security Analysis

1) Security key

Let  $X_n = \{1, 2, \dots, n\}$  be a permutation of  $n$  separate lists. For  $n \in \mathbb{N}^*$ , denote by  $E_n$  the set of the possible permutations of  $X_n$ ,  $E_n = \langle X_1; X_2; X_3; X_4; \dots; X_n \rangle$ .

Counting the permutations of  $X$  back to enumerate all  $n$ -tuples formed of integers from  $1$  to  $n$  in some order. There are  $n$  choices for the first term of the permutation. Then for each of these first choice, there are  $n-1$  possibilities for the second choice,  $n-2$  to the third, and so on. Finally, according to this principle the cardinal of  $E_n$  is  $n!$

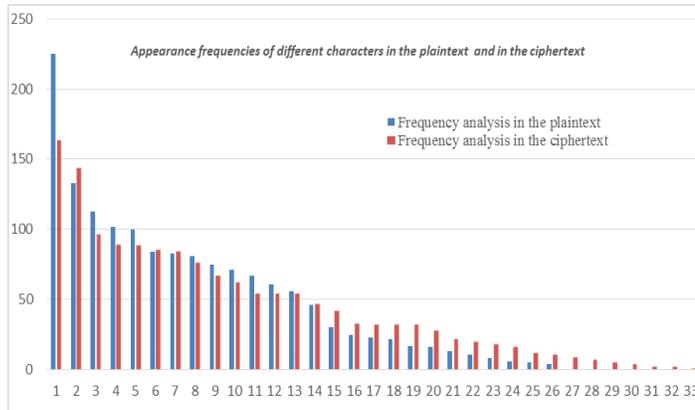


Fig. 4. Graphical representation of the appearance frequencies of different characters in the plaintext and the cipher text with *STSC*.

TABLE. I. RECAPITULATIVE OF THE RESULTS OF THE CONVERGENCE SYSTEM

Size of plaintext	Neighbors	Size of Blocks						
		5	6	7	9	10	11	12
1000 Characters	20	47	57	44	55	64	79	91
	30	59	60	78	71	60	75	89
	40	67	54	90	78	76	89	98
3000 Characters	20	42	59	76	63	98	86	73
	30	56	48	79	98	120	93	112
	40	66	54	96	102	80	115	104
6000 Characters	20	46	66	59	75	92	84	106
	30	61	63	72	88	90	81	95
	40	71	72	86	99	93	82	99
10000 Characters	20	56	61	77	73	93	107	90
	30	58	67	81	70	83	85	94
	40	53	51	80	79	76	87	111

Denote by *NL* the maximum number of different lists constituting our key, *KN* the maximum number of different keys generated by *STSC* and by *PA* the probability to encounter the right key is equal to the inverse of *KN*.

Table 2 summarizes the relation between the number of different blocks and the security of our approach.

It is noted that even in the case where the size of the text desired to be encrypted is small and the number of different blocks is reduced, the probability to fall on the correct key is very small or almost nil. We also note that the security of our approach clearly increases along with the number of different blocks.

2) Complexity of Brute Force Attack

The key length is an important security parameter [17]. Generally, the level of security of the encryption system is based on the size of the keys used (the longer the key size, the more robust the encryption system). The key to our system is composed of two elements: the Tabu-key and the block size '*k*'.

- The *Tabu-key* size is the product of the number of different blocks (NDB) and 8 bits.

TABLE. II. SUMMARY SHOWING THE RELATION BETWEEN THE NUMBER OF DIFFERENT BLOCKS AND THE SECURITY OF OUR APPROACH

Size of blocks	<i>NL</i>	<i>KN</i>	<i>PA</i>
5	$2^5$	32!	$3,80 e^{-36}$
7	$2^7$	128!	$2,59 e^{-216}$
9	$2^9$	512!	$2,87 e^{-1167}$
10	$2^{10}$	1024!	$1,84 e^{-2640}$
11	$2^{11}$	2048!	$5,97 e^{-3895}$

TABLE. III. RECAPITULATIVE PRESENTATION OF THE RESULTS OF THE KEYS SIZE GENERATED BY STSC AND THE COMPLEXITY OF THE BRUTE-FORCE ATTACK

Size of Plaintext	Data	Size of blocks					
		5	6	7	9	10	11
1000 characters	NDB	26	55	109	297	310	485
	STSC Key (bit)	216	448	880	2384	2488	3888
	Complexity of BFA	$2^{216}$	$2^{448}$	$2^{880}$	$2^{2384}$	$2^{2488}$	$2^{3888}$
3000 characters	NDB	27	59	115	315	324	599
	STSC Key (bit)	224	480	928	2528	2600	4800
	Complexity of BFA	$2^{224}$	$2^{480}$	$2^{928}$	$2^{2528}$	$2^{2600}$	$2^{4800}$
6000 characters	NDB	30	68	116	335	378	629
	STSC Key (bit)	248	552	936	2688	3032	5040
	Complexity of BFA	$2^{248}$	$2^{552}$	$2^{936}$	$2^{2688}$	$2^{3032}$	$2^{5040}$

We calculate the number of different blocks existing in the texts to determine the size of the encryption STSC key.

In the case of keys used by the system STSC the length is given in bits. In this case, the number of possibility to explore for a brute force attack (BFA) is in the order of  $2^N$  where  $N$  is key length in bit, since the key is randomly generated.

Table 3 summarizes the results of the key size generated by STSC and the complexity of the brute-force attack.

With the current technology a 128-bit key length is already a limit impossible to achieve. If we compare the minimum key length of our system with the recommended size for symmetric systems that ensure a basic security, we can deduce that our system is able to resist against the brute-force attacks more than most other existing systems. The attacker must consider other cryptanalysis strategies if they exist. It should nevertheless take into account that the power of computers is increasing every day and an indecipherable message today can be decipherable in the future.

D. Comparison Between the Performances of STSC System with Existing Systems

To illustrate the effectiveness and performance of the new STSC system compared to older systems, we tested our program on multiple messages with different sizes and we recorded the number of iterations required for the system to converge.

The comparison is based not only on the quality of the results but also on the speed of convergence and computation.

Table 4 summarizes the values of the parameters used by the SEC, SMC and STSC systems to encrypt a text whose size is 6000 characters.

Table 5 presents the results obtained by applying the three precited algorithms.

Fig. 5 to 7 shows graphs of different values of the evaluation function.

TABLE. IV. VALUES OF PARAMETERS RELATING TO SEC, SMC AND STSC SYSTEMS

Parameters	SEC	SMC	STSC
Population size	30	30	-
Probability of crossover	0.7	0.7	-
Probability of mutation	0.03	0.03	-
Size of neighborhood	-	10	50
Size of the tabu list	-	-	100
Aspiration Criterion	-	-	50
Nombre d'itérations Max	100	100	200

TABLE. V. SIMULATION RESULTS OBTAINED BY SEC, SMC AND STSC SYSTEMS

Optimization methods	Evolutionary algorithm (SEC)	Memetic algorithm (SMC)	Tabu search (STSC)
Number of iterations	49	28	88
Optimum Global	5	20	50
Execution time (ms)	69	55	47
Time for a single iteration (ms)	1,40	1,96	0,53

They actually present the best configuration for which the global minimum is obtained.

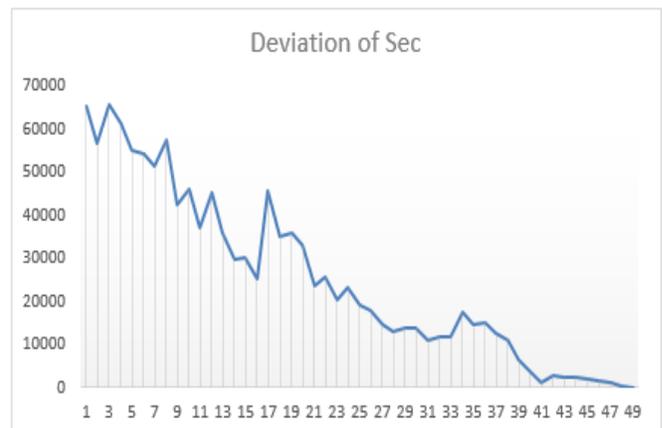


Fig. 5. Evolution of the cost by applying the SEC algorithm.

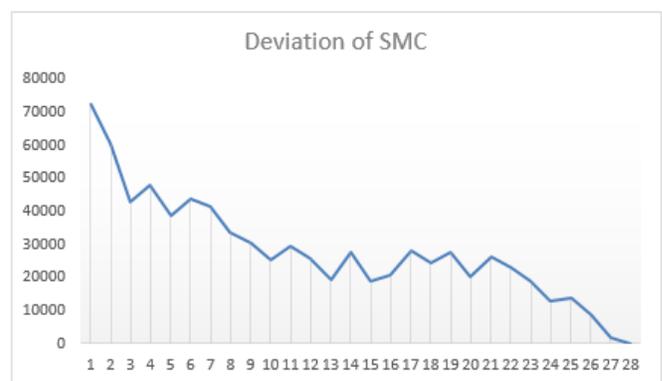


Fig. 6. Evolution of the cost by applying the SMC algorithm.

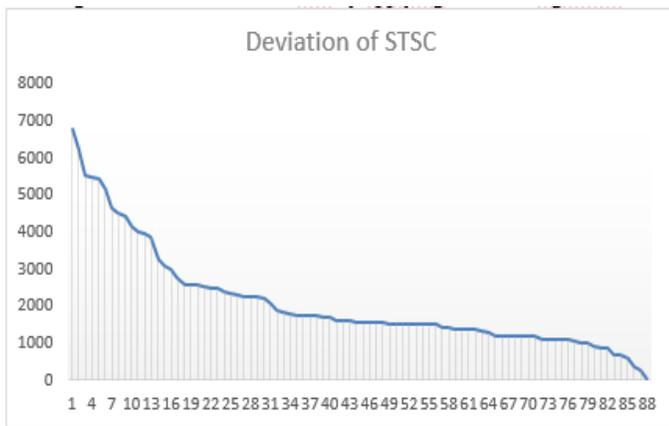


Fig. 7. Evolution of the cost by applying the STSC algorithm.

From Fig. 5, we can notice that the convergence of the SEC system is achieved in the 49th iteration to a global minimum equal to 5. From Fig. 6, the SMC system stops at the 28th iteration with an optimum value of 20 and from Fig. 7 we can see that the best configuration for the STSC system, giving the optimum overall equal to 50, is reached after 88 iterations.

From Table 5, the encryption system STSC turns out faster in terms of computation time than the other two SEC and SMC systems.

## V. CONCLUSION

Tabu search is a very efficient new meta-heuristic; it can solve a wide range of problems.

In this article, the first adaptation of the meta-heuristic “Tabu search” cryptography was presented. The proposed algorithm uses a variable-length encoding to represent a symbol of the data input, which allows the encryption of any kind of information (text, image, sound, etc.).

Our system generates a secret key that we call “*Tabu-key*” which has the essential qualities to be efficient and able to resist against the brute-force attacks. According to the results of the occurrence frequencies of the characters obtained, we have shown that this method blocks the way against all the attacks which are based on the study of the occurrence frequencies of characters in a cipher text. We can also increase the security of our system by combining it with another encryption method such as [18]-[22].

## REFERENCES

[1] Delfs H, Knebl H. Introduction to Cryptography: Principles and Applications. Second edition. Springer. United States of America, March 2007.

[2] Colin B. Reeves. Modern Heuristic Techniques for Combinatorial Problems. JohnWiley & Sons. Great Britain, 1993.

[3] Glover F, Laguna M. Tabu Search. Kluwer Academic Publishers, Norwell, MA, 1997.

[4] Glover F, Kelly J.P, Laguna M. Genetic Algorithms and Tabu Search: Hybrids for Optimization. Computers and Operations Research 1995; 22: 111 – 134.

[5] Omary F. Applications des algorithmes évolutionnistes à la cryptographie. University of science, Rabat, Morocco, 2006.

[6] Omary F, Tragha A, Lbekkouri A, Bellaachia A, Mouloudi A. An Evolutionist Algorithm to Cryptography. Brill Academic Publishers – Lecture Series and Computational Sciences 2005; 4: 1749-1752.

[7] Omary F, Tragha A, Bellaachia A, Mouloudi A. Design and Evaluation of Two Symmetrical Evolutionist-Based Ciphering Algorithms. International Journal of Computer Science and Network Security (IJCSNS) 2007; 7: 181-190.

[8] Glover F, Tabu search, part I. ORSA. Journal of Computing 1989; 1: 190-206.

[9] Ji M, Tang H. Global Optimizations and Tabu Search Based on Memory. Applied Mathematics and Computation 2004; 159: 449 - 457.

[10] Hanafi S. On the Convergence of Tabu Search. Journal of Heuristics 2001; 7: 47-58.

[11] Hertz A, Taillard E, Werra D. A Tutorial on Tabu Search. Proc. of Giornate di Lavoro AIRO 1995; 95: 13-24.

[12] Hertz A, Widmer M. Guidelines for the use of meta-heuristics in combinatorial optimization. European Journal of Operational Research 2003; 151: 247-252.

[13] Mouloudi A, Omary F, Tragha A, Bellaachia A. An Extension of evolutionary Ciphering System. International Conference on hybrid Information Technology 2006; 1: 314-321.

[14] Labouret G. Introduction à la cryptographie. Support du cours du cabinet Hervé Schauer (HSC). France, 2001.

[15] Stinson D. cryptography theory and practice. Discrete Mathematics and Its Applications. Canada, 2003.

[16] Schneier B. Applied Cryptography. John Wiley & Sons, France, 1996.

[17] Florin G, Natkin S. Les techniques de la cryptographie. CNAM. France, 2002.  
<http://deptinfo.cnam.fr/Enseignement/DESS/surete/securite/courcry.pdf>

[18] Omary F, Tragha A, Lbekkouri A, Bellaachia A, Mouloudi A. A New Ciphering Method Associated with Evolutionary Algorithm. Computational Science and Its Applications; ICCSA 2006; Lecture Notes in Computer Science 2006; 3984: 346-354.

[19] Kaddouri Z, Mise en œuvre de nouvelles techniques pour la securite informatique basees sur les algorithmes evolutionnistes et les fonctions de hachage. University of science, Rabat, Morocco, 2014.

[20] Kaddouri Z, Omary F, Abouchouar A. Binary Fusion Process to the Ciphering System “Sec Extension to Binary Blocks”. Journal of Theoretical and Applied Information Technology 2013; 48: 067-075.

[21] Kaddouri Z, Omary F, Abouchouar A, Daari M. Balancing Process to the Ciphering System Sec. Journal of Theoretical and Applied Information Technology 2013; 52: 092-093.

[22] Kaddouri Z, Omary F. New Symmetrical Ciphering Approach Based on Memetic Algorithm. International Journal of Engineering and Technology 2014; 6: 2728-2737.

# Hearing Aid Method by Equalizing Frequency Response of Phoneme Extracted from Human Voice

Kohei Arai<sup>1</sup>

1 Graduate School of Science and Engineering  
Saga University  
Saga City, Japan

Takuto Konishi<sup>1</sup>

1 Graduate School of Science and Engineering  
Saga University  
Saga City, Japan

**Abstract**—Hearing aid method by equalizing frequency response of phoneme which is extracted from human voice is proposed. One of the problems of the existing hearing aid is poor customization of the frequency response compensation. Frequency response characteristics are different by the person who need hearing aid. The proposed hearing aid is based on frequency response equalization by phoneme by phoneme. Frequency characteristics of phoneme are to be equalized. This is the specific feature of the proposed hearing aid method. Through experiments, it is found that the proposed hearing aid by phoneme is superior to the conventional hearing aid.

**Keywords**—Hearing aid; phoneme; frequency response; equalization filter; hidden markov model (HMM)

## I. INTRODUCTION

In general, hearing capability of human voices is getting bad for elderly persons since a high-frequency response of elderly persons' ears is getting poor. Hearing capability is defined with the well-known averaged hearing capability level that is defined as Averaged value of hearing capability for human voices regarding frequency components ranged from 500 Hz to 4000 Hz. In accordance with the definition, 25-40 dB of loudness of human voices are difficult to hear slightly when human voice is not loud while 40-70 dB of loudness of human voices are difficult to hear when human voice is normal level.

Earlier devices, known as ear trumpets or ear horns [1], [2], were passive funnel-like amplification cones designed to gather sound energy and direct it into the ear canal. After that not so small number of methods have been proposed so far [3]-[10].

Mobile device based personalized equalizer for improving the hearing capability of human voices for elderly persons are proposed. Through experiments, it is found that the proposed equalizer does work well for improving hearing capability by 2 to 55% of voice Recognition success ratio. According to the investigation of the frequency component analysis and formant detections, most of the voice sounds have the formant frequencies for the first to third frequencies within the range of 3445 Hz. Therefore, a nonlinear equalizing multiplier is better to enhance the frequency components for the first to third formants. The experimental results with the voice above input experiments show that a good Percent Correct Recognition: PCR is required for 0 to more than 8000 Hz of frequency components. Also, 8162 Hz cut off frequency

would be better for both noise suppressions and keeping a good PCR [11].

As I described above, hearing capability is getting deteriorated for aged persons. It is called "Senile deafness". In Japan, around 18% of peoples whose age ranged from 65 to 74 have a trouble on hearing capability while 40% of peoples whose age is more than 74 have a trouble on hearing capability. There are some young peoples who have a trouble on hearing capability for some specific frequency component. Although they need a hearing aid, most of they do not like to have such conventional hearing aid due to some reasons. It does not look good. Hearing capability, frequency response varied for time being. Hearing capability is different by person. There are some other reasons.

Because of these reasons, a customization of hearing aid is required. Also, equalization of specific spectrum components is required. Furthermore, it would be better to equalize specific frequency component by phoneme by phoneme if they would like to hear human voices. Therefore, human voice hearing capability improvement method by equalizing frequency response equalization by phoneme by phoneme is proposed. This is the specific feature of the proposed hearing aid method.

The following section describes the proposed method for equalization followed by some experiments. Then conclusions are described together with some discussions and future research works.

## II. PROPOSED METHOD

### A. Frequency Response Model

Fig. 1 shows the cochlea of human ear model.

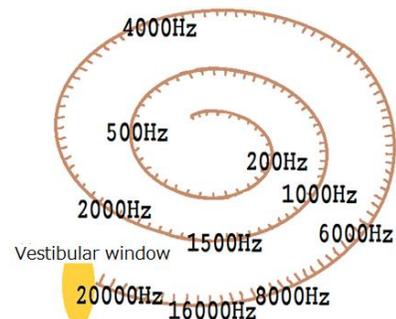


Fig. 1. Cochlea of human ear model.

From the vestibular window to the end of the cochlea, frequency response is varied from high to low frequency components ranged from zero to 20KHz around. Usually, high frequency response is getting deteriorated by age. On the other hand, some frequency response is degraded for the young generation's deafness.

**B. Japanese Phoneme Model**

Frequency range of each Japanese phoneme is shown in Table 1. In Japanese, there are just 23 of phoneme. The number of phoneme is different by Language. The number of phoneme of Japanese is smallest followed by Germany (the number of phoneme is 25).

It is considerably certain that it would better to equalize by phoneme by phoneme because frequency component of each phoneme is different each other. This is the fundamental idea of the proposed hearing aid. Also, it is realized by using smartphone or i-phone as an application software installed on the mobile devices. Therefore, it can be customized by human and may be changed the equalization characteristics even if their frequency response is changed for time being. Also, it can be worked in a real-time basis because the equalization filter can be created in prior to use.

**Auditory Steady-State Response:** ASSR [12] allows to measure frequency response of human ear objectively (Galambos et al. (1981) [13], Rickards et al. (1994) [14], Kuwada et al. (1986) [15]). During sleep, frequency response can be measured using ASSR.

TABLE I. FREQUENCY RANGE OF EACH JAPANESE PHONEME<sup>1</sup>

		Frequency Range(Hz)
Vowel	a	0~1500
	i	0~1000,4000~5000
	u	0~700
	e	0~1000,2000~3000
	o	0~1000
Consonant	k	900~3500
	s	4000~5100
	t	4000~5100
	c	~
	n	3900~4900
	m	0~400
	r	0~1000,4000~4500
	g	0~500,2000~5000
	z	0~500,4000~5000
	d	0~500,3900~5000
	b	0~1000
	p	200~700
	j	0~1000,4000~5000
	w	~
Special Mora	n	0~400
	q	~
	H	~

<sup>1</sup> <http://www.geocities.jp/myonseil/>

It is proposed to measure responses by input 23 of different phoneme to human ear using ASSR. Then appropriate equalizer for each phoneme is designed and installed it to smartphone or i-phone in prior to use.

**C. Procedure of the Proposed Design of Equalization**

Before using the proposed equalizer, customization of the equalization is required. The most appropriate equalization filter response is designed as follows:

- 1) Frequency response characteristic of each phoneme is measured with ASSR.
- 2) Equalization filter is designed by each phoneme.

Phoneme is extracted from the acquired voice signals based on Hidden Markov Model: HMM<sup>2</sup> which is shown in Fig. 2. "Julius" software which is developed by Julius development team composed with Kyoto University, Nagoya Institute of Technology, etc. which allows speech recognition.<sup>3</sup>

First, input voice signals are divided into the frames (25 ms in this case) with the pre-assigned short term shift of the signals (10 ms in this case) as shown in Fig. 3.

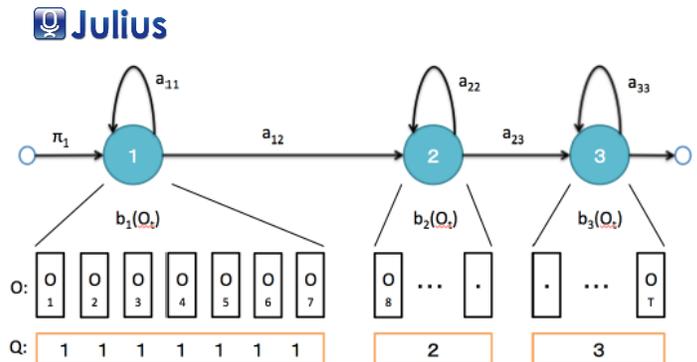


Fig. 2. Example of the well known Hidden Markov Model: HMM.

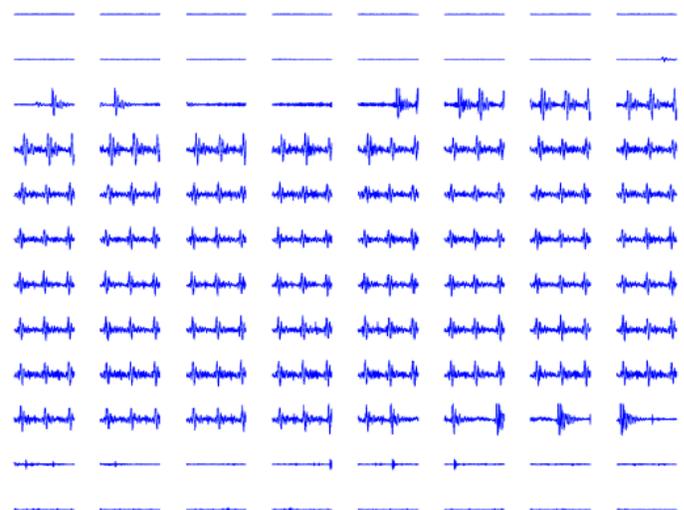


Fig. 3. Divided input voice signal into frames with 25ms of width.

<sup>2</sup> [https://en.wikipedia.org/wiki/Hidden\\_Markov\\_model](https://en.wikipedia.org/wiki/Hidden_Markov_model)

<sup>3</sup> <https://julius.osdn.jp/juliusbook/ja/julius.html>

After that, phoneme is extracted from the frame signal with the quality assessed results “n\_score” as shown in Fig. 4. The frames are attached frame ID and assessed frames are attached “unit number”. These are candidates of the phoneme. The most reliable phoneme is selected from the candidates. In the case of Fig. 4, #2 of units are selected depending on the assessed “n\_score”.

Input voice signals are equalized using previously designed equalizing filters by each phoneme. The equalizing filter is designed as a bandpass filter as shown in Fig. 5. Such bandpass filter can be synthesized by composing low-pass, bandpass and high-pass filters. The low-pass filter suppresses the existing noises while bandpass filter enhances the required frequency response.

The high-pass filter suppresses a low frequency noise. Another method for creating equalizing filter is a composition of low-pass and high-pass filter which are shown in Fig. 6. By combine the two low-pass and high-pass filters, an arbitrary frequency response of equalizing filter can be designed.

The filter responses are candidates of the low-pass filters (see Fig. 7). From these candidates, calm frequency response of filter is selected.

id:	from	to	n_score	unit
[ 0	13]	-20.464859	silB #1	
[ 14	14]	-25.480347	silB #2	
[ 15	19]	-28.562841	silB #3	
[ 20	20]	-38.377106	ky #1	
[ 21	31]	-36.201962	ky #2	
[ 32	35]	-37.868408	ky #3	
[ 36	40]	-34.720387	o: #1	
[ 41	55]	-25.332415	o: #2	
[ 56	56]	-31.801636	o: #3	

Fig. 4. Phoneme is extracted from the frame signal with the quality assessed results “n\_score”.

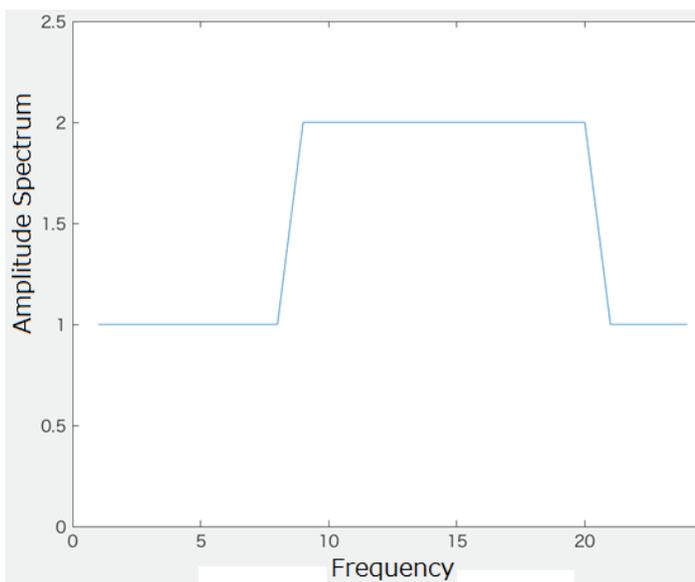
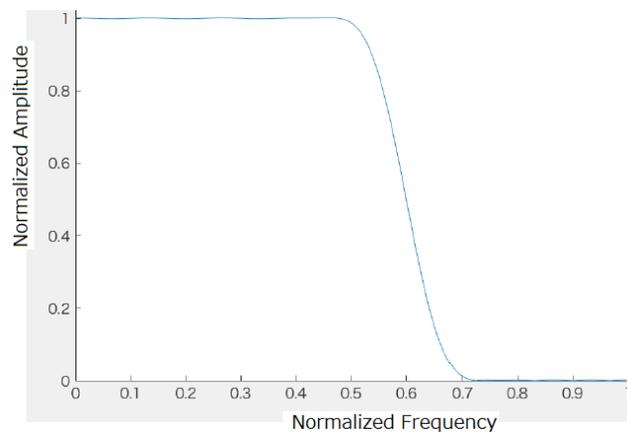
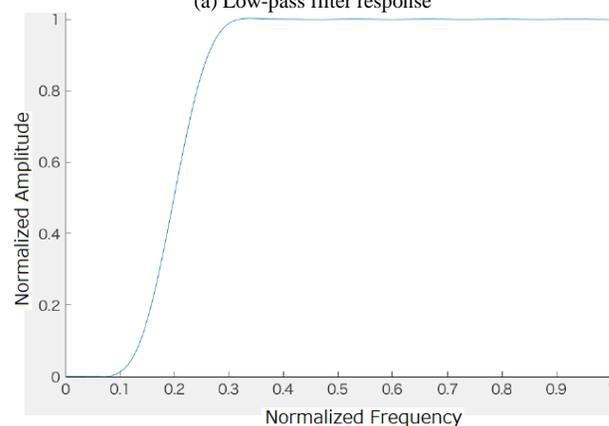


Fig. 5. Frequency response of the required equalizing filter.

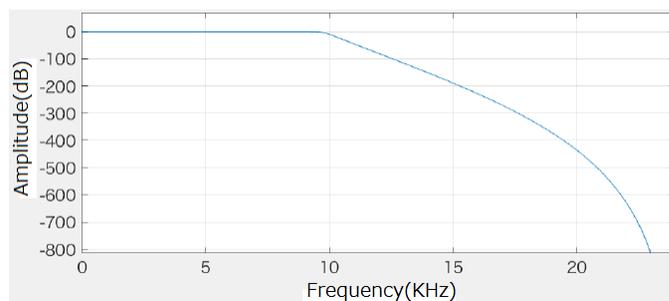


(a) Low-pass filter response

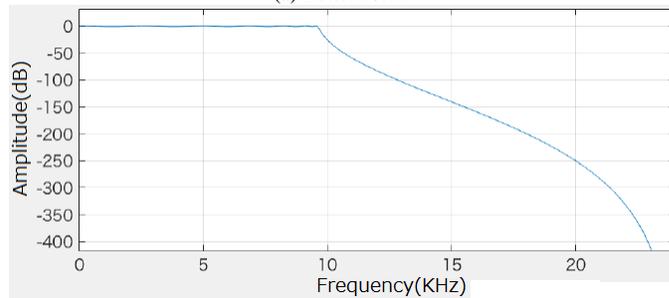


(b) High-pass filter response

Fig. 6. Arbitrary bandpass filter frequency response creation with low-pass and high-pass filters.



(a) Butterworth<sup>4</sup>



(b) Cevichef I<sup>5</sup>

<sup>4</sup> [https://en.wikipedia.org/wiki/Butterworth\\_filter](https://en.wikipedia.org/wiki/Butterworth_filter).

<sup>5</sup> <https://ja.wikipedia.org/wiki/%E3%83%81%E3%82%A7%E3%83%93%E3%82%B7%E3%82%A7%E3%83%95%E3%83%95%E3%82%A3%E3%83%A B%E3%82%BF>.

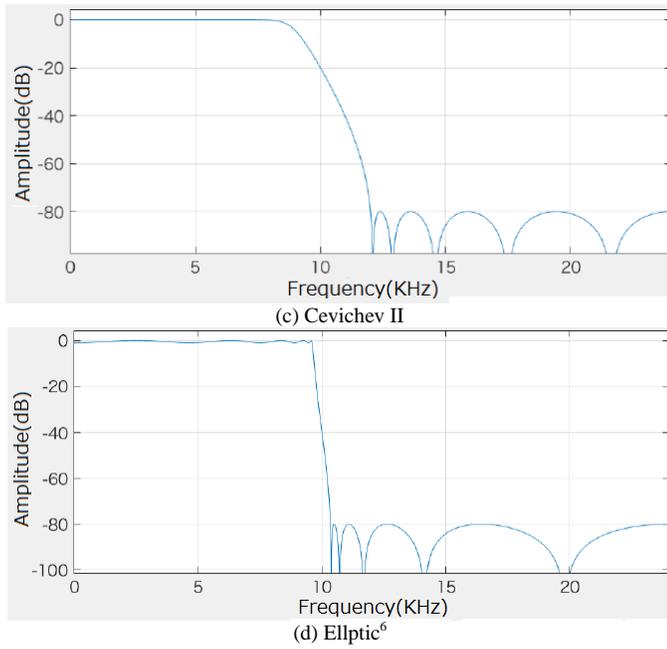


Fig. 7. Candidates of the low-pass filter of frequency responses for equalizing filter.

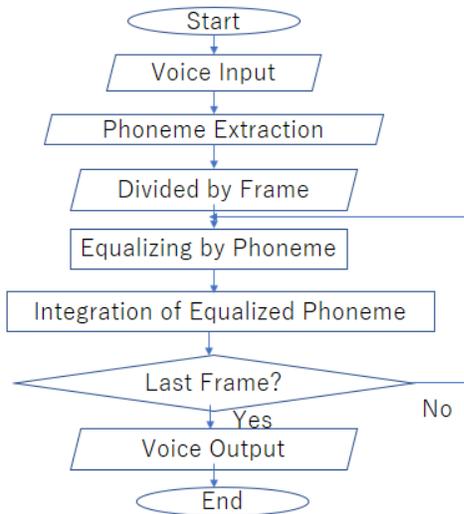


Fig. 8. Detailed flow chart of the proposed procedure.

The detailed flow chart of the proposed procedure is shown in Fig. 8.

After the voice is input in the PC with microphone, phoneme is extracted from the input voice signal followed by division of phoneme by 25ms of frame. Then equalization filter is retrieved by phoneme database followed by integration of the equalized phoneme until the end of the divided frames. After that, the equalized voice signal is output from the PC with speaker.

<sup>6</sup> [https://en.wikipedia.org/wiki/Elliptic\\_filter](https://en.wikipedia.org/wiki/Elliptic_filter)

### III. EXPERIMENTS

#### A. Experimental Environment

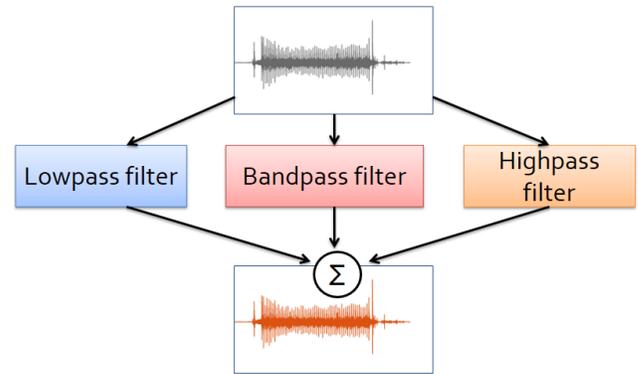
Experimental environment is shown in Table 2. The entire program used for the experiment is based on Matlab.

TABLE II. EXPERIMENTAL ENVIRONMENT

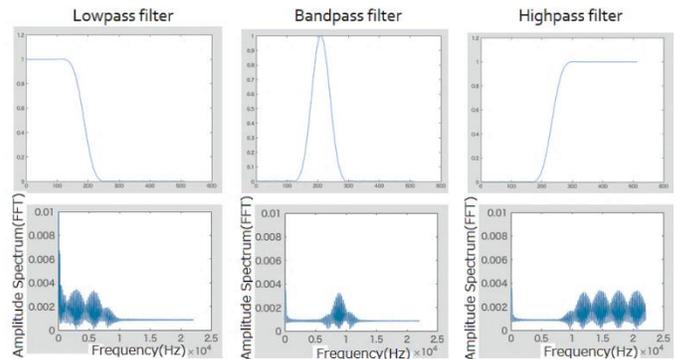
PC	MacBook_Pro_Retina_Mid_2014
OS	MacOS_X_10.9.4
CPU	2.6GHz_Intel_Core_i5
Main_Memory	8GB_1600MHz_DDR3
Programing_Lanbguage	Matlab
Software	MATLAB_R2015b

#### B. Preliminary Experiment

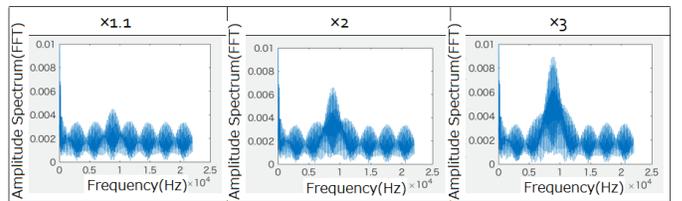
The basic idea behind the proposed equalizing filter is illustrated in the Fig. 9(a). Example of the designed low-pass, high-pass and bandpass filters are shown in Fig. 9(b). Meanwhile, specific frequency ranges can be enhanced as shown in Fig. 9(c).



(a) Illustrative view of the basic idea of the proposed equalizing filter



(b) Three filters



(c) Arbitrary frequency ranges are enhanced

Fig. 9. Basic idea of the proposed equalizing filter and example of the frequency responses of the designed low-pass, high-pass and bandpass filters.

C. Experimental Results

One of the examples of actual spectrum of phoneme is shown in Fig. 10. This is an example of “a”. There are peaks which are named as Formants (from the first to n-th formants) which represent features of the input voices.

Appropriate frequency ranges which must be enhanced are determined with the formants. These formants are estimated with envelopes of frequency spectrum of each phoneme. Then appropriate filter response can be designed by the method.

Fig. 11(a) shows the frequency responses with frequency enhancement while Fig. 11(b) shows the frequency responses without enhancement. #2 in Fig. 4 must be enhanced while #1 and #3 has not to be enhanced. The left image shows #1 and #3 of frequency response while the right image shows #2 of frequency response which must be equalized.

The processed voice signals by the proposed frequency response equalization are shown in Fig. 12. The left image is the original voice input signal while the right image shows the reconstructed output voice signal after the frequency equalization. These are corresponding to the voice signals which are shown in Fig. 11(a) of the left and the right images, respectively.

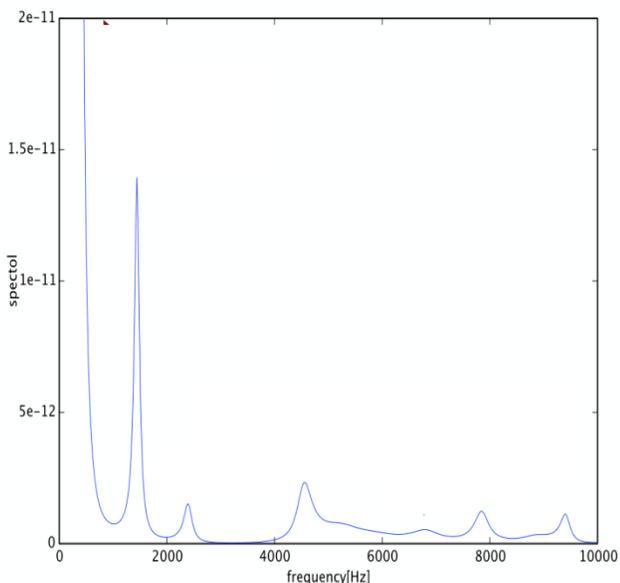
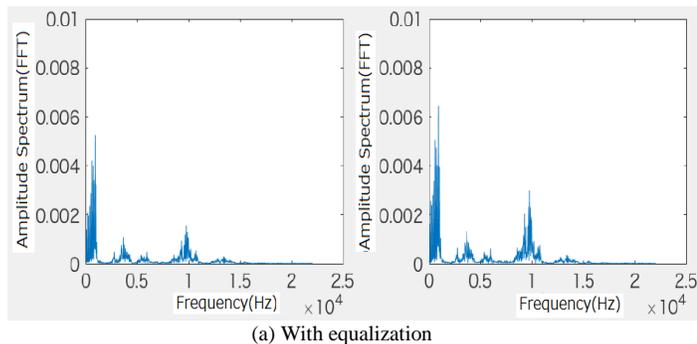
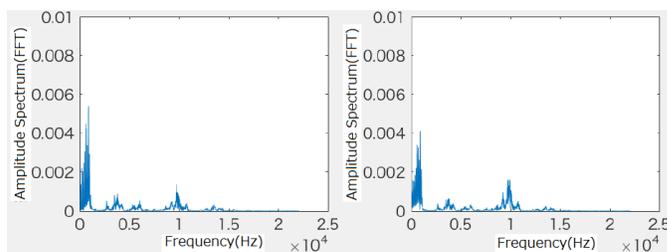


Fig. 10. Example of formant of “a”.



(a) With equalization



(b) Without equalization

Fig. 11. Frequency responses with and without frequency enhancement.

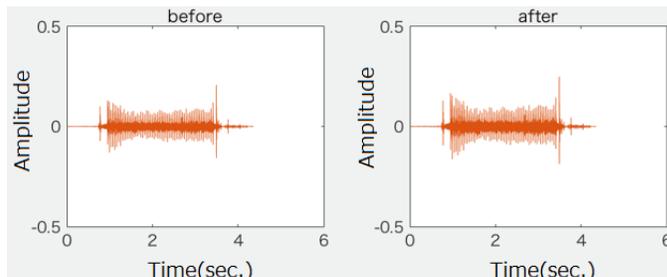


Fig. 12. Example of the processed voice signals by the proposed frequency response equalization.

Four of patients participate a validation test for the proposed system. “Kyo-wa Ii-Tenki-da” in Japanese (“It is fine day” in English) is pronounced by the user. 44,100Hz (Sampling frequency) / 16bit (Quantization bit) / monaural voice signal is created. Also, a degraded input voice signal is created by using low-pass filter with “butter-worse” filter with the cut-off frequency at the 5 KHz. This is called as #1 input voice signal hereafter. The #2 input voice signal is also created with conventional frequency equalization with high-pass filter (the cut-off frequency is at 5 KHz). Another #3 input voice signal is created with the proposed method of frequency equalizer. The four patients hear these three input voice signals and then evaluate the quality of voice with 5 grades. Table 3 shows the evaluation results.

As the results from the evaluation experiments for three input voice signals, it is found that the proposed method shows superior performance to the other two degraded voice signals and the restored voice signal with conventional high-pass filter about 10 points. It is noticed that some of the consonances are not clear enough though. Also, it is noticed that #3 input voice is not so natural since reconstruction is made some sound defects caused by the combining the different frame signal peace of phoneme for the proposed frequency equalization method. In comparison to the conventional method, the reconstructed voice signal by the proposed method is not so noisy. This is one of the features of the proposed method.

TABLE III. EVALUATED RESULTS FOR THREE INPUT VOICE SIGNALS

Input_Voice_Signal	Score	Comments
#1 Voice_Signal	2.67	Relatively_unclear
#2 Conventional	3.5	Consonance_is_not_clear_enough, Noisy
#3 Proposed	3.83	Comparatively_consonance_is_clear, Not_so_noisy

#### IV. CONCLUSION

Human voice hearing capability is improved by equalizing frequency response equalization by phoneme by phoneme. One of the problems of the existing hearing aid is poor customization of the frequency response compensation. Frequency response characteristics are different by the person who need hearing aid. The proposed hearing aid is based on frequency response equalization by phoneme by phoneme. Through experiments, it is found that the proposed hearing aid by phoneme is superior to the conventional hearing aid by the factor of 9.4 %.

It is found that the proposed method shows superior performance to the other two degraded voice signals and the restored voice signal with conventional high-pass filter about 10 points. It is noticed that some of the consonances are not clear enough though. Also, it is noticed that #3 input voice is not so natural due to the fact that reconstruction is made some sound defects caused by the combining the different frame signal piece of phoneme for the proposed frequency equalization method. In comparison to the conventional method, the reconstructed voice signal by the proposed method is not so noisy. This is one of the features of the proposed method.

Further investigations are required for simultaneous estimation of cornea curvature center and cornea radius, noise removal of the depth image.

#### ACKNOWLEDGMENT

The authors would like to thank the fourth group members of the Department of Information Science, Saga University for their contribution to the experiments.

#### REFERENCES

- [1] Bentler Ruth A., Duve , Monica R. (2000). "Comparison of Hearing Aids Over the 20th Century". *Ear & Hearing* 21 (6): 625–639. doi:10.1097/00003446-200012000-00009.
- [2] Bentler, R. A.; Kramer, S. E. (2000). "Guidelines for choosing a self-report outcome measure". *Ear and hearing* 21 (4 Suppl): 37S–49S. doi:10.1097/00003446-200008001-00006. PMID 10981593. edit
- [3] Jack Katz; Larry Medwetsky; Robert Burkard; Linda Hood (2009). "Chapter 38, Hearing Aid Fitting for Adults: Selection, Fitting, Verification, and Validation". *Handbook of Clinical Audiology* (6th ed.). Baltimore MD: Lippincott Williams & Wilkins. p. 858. ISBN 978-0-7817-8106-0.
- [4] K. Sickel, Shortest Path Search with Constraints on Surface Models of In-ear Hearing Aids 52. IWK, Internationales Wissenschaftliches Kolloquium (*Computer science meets automation Ilmenau 10. - 13.09.2007*) Vol. 2 Ilmenau : TU Ilmenau Universitätsbibliothek 2007, pp. 221-226.
- [5] K. Sickel et al., Semi-Automatic Manufacturing of Customized Hearing Aids Using a Feature Driven Rule-based Framework *Proceedings of the*

*Vision, Modeling, and Visualization Workshop 2009* (Braunschweig, Germany November 16–18, 2009), pp. 305-312.

- [6] Dave Fabry, Hans Mülder, Evert Dijkstra (November 2007). "Acceptance of the wireless microphone as a hearing aid accessory for adults". *The Hearing Journal* 60 (11): 32–36. doi:10.1097/01.hj.0000299170.11367.24.
- [7] Hawkins D (1984). "Comparisons of speech recognition in noise by mildly-to-moderately hearing-impaired children using hearing aids and FM systems". *Journal of Speech and Hearing Disorders* 49 (4): 409. doi:10.1044/jshd.4904.409.
- [8] Ricketts T., Henry P. (2002). "Evaluation of an adaptive, directional-microphone hearing aid". *International Journal of Audiology* 41 (2): 100–112. doi:10.3109/14992020209090400.
- [9] Lewis M Samantha, Crandell Carl C, Valente Michael, Horn Jane Enrietto (2004). "Speech perception in noise: directional microphones versus frequency modulation (FM) systems". *Journal of the American Academy of Audiology* 15 (6): 426–439. doi:10.3766/jaaa.15.6.4.
- [10] Exemption from Preemption of State and Local Hearing Aid Requirements; Applications for Exemption, Docket No. 77N-0333, 45 Fed. Reg. 67326; Medical Devices: Applications for Exemption from Federal Preemption of State and Local hearing Aid Requirements, Docket No. 78P-0222, 45 Fed. 67325 (Oct. 10, 1980).
- [11] Kohei Arai, Takuto Konishi, Mobile device based personalized equalizer for improving hearing capability of human voices in particular for elderly persons, *International Journal of Advanced Research on Artificial Intelligence*, 4, 6, 23-27, 2015.
- [12] American Academy of Pediatrics. Task Force on Newborn Hearing Screening. Newborn and infant hearing loss: detection and intervention. *Pediatrics* 103:53-68., 1999.
- [13] Galambos R, Makeig S, Talmachoff PJ. A 40 Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A*, 1981.
- [14] Rickards FW, Tan LE, Cohen LT, et al. Auditory steady-state evoked potentials in newborns. *Br J Audiol* 28:327-337., 1994.
- [15] Kuwada S, Batra R, Maher VL. Scalp potentials of normal and hearing-impaired subjects in response to sinusoidal, 1986.

#### 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 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 37 books and published 570 journal papers. He received 30 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> .

# New Divide and Conquer Method on Endmember Extraction Techniques

Ihab Samir, Bassam Abdellatif

Data Reception, Analysis and Receiving  
station Affairs Division  
National Authority of Remote Sensing and Space Sciences  
Cairo, Egypt

Amr Badr

Computer Science department,  
Faculty of Computer & Information Sciences  
Cairo University  
Cairo, Egypt

**Abstract**—In hyperspectral imagery, endmember extraction (EE) is a main stage in hyperspectral unmixing process where its role lies in extracting distinct spectral signature, endmembers, from hyperspectral image which is considered as the main input for unsupervised hyperspectral unmixing to generate the abundance fractions for every pixel in hyperspectral data. EE process has some difficulties. There are less distinct endmembers than its mixed background; also, there are endmembers that have rare occurrences in data that are considered as difficulties in EE process. In this paper, we propose a new technique that uses divide and conquer method for EE process to find out these difficult (rare or less distinct) endmembers. divide and conquer method is used to divide hyperspectral data scene to multiple divisions and take each division as a standalone scene to enable endmember extraction algorithms (EEAs) to extract difficult endmembers easily and finally conquer all extracted endmembers from all divisions. We implemented this method on real dataset using three EEAs: ATGP, VCA, and SGA and recorded the results that outperform the results from usual endmember extraction techniques methods in all used algorithms.

**Keywords**—Endmember extraction algorithm (EEA); endmember extraction (EE); automatic target generation process (ATGP); hyperspectral imagery; simplex growing algorithm (SGA); hyperspectral unmixing; vertex component analysis (VCA); divide and conquer method

## I. INTRODUCTION

Endmember extraction is considered to be an important and crucial step in hyperspectral data exploitation. A pixel in hyperspectral data may be either a pure pixel or mixed pixel. A pure pixel represents an endmember (EM) that exists in the scene. A mixed pixel contains multiple contributions from a group of different endmembers that exists in the scene. Therefore, endmember is considered as a pure signature for a class [1]. Generally, an endmember is not a pixel; it is a spectral signature which is specified completely by the spectrum of a single material substance.

Several endmember extraction methods have been developed to extract pure pixels from hyperspectral data. Here we use three different algorithms for extracting endmembers from hyperspectral data. The first one is Automatic Target Generation Process (ATGP) that finds its targets by using a sequence of orthogonal subspaces with the maximal orthogonal projections [2], [5], [7], [8] where ATGP

considered the unsupervised version of Orthogonal Subspace Projection (OSP) algorithm. The second used algorithm is the Simplex Growing Algorithm (SGA) [3], [8] which finds its endmembers by growing a simplex, vertex by vertex, until it reaches the required endmembers represented by vertices of simplex. The last used algorithm is the Vertex Component Analysis (VCA) [4], [8], it is an OP-based EEA that is characterized by computational complexity reduction by replacing simple volume calculation with OP and growing nonnegative convex hulls, vertex by vertex, until it builds a  $p$ -vertex convex hull ( $p$  denotes the endmembers required to be extracted).

Authors in [6], demonstrate some EEAs as ATGP, VCA, and SGA and demonstrate their efficiency by using different criteria as sequential or parallel implementation, dimensionality reduction, etc. ATGP, VCA, SGA are most widely used in EE [8]. They are similar in their design but different in preprocessing steps.

Some researches work in spatial and spectral information of hyperspectral data to enhance EEAs. Over segmentation based method introduced in [9], exploit spatial and spectral information to enhance computational performance for EEA. A new enhancement for EEAs is suggested in [10] that gives guidance to EE process for spatially homogenous regions and consequently to enhance performance of unmixing process.

This paper contributed to enabling EEAs to find difficult endmembers where EEAs alone couldn't find them without using this proposed method.

This paper is organized as follows. Section 2 introduces Linear Mixture Model. Section 3 describes the proposed method. Dataset used is introduced in Section 4. Results and discussions are provided in Section 5. The conclusion is given in Section 6.

## II. LINEAR MIXTURE MODEL

Linear mixture model is a well-known approach used for determination and quantification of materials in hyperspectral images. Hyperspectral image consists of pixels where every pixel is represented by a vector of values for each spectral band which, in its turn, is the reflectance of the material in a specific wavelength.

Let  $\mathbf{r}$  be an  $L \times 1$  column vector in a hyperspectral image where  $L$  refers to the number of bands. Suppose that there are

$p$  materials in the hyperspectral image and  $\mathbf{M} = [m_1 m_2 \dots m_p]$  is an  $L \times p$  matrix of material signature, where  $m_j$  is an  $L \times 1$  column vector of the  $j^{\text{th}}$  material signature in the hyperspectral image. Assume that  $\mathbf{a}$  is a  $p \times 1$  abundance column vector denoted as  $(a_1, a_2, \dots, a_p)^T$  which associated with  $\mathbf{r}$  ( $a_k$  represents the abundance fraction of the  $k^{\text{th}}$  signature exist in the pixel vector  $\mathbf{r}$ ).

Linear unmixing can solve this mixed pixel problem. It assumes that spectral signature  $\mathbf{r}$  can be represented by a linear regression model as in (1) where  $\mathbf{r}$  is linearly mixed by  $p$  material signatures.

$$\mathbf{r} = \mathbf{M}\mathbf{a} + \mathbf{n} \tag{1}$$

Where  $\mathbf{n}$  is noise. In unsupervised hyperspectral unmixing process, hyperspectral image pixel represented by  $\mathbf{r}$  with  $\mathbf{M}$  and  $\mathbf{a}$  are unknown. Endmember extraction algorithms come to extract  $\mathbf{M}$  matrix from hyperspectral image to be used as an input in a linear unmixing method that plays its role to unmix the unknown abundance fractions matrix by an inverse of the linear mixture model.

### III. PROPOSED METHOD

From the spectral viewpoint, endmembers in the scene have distinct signatures. These endmembers are the target of any EEA regardless its design and implementation. EEAs used extract all vertices in the simplex as endmembers as shown in Fig. 1 where vertices of great triangle are E1, E4, and E5 but vertices in small triangle are E1, E2, and E3. Endmembers set in small triangle is different from endmembers set in bigger triangle according to EEAs viewpoint. Notice E2, and E3 cannot be extracted from the bigger triangle unless we divide data into sections that will raise the probability of extracting them using different EEAs used.

In this section, a new technique that uses divide-and-conquer method in endmember extraction algorithms is proposed.

Not necessarily that all extracted pixels are pure pixels and represent material signature resident in hyperspectral scene. Usually, some of the extracted pixels, using the EEAs, are mixed. This is normal because each EEA has its strategy in finding endmember set. EEAs suffer from not finding all materials signatures. The proposed technique tries to solve this problem and enhance EEAs results. To test the method, we used real dataset (as explained in the next section) along with its ground truth abundant matrix. Fig. 2 explains the workflow used in the proposed technique.

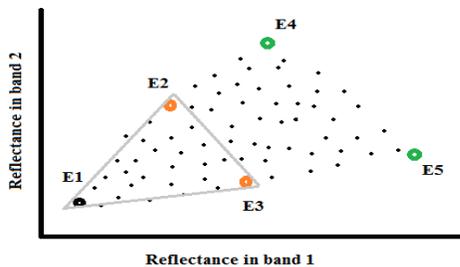


Fig. 1. 2-dimensional plot for many pixels includes 5 endmembers.

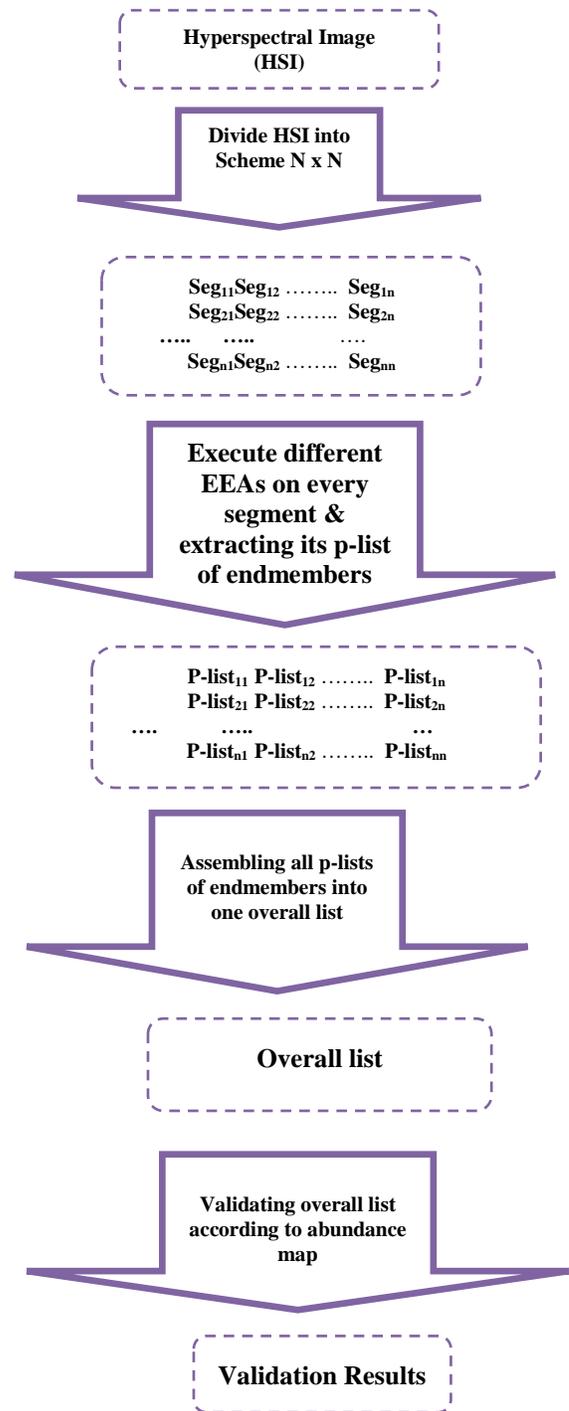


Fig. 2. Workflow used in the proposed technique.

There are five stages in the workflow, beginning with hyperspectral image (HSI). To begin, HSI is divided spatially into scheme  $N \times N$  to create  $N^2$  equivalent segments, where every segment is considered as a standalone HSI.

Following this, EEAs are applied into every segment to extract  $p$  endmembers and record them into  $p$ -list ( $p$  refers to the expected number of endmembers in HSI).

In the subsequent stage, all  $p$ -lists that are created are assembled into overall-list that contains  $p \times N^2$  endmembers.

Next, each endmember in the overall-list is validated using abundance map.

A. The different schemes used in the proposed method are as the following:

- No division scheme: dataset remains as one segment and EEAs are applied on the full dataset and results are recorded.
- Scheme 2×2: dataset is divided spatially into 2 by 2 grids that yield 4 equivalent segments.
- Scheme 3×3: dataset is divided spatially into 3 by 3 grids that yield 9 equivalent segments.
- Scheme 4×4: dataset is divided spatially into 4 by 4 grids that yield 16 equivalent segments.

#### IV. REAL DATASET (JASPER RIDGE)

Jasper Ridge is one of the popular datasets used in hyperspectral data analysis [11]-[12]. Jasper Ridge is a cube of data consists of 512 rows × 614 columns × 224 bands. Its spectral range is starting at 0.38 micron and ending at 2.5 micron. For simplicity, we cut a subset from the original dataset consisting of 100 rows × 100 columns as shown in Fig. 3.

This subset is starting from pixel at 105th row and 269th column from the whole dataset. Because of some effects of atmosphere and water vapor absorptions, 26 bad bands are discarded from total 224 bands as follows: 1:3,108:112,154:166,220:224. The number of remaining bands is 198 were used for analysis.

There are four endmembers in Jasper Ridge data: *Tree*, *Water*, *Soil*, and *Road*. Their abundance images are shown in Fig. 4(a). Jasper Ridge dataset has an abundance map that restricted by Abundance Non-negativity Constraint (ANC) and Abundance Sum-to-one Constraint (ASC). Due to some noise and other calibration problems, we suppose that the pixel which has abundance fraction greater than 90 % is considered as a pure pixel. Fig. 4(b) illustrates pure pixels for every endmember in the map.

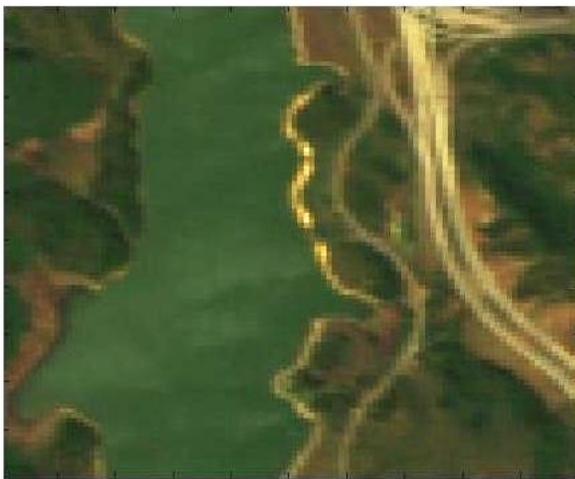


Fig. 3. Jasper Ridge subset consisting of 100rows x 100columns.

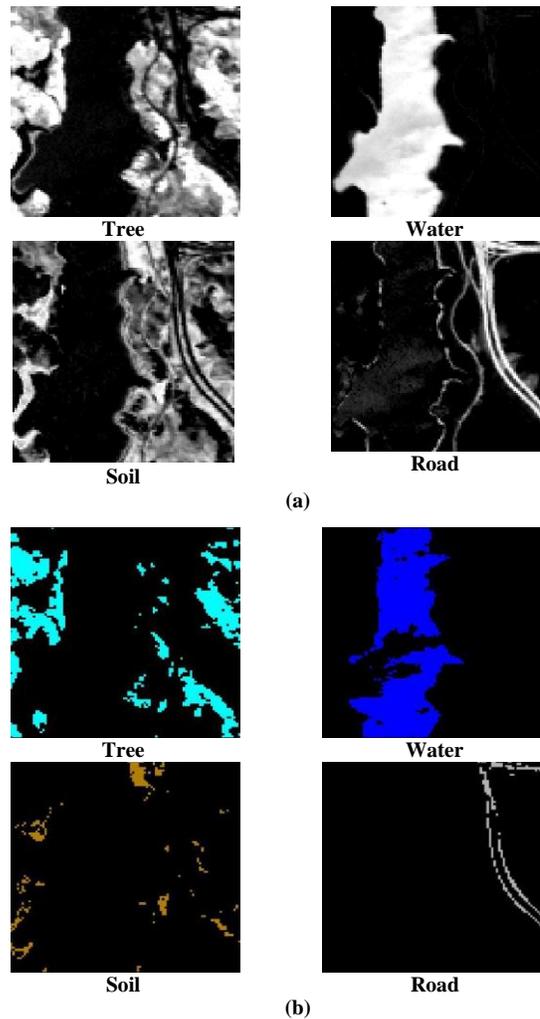


Fig. 4. (a) Abundance images for four endmembers.of Jasper Ridge dataset (b) Pure pixels for each endmember in Jasper Ridge dataset.

#### V. RESULTS AND DISCUSSIONS

In this section, a full description for experiments is executed on Jasper Ridge dataset and synthetic dataset. It gives an overall analysis which demonstrates that results from EEAs using the proposed method D&C outperforms results from EEAs without D&C.

There are three EEAs used in the experiments (ATGP, SGA, and VCA). ATGP is a deterministic algorithm, where it can extract the same set of endmembers for different runs, so it was executed only one time. As opposed to ATGP, VCA and SGA are random algorithms, so they were executed three different runs and their results were recorded separately.

Applying an EEA on any dataset using No Division Scheme will give set of  $p$  extracted endmembers (where  $p$  is the number of endmembers resident in this dataset). Using Scheme 2×2, there are  $4 \times p$  extracted endmembers by applying any EEA. Also in Scheme 3×3 and Scheme 4×4 there are  $9 \times p$  and  $16 \times p$  extracted endmembers respectively.

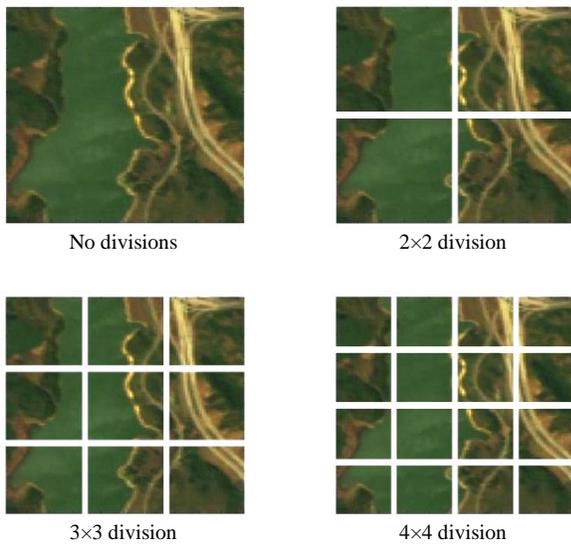


Fig. 5. Different division schemes for Jasper Ridge dataset.

Fig. 5 illustrates different division schemes used on Jasper dataset. The workflow, introduced before, will be applied on every scheme and the extracted endmembers sets will be gathered into one overall set for every unique scheme.

Table 1 illustrates number of divisions and number of expected extracted endmembers according to different schemes used in Jasper Ridge dataset.

Tables 2 to 4 demonstrate results from applying EEAs of different division schemes on Jasper dataset.

According to abundance image for Jasper Ridge dataset, all pixels extracted by EEAs from Jasper dataset are validated as follows:

**A. No Divisions Scheme Results**

Table 2 demonstrates results extracted from EEAs by using No Divisions Scheme. ATGP extracted four pixels; two of them were pure pixels one for *Tree* and another for *Soil*. The other two pixels were mixed pixels, and ATGP couldn't extract any pure pixels for *Water* or *Road*.

VCA #1 as VCA #2, they extracted two pure pixels one for *Water* and another for *Soil*, but also they couldn't extract any pure pixels for *Tree* and *Road*. But, VCA #3 extracted three pure pixels from the four extracted pixels, only *Road* couldn't be extracted.

All SGA runs extracted here the same two pure pixels, *Tree* and *Soil*, while *Water* and *Road* didn't have any pure pixels with SGA.

TABLE I. NUMBER OF DIVISIONS AND NUMBER OF EXTRACTED ENDMEMBERS IN EVERY SCHEME

	No Divisions	2x2 Scheme	3x3 Scheme	4x4 Scheme
# of Divisions	1	4	9	16
# of extracted endmembers	4	16	36	64

TABLE II. EXTRACTED PURE PIXELS USING NO DIVISION SCHEME

EMs	ATGP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	1	0	0	1	1	1	1
Water	0	1	1	1	0	0	0
Soil	1	1	1	1	1	1	1
Road	0	0	0	0	0	0	0
# of extracted pure EMs	2/4	2/4	2/4	3/4	2/4	2/4	2/4
# of extracted pure materials	2	2	2	3	2	2	2

TABLE III. EXTRACTED PURE PIXELS USING 2x2 SCHEME

EMs	ATGP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	3	4	4	4	4	3	3
Water	0	3	3	3	0	0	0
Soil	2	2	2	2	2	2	1
Road	0	0	0	0	0	0	0
# of extracted pure EMs	5/16	9/16	9/16	9/16	6/16	5/16	4/16
# of extracted pure materials	2	3	3	3	2	2	2

TABLE IV. EXTRACTED PURE PIXELS USING 3x3 SCHEME

EMs	ATGP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	8	10	11	9	10	8	9
Water	0	5	4	3	0	0	0
Soil	5	4	5	4	4	4	5
Road	2	1	1	1	1	1	1
# of extracted pure EMs	15/36	20/36	21/36	17/36	15/36	13/36	15/36
# of extracted pure materials	3	4	4	4	3	3	3

TABLE V. EXTRACTED PURE PIXELS USING 4x4 SCHEME

EMs	AT GP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	15	15	16	15	13	13	14
Water	2	10	9	9	2	2	2
Soil	6	4	5	4	4	5	5
Road	3	3	3	3	3	2	3
# of extracted pure EMs	26/6 4	32/ 64	33/ 64	31/ 64	22/ 64	22/ 64	24/ 64
# of extracted pure materials	4	4	4	4	4	4	4

TABLE VI. COMPUTATIONAL TIME FOR ATGP, VCA, & SGA USING DIFFERENT DIVISION SCHEMES (TIME IN SECONDS)

Division Scheme	ATGP	VCA*	SGA*
No Division	0.339	0.148	3.745
Division 2x2	0.311	0.211	3.658
Division 3x3	0.313	0.289	3.522
Division 4x4	0.228	0.364	3.105
* Average computational time of three runs			

B. 2x2 Divisions Scheme Results

Table 3 shows results after applying EEAs using 2x2 Divisions Scheme. ATGP extracted four groups where each group contains four pixels with a total of 16 pixels that should be extracted as endmembers. Five of sixteen were pure pixels which represent only *Tree* and *Soil*, and the other 11 pixels were mixed pixels. ATGP is still not able to find pure pixels for *Water* and *Road* signature. All VCA runs had same results, where they extracted all materials signatures except for *Road* signature. SGA as ATGP couldn't find *Water* and *Road* signatures.

C. 3x3 Divisions Scheme Results

According to 3x3 Divisions Scheme, results extracted after applying EEAs is listed in Table 4. ATGP was able to find two pure pixels for *Road*. Also, results of VCA were improved and all materials signatures are extracted. SGA could extract one pure pixel for *Road* signature as ATGP and continued to be unable to extract any pure pixels for *Water* spectral signature.

D. 4x4 Divisions Scheme Results

Finally, Table 5 shows results for extracted pixels by EEAs using 4x4 Divisions Scheme where this scheme set appropriate conditions for different EEAs to find pure pixels for all materials signatures in dataset.

E. Computation Time

Different division schemes divide dataset into different number of divisions as shown in Table 1, but by increasing the number of divisions, the division size get smaller. This section describes the change in computational time for different used divisions. Table 6 illustrates computational time consumed in seconds for different used EEAs using different division schemes where its content is reflected by Fig. 6.

In ATGP, computational time using NxN Divisions Scheme declines towards increasing N but it's a bit disturbing in VCA, where time slightly increases. It is noticeable that SGA slightly decreases in time consumption by incrementing N. It's worth noting that computational time of ATGP and SGA decline towards more divisions for dataset, but time for VCA slightly increases.

F. No Division Scheme vs. Different Others Schemes from Viewpoint of Extracted p

In the first experiments, No Division scheme used in extracting only 4 endmembers (p = 4), where the expected number of endmembers in dataset is 4 (*Tree*, *Water*, *Soil*, and *Road*). Also each division, in the other division schemes, is used in extracting 4 endmembers.

It is a fair comparison among different division schemes in terms of giving the suitable chance to extract p endmembers from each different division where division is considered as a standalone scene. But it is not a fair comparison in terms of the number of total endmembers extracted that equals to p x N<sup>2</sup> for NxN Divisions Scheme used.

In this experiment, No Division scheme used to extract the same numbers of total extracted endmembers from different other division schemes. According to experiments conducted on whole dataset (No Division Scheme) with p = 16, 36, & 64, we discuss the extracted results and the computational time taken in the following two sections:

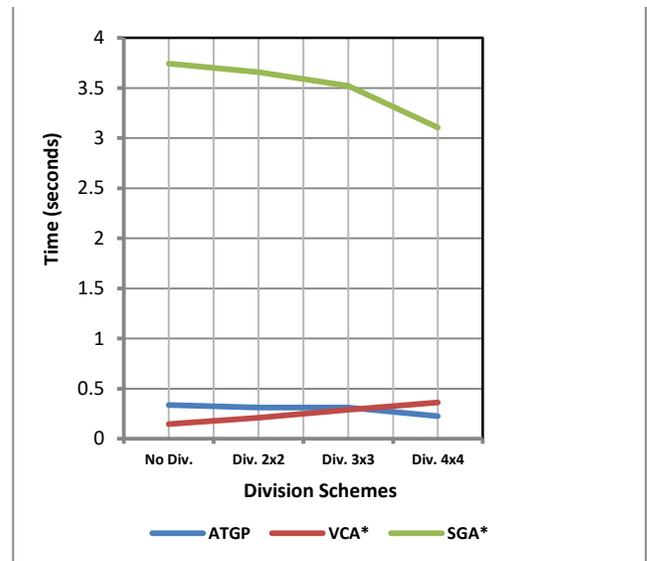


Fig. 6. Computational time consumed for different EEAs using different division schemes. (\* Average computational time of three runs).

1) *Extracted results using different values for p*

Firstly, in Table 7, an experiment conducted on whole Jasper dataset with  $p = 16$ . ATGP and SGA extracted pure pixels for all materials signatures expect for *Water*. Although there are 16 extracted pixels, but 10 of them are mixed pixels. VCA results varied in extracting materials signatures where VCA #1 extracted all materials signatures, VCA #2 extracted all except *Road*, and VCA #3 only extracted *Tree* and *Soil* signatures.

An experiment conducted in Table 8 with  $p = 36$ . ATGP and SGA also (as results using  $p = 16$ ), were agreed on the same extracted results where *Water* wasn't extracted yet, but all other materials signatures were extracted. VCA extracted all materials signatures sometimes including *Water* and sometimes without it.

Finally, using  $p = 64$ , Table 9 lists the results. ATGP and SGA were unable to extract *Water* signature and VCA were able to extract pure pixels for all materials signatures.

2) *Computational time consumed using different values for p*

Computational time taken for experiments conducted for different values of  $p$  ( $p = 16, 36, \& 64$ ) are tabulated in Table 10. By comparing computational time consumed for different EEAs used, we found that VCA increased linearly which considered the least growing algorithm in computational time. ATGP had great increments in time by increasing  $p$ . SGA increases dramatically which indicates the difficulty of its implementations as  $p$  increases. All EEAs increased in computational time using different values for  $p$  without resorting to divide data spatially.

TABLE VII. EXTRACTED PURE PIXELS USING NO DIVISION SCHEME (P=16)

EMs	ATGP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	3	3	3	2	3	3	3
Water	0	1	1	0	0	0	0
Soil	2	1	1	1	2	2	2
Road	1	1	0	0	1	1	1
# of extracted pure EMs	6/16	6/16	5/16	3/16	6/16	6/16	6/16
# of extracted pure materials	3	4	3	2	3	3	3

TABLE VIII. EXTRACTED PURE PIXELS USING NO DIVISION SCHEME (P=36)

EMs	ATGP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	5	7	7	5	5	5	5
Water	0	0	1	1	0	0	0
Soil	4	4	1	2	4	4	4
Road	2	1	3	5	2	2	2
# of extracted pure EMs	11/36	12/36	12/36	13/36	11/36	11/36	11/36
# of extracted pure materials	3	3	4	4	3	3	3

TABLE IX. EXTRACTED PURE PIXELS USING NO DIVISION SCHEME (P=64)

EMs	ATGP	VCA			SGA		
		#1	#2	#3	#1	#2	#3
Tree	8	8	9	10	8	8	8
Water	0	1	1	2	0	0	0
Soil	9	7	5	2	9	9	9
Road	3	5	4	5	3	3	3
# of extracted pure EMs	20/64	21/64	19/64	19/64	20/64	20/64	20/64
# of extracted pure materials	3	4	4	4	3	3	3

TABLE X. COMPUTATIONAL TIME FOR ATGP, VCA, & SGA USING NO DIVISION SCHEMES FOR DIFFERENT VALUES OF P (TIME IN SECONDS)

Value of P	ATGP	VCA*	SGA*
4	0.218	0.143	3.114
16	1.136	0.322	19.501
36	2.729	0.434	84.152
64	4.772	0.607	286.066
* Average computational time of three runs			

As opposed to using different division schemes which showed that computational times consumed were declined as ATGP and SGA or at most slightly increased as VCA. It is noted that division schemes showed superiority in the consumption of less computational time and improve the results of EEAs.

VI. CONCLUSION

Unsupervised hyperspectral unmixing process needs endmember extraction process prior to extract endmembers resident in hyperspectral scene. EEA suffers from finding less distinct and scarce endmembers in the scene. Our proposed method divided dataset into equivalent sections where each section represented as a standalone dataset, and applied EEAs on each section and the extracted endmember sets for the same division scheme were grouped into one overall set. VCA could find pure pixels that represent all materials signatures in smaller homogeneous division, while ATGP and SGA could find them in even smaller divisions.

By comparing all overall sets for different division schemes, we found that dividing data into sections can help EEAs to find rare and less distinct endmembers where computational time consumed decreases as in ATGP and SGA and at most increases slightly as in VCA.

We often need to increase  $p$  value to make EEAs more capable of finding pure pixels in hyperspectral image. But it takes great computational time and doesn't guarantee finding pure pixels that represent all materials signatures in the scene.

Results could be enhanced using different division schemes, not only for enhancement of finding pure pixels, but also in decreasing the computational time consumed.

We divided the data into 4 sections, 9 sections and 16 sections but didn't need more divisions. But how far will we stop the data divisions!! This work can be extended by creating stop condition for more divisions.

#### ACKNOWLEDGMENT

The author would like to thank Marwa Moustafa for her great efforts.

#### REFERENCES

- [1] A. Plaza, P. Martínez, R. Pérez and J. Plaza, "A Quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data", *IEEE Trans. on Geoscience and Remote Sensing*, vol. 42, no 2, pp.650-663, March 2004.
- [2] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery", *IEEE Trans. Aerosp. Electron. Syst.* 39, 1232–1249 (2003).
- [3] C.-I. Chang, C.-C. Wu, W.-M. Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm", *IEEE Trans. Geosci. Remote Sens.* 44, 2804–2819 (2006).
- [4] J. M. P. Nascimento and J. M. Bioucas Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data", *IEEE Trans. Geosci. Remote Sens.* 43, 898–910 (2005).
- [5] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection", *IEEE Trans. Geosci. Remote Sens.* 32, 779–785 (1994).
- [6] Q. Du, N. Raksuntorn, N. Younan and R. King, "End-member extraction for hyperspectral image analysis", *Appl. Opt.*, vol. 47, no. 28, pp. 77-84, 2008.
- [7] C. I. Chang, C. Gao and S. Y. Chen, "Recursive Automatic Target Generation Process in Subpixel Detection", in *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1848-1852, Sept. 2015.
- [8] C. I. Chang, S. Y. Chen, H. C. Li, H. M. Chen and C. H. Wen, "Comparative Study and Analysis Among ATGP, VCA, and SGA for Finding Endmembers in Hyperspectral Imagery", in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4280-4306, Sept. 2016.
- [9] F. Kowkabi, H. Ghassemian and A. Keshavarz, "Enhancing Hyperspectral Endmember Extraction Using Clustering and Oversegmentation-Based Preprocessing", in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 6, pp. 2400-2413, June 2016.
- [10] A. Ertürk, "Enhanced Unmixing-Based Hyperspectral Image Denoising Using Spatial Preprocessing", in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2720-2727, June 2015.
- [11] F. Zhu, Y. Wang, B. Fan, S. Xiang, G. Meng and C. Pan, "Spectral Unmixing via Data-Guided Sparsity", in *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5412-5427, Dec. 2014.
- [12] F. Zhu, Y. Wang, S. Xiang, B. Fan, C. Pan, "Structured sparse method for hyperspectral unmixing", *ISPRS J. Photogram. Remote Sens.*, vol. 88, pp. 101-118, Feb. 2014.

# Real-Time Analysis of Students' Activities on an E-Learning Platform based on Apache Spark

Abdelmajid Chaffai  
RITM LAB, CED ENSEM  
Hassan II University  
Casablanca, Morocco

Larbi Hassouni  
RITM LAB, CED ENSEM  
Hassan II University  
Casablanca, Morocco

Houda Anoun  
RITM LAB CED ENSEM  
Hassan II University  
Casablanca, Morocco

**Abstract**—Real time analytics is the capacity to extract valuable insights from data that comes continuously from activities on the web or network sensors. It is largely used in web based business to drive decisions based on user's experiences, such dynamic pricing and personalized advertising. Many universities have adopted web based learning in their learning process. They use data-mining techniques to better understand students' behavior, and most of the tools developed are based on historical and stored data, and do not allow real time reactivity. Online activities of learners generate at high speed a huge amount of data in form of users' interactions which have all characteristics to be considered as Big data. Deal with volume and velocity of these data in order to inform and enable decisions-makers to act at right time lead us to use new methods to capture E-Learning data, and process it in real time.

This paper focuses on the design and implementation of modern and hybrid real time data pipeline architecture using Apache Flume to collect data, Apache Spark as an unified engine computation for performing analytics on students' activities data and Apache Hive as a data warehouse for storing the processed data and for use by various reporting tools. To conceive this platform we conduct an experiment on Moodle database source.

**Keywords**—Real time analytics; e-learning; big data; Hadoop; spark; Moodle; change data capture; streaming; data visualization clustering

## I. INTRODUCTION

E-Learning is a revolutionary and very promising field that brings about a radical change in the field of learning. Web based technologies are used to create virtual classrooms with attractive materials and resources, and provides a wide range of solutions that support the learning process and services that are accessible anytime from anywhere.

Interactions of students with an E-Learning platform often come in three forms:

- Learner-learner
- Learner-instructor
- Learner-content

Learning Analytics (LA) is a recent field of research and development of tools and technologies that help to analyze and understand the interactions of learners with educational resources. In the first international Conference on Learning Analytics and Knowledge (LAK 2011), it was defined as "the measurement, collection, analysis and reporting of data about

learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" [1]. A related domain is Educational data mining (EDM) which is a data-driven field defined in the community site [2] as "Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings, and using those methods to better understand students, and the settings which they learn in." LA aims to improve E-Learning [3] based on the analysis of the learners' behavior during their interactions with the course.

As more the learning in Higher education are occurring on the web, online activities generate, at high speed, huge amounts of information in the form of users' traces. Deal with volume and velocity of data in order to extract valuable information that can support real-time decision making, lead us to design a modern and flexible architecture that can manage and scale to the continuous stream data.

In this work, we propose a solution for the near-real time analysis of students' activities on a web based learning platform, the most widely used in Moroccan higher education institutions which is Moodle. For this, we have built a system of complete Data Analytics Pipeline which is composed of three main layers. The first layer ensures the data capture from Moodle database. The second layer performs real time processing. The third layer provides a flexible data persistence which can be used by different reporting tools.

All operations are executed in a distributed environment on inexpensive hardware. We use open source technologies such Apache Spark as the main computing engine and Hive to conceive the data warehouse on top of Hadoop cluster.

Analyzing data stream that come continuously from the Moodle platform can greatly help us to track students' progress in courses and detect the students at risk. It can also allow us to monitor the daily health of the E-Learning platform by using fresh reports which can be useful to deduct smart ideas in order to redefine the decisions strategies at right time by adjusting and improving the courses content that respond to students' needs.

The rest of the paper is structured as follows. Section 2 defines big data analytics. Section 3 discusses related work. Section 4 presents the tools used in our work. Section 5 presents the structure of the proposed system and the data processing methodology adopted. Then, it presents and

discusses the experiment results. Finally, Section 6 concludes the paper and describes the future research directions.

## II. BIG DATA ANALYTICS

Big data [4] is a huge amount of data that is generated from various sources. It may be structured when data come from flat files or relational databases, and unstructured or semi structured when data come from the web activities or equipment sensors. Acquiring data very fast does not create the value to the business [5], it needs additional efforts to be meaningful. Big data analytics is the process to apply statistical analysis, data mining, predictive analytics, and text mining on large amount of data using a distributed platform. It depends on speed at which data arrives, and can be divided in two categories [6]:

1) *Batch Processing*: Computation and analysis are applied on data that comes in big batches, then fixed and stored in distributed file system. This type of processing is largely used to learn from historical data by using clustering and classification techniques to create machine learning models which can be applied on new data.

2) *Streaming Processing*: Computation and analysis are applied in real time on recent data that come in continuous records. There are two distinct approaches to analyze live data. The first is to process each record individually, and the second is to split the input data in discretized units called mini-batch according to the interval batch. Stream processing solution must be connected with the source in real time in order to continuously ensure the capturing data.

## III. RELATED WORK

P.K. Udipi et al. [7] proposed a smart learning system model, they describe the possibilities to integrate the E-Learning paradigm with the big data analytics concept and smart utilization. The proposed system contains three layers of different technology framework. The first layer is an E-Learning framework which contains the information and data of user performance evaluation. The second layer is a big data framework which performs a set of different tasks like data extraction, data process and analysis. The third layer is a smart technology framework which enables support of technology need for capturing, predicting, analyzing, decision making and initiates necessary actions as control parameters.

B. Logica et al. [8] lead a study where they discuss the benefits of the use of big data technologies, in order to resolve the problem of managing the massive increase in the produced data volume in educational setting and extracting value from these data to enhance the learning process. They proposed a model for big learning data on cloud architecture based on Hadoop cluster, which can be integrated with the existing Learning Management System (LMS) that the universities usually already own. The different levels of the proposed architecture are designed for collecting any type of data, processing them using Hadoop cluster, performing classification on data stored, and exploring unstructured data using the graphical Gephi tool.

Sunita B Aher et al. [9] proposed a framework for recommendation of courses in E-Learning system Moodle.

They use the enrolled data related to a specific set of courses collected from Moodle database. They use different machine learning algorithms: classification, association rules, and clustering to produce a final model for recommendation. All steps of building the dataflow and model are performed on Weka.

Yassine Tabaa et al. [10] described a learning analytics system for MOOCs based on Hadoop cluster deployed on a private cloud. The main core component of this system is the analytics engine which relies on Map and Reduce model programming, for performing many different analytics jobs, on data that comes from relational database, by using a data integrator based on Apache Sqoop for bulk transfer data from sql sources to HDFS. The analytics platform can help the decision-makers to early identify the students at risk.

San et al. [11] conducted a study in the field of smart grid research. They proposed a complete automation system, where large pool of sensors is embedded in the existing power grids system for controlling and monitoring it by utilizing modern information technologies. Data used in the experiment is in form of times series data available from Texas Synchrophasor Network. The proposed solution uses Apache Kafka to ingest data in real time into the processing layer based on Apache Spark, responsible to perform analytics in fast way. Computation is done in parallel across the cluster of machines.

## IV. TOOLS

### A. Apache Hadoop

Apache Hadoop [12], [13] is an open source framework written in java. It is used to build a cluster for both distributed storage and computation on inexpensive hardware. Hadoop is a master slave architecture which hides technical complexity with high level abstraction in terms of network I/O operations management, fault tolerance and easy horizontal scalability. The main subsystems of Hadoop distributed system are (see Fig. 1):

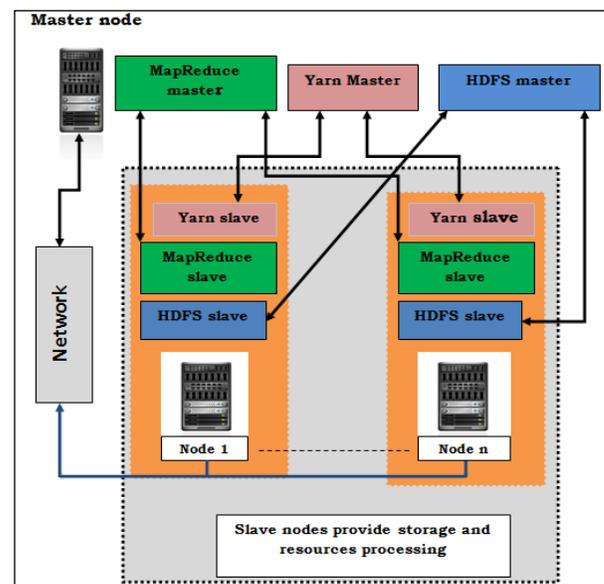


Fig. 1. Hadoop architecture.

- **HDFS:** It is a distributed file system, inspired from the Google file system GFS [14]. When data is stored in HDFS, it is divided into a set of blocks over different nodes of the cluster. The default size of a block is 128 MB.
- **Yarn (Yet Another Resource Negotiator):** It is a distributed resource manager introduced in Hadoop version 2.
- **MapReduce:** It is a distributed and batch-based computing model developed after Google paper on MapReduce [15]. It allows parallelizing the job in small functions map and reducing, and moving the tasks to data locality across a cluster. MapReduce in Hadoop version 2 runs as Yarn application.

### B. Apache Hive

Apache Hive [16], [17] is a data warehouse and an analysis system initially developed at Facebook [18]. It allows query and manage large datasets stored in Hadoop distributed cluster using a language called Hive Query Language (HQL) similar to SQL. Hive converts the queries in one or more MapReduce jobs that are executed on Hadoop cluster and returns the results to the user. Hive stores all metadata in a relational database, and uses by default Derby which is an embedded Java relational database. We used MySQL because Derby cannot be used in a multi-user environment.

### C. Apache Spark

Apache Spark [19], [20] is an open source distributed Framework built in Scala, developed at the University of California Berkeley. It is a Java Virtual Machine designed for fast data processing in the main memory of nodes in the cluster. It can interact with HDFS and Hive and can run as YARN application. The strength of spark resides in its programming model based on high level abstraction of representing a data structure in cluster memory called Resilient Distributed Dataset. RDD [21] is the main component of Spark core. It is resilient because it is capable to rebuild data in case of failures in cluster. RDD is an immutable distributed collection of objects partitioned across different nodes of cluster, and can be created in different ways from external sources or in local and from transformations or actions on existent RDDs. Spark contains several components built on its core like Spark SQL, Spark Streaming, MLlib (Machine Learning library), and GraphX (graph processing), thus it offers to programmers an unified programming [22] platform. This is the main motivation for choosing it to build our system. It allows data sharing between jobs instead of storing intermediate results in the disk compared to MapReduce; and is well suited for iterative operations. Any application submitted to Spark cluster which is master/workers architecture, activates five elements in the following order (see Fig. 2):

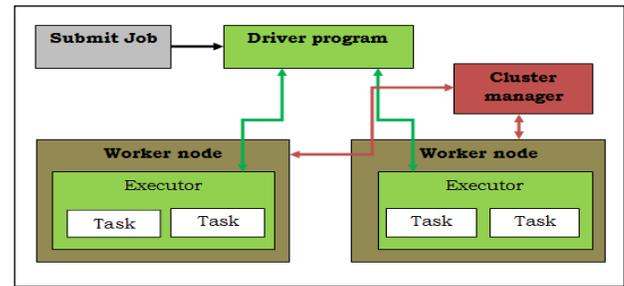


Fig. 2. Spark components architecture.

- 1) **Driver program:** Any program submitted to spark starts with an instantiation of SparkContext object, which is the main entry point to use spark library. SparkContext object use an instance of SparkConf which allows setting parameters and the required resources to run the application.
- 2) **Workers:** Worker is a slave node, which provides resources such computing (CPU), storage, and memory.
- 3) **Cluster manager:** Spark uses a cluster manager to allocate cluster resources for executing a job, and manage the resources across the cluster of worker nodes.
- 4) **Executor:** Each application has its own executors. Executor is a Java virtual machine process which is created on a worker for executing tasks.
- 5) **Task:** This is the smallest unit work of executor that will be sent to one executor which is launched to compute a RDD partition.

Spark streaming library [23], [24] allows consuming live data; it divides the stream in mini batch into time periods equal to batch interval. After every batch it produces a DStream (see Fig. 3), which is a sequence of Resilient Distributed Dataset (RDD). From there, live data can be processed by spark library in the same way like a batch processing.

### D. Apache Flume

Apache Flume [25] is a distributed service designed to ingest the streaming data into Hadoop storage system. The data loading process is triggered by an event using an event driven pipeline architecture based on the principle of the data flow. The Event flows from Source to Channel to Sink (see Fig. 4), orchestrated by a JVM process called “flume agent” responsible to manage the following components:

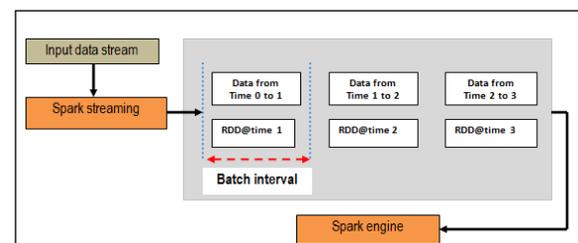


Fig. 3. Discretized stream abstraction.

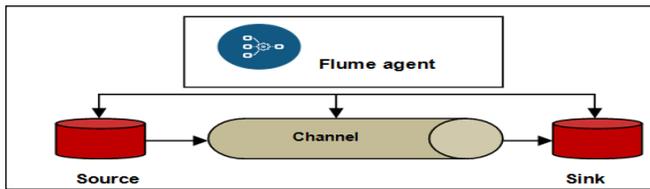


Fig. 4. Dataflow architecture in Apache Flume.

1) *Sources*: It allows connecting to the data sources and collecting the events. There are different types of sources, below we cite the main used in real world applications:

- SpoolingDirectorySource: Retrieves the contents of log files that arrive in a directory.
- ExecSource: Executes a bash command. The most used one is tail command which, when it is executed, retrieves the last line from a log file.
- SyslogSource: Redirects the logs data from a syslog server to Flume.
- AvroSource: Allows setting the flume agent to listen on a TCP port and pump logs in Avro format.

2) *Channel*: It ensures the storage of collected data and the fault tolerance in case of failure of the flume agent. Channel keeps the event until a sink consumes it. Flume provides three types of storage:

- FileChannel: Persists data on a file system.
- MemoryChannel: Stores data in memory for better performance.
- JDBCChannel: Uses a JDBC as persistence solution.

3) *Sink*: Removes and consumes the event from the channel, and moves it to the external destination. Below, we cite two types of sinks:

- AvroSink: Redirects data in Avro format to a distant TCP port.
- HDFSSink: Delivers and writes the events to a local file system.

## V. EXPERIMENT

### A. Description

We have conducted this experiment since October 2016. We started by setting up our system around the Moodle E-Learning platform which is used in HASSAN II University. In collaboration with a team of teachers, we have submitted and published the following courses:

- Object Oriented Programming with java.
- Programming in C ++.
- Programming in Android.
- Software Analysis with UML.

We have authorized access to the courses only to three groups of students. We gave them the following services:

- Consult the course content and download as PDF.
- View the videos.
- Click on web links to references in relation to content.
- Take tests.

Our goal is to build a real time data pipeline system around the existent data source; this system ensures the following tasks:

- Data Integration in two modes offline and online.
- In- memory Data processing.
- Storage of data aggregation and result in a distributed data warehouse in real time. This data warehouse is flexible in order to interact and respond to queries performed by different client applications such the reporting and analysis tools.

### B. Data

Moodle stores its data in the relational database natively Mysql. We used Workbench to visualize the schema of the database which contains about 250 tables. The tables we are interested in are those which contain data profile about students and courses such as mdl\_course, mdl\_user, and those which contain information about interactions with the platform and more specifically with courses such as mdl\_log\_store\_standard, mdl\_lastaccess, mdl\_quizz\_attempts.

### C. Data Integration Methodology

To capture data changes on the Moodle database in order to integrate them in our system processing layer, different approaches are possible depending on the data changes' nature (INSERT or UPDATE) carried out during students' activities on the Moodle platform.

The tables required in our context are divided into two categories:

1) Tables whose content does not change during the web activity, such as mdl\_course, mdl\_user, mdl\_groups, and mdl\_role\_assignments. These tables are used as the reference where we can retrieve the profile information about users and courses, such as username, course name, etc.

We perform the batch replication of these tables to Hive data warehouse by using Apache Sqoop [26] which is an efficient tool designed to transfer bulk data between a relational databases and HDFS. Sqoop allows to extract the content of table using SQL queries, import the updates made in a database and export the result to Hive data warehouse. Several solutions in big data management and analytics use Sqoop as main part of data ingestion.

#### Code Example :

-Create database in Hive called moodle-experiment from hive terminal:

```
Hive> create database moodle_experiment
```

-Bulk transfer data from mdl\_user to our data warehouse

```
moodle_experiment
$ sqoop import --connect jdbc:mysql://URL/moodle_db
--username root -P
--table mdl_user
--hive-import
--hive-table moodle-experiment.users -m 1
```

2) Tables whose content changes during the web activity, such as mdl\_logstore\_standard\_log , mdl\_user\_lastaccess.

The mdl\_log\_store\_standard table is the table where Moodle inserts rows in an incremental way. The other tables undergo changes in their columns like mdl\_user\_lastaccess, to capture data from these tables; we created Mysql triggers on these tables to capture the transactions occurring on them, and incrementally populate new tables we created in Moodle database.

Intercept recent data from different tables via multiple Flume agents generate a lot of streams. To organize the data traffic in subjects and manageable categories, we need a middleware or a central hub capable to interact with Spark and enable real-time data processing. For this, we use Apache Kafka [27] as a pivot point in our system to receive records from Flume and push them into Apache Spark.

Kafka is a distributed persistent subscribe messaging system initially developed at LinkedIn. Kafka stores streams of events in categories called topics. A topic is a logical collection that will receive data from Flume in our context. Kafka uses Zookeeper [28] to manage its components and check the operations status.

We created manually different topics in Kafka cluster using the script Kafka-topics.sh which is a part of Kafka bin files.

Example:

- Creation of a topic named log\_action\_1

```
kafka-topics.sh --create --zookeeper localhost:2181 --
replication-factor 1--partitions 1 --topic log_action_1
```

We use Apache Flume to intercept the latest lines in the tables, in order to interact with both Moodle database and Kafka cluster, by adding to the flume library the following jar files: Flume-ng-sql [29], mysql connector, kafka\_2.11-0.10.0.0.

The created topic log\_action\_1 receives fresh records from Flume via a customized flume-agent configuration file where we set the parameters of source, channel, sink and topic (=log\_action\_1).

Example:

```
#flume-agent configuration file
#channel & source
agent.channels = ch1
agent.sinks = kafkaSink

agent.sources = sql-source
agent.channels.ch1.type = memory
agent.channels.ch1.capacity = 1000000
agent.sources.sql-source.channels = ch1
```

```
agent.sources.sql-source.type =
org.keedio.flume.source.SQLSource

# database
agent.sources.sql-source.connection.url
=jdbc:mysql://URL/moodle_experiment
agent.sources.sql-source.user = root
agent.sources.sql-source.password = password
agent.sources.sql-source.table =
moodl_experiment.mdl_logstore_standard_log
#select colums to intercept
agent.sources.sql-source.columns.to.select =
courseid,userid,action
agent.sources.sql-source.incremental.column.name = id
agent.sources.sql-source.incremental.value = 0
agent.sources.sql-source.run.query.delay=10000
agent.sources.sql-source.status.file.path = /var/lib/flume
agent.sources.sql-source.status.file.name = sql-source.status
agent.sinks.kafkaSink.type=org.apache.flume.sink.kafka.Kafka
Sink
#topic
agent.sinks.kafkaSink.brokerList=master:9093
agent.sinks.kafkaSink.topic= log_action_1
agent.sinks.kafkaSink.channel=ch1
agent.sinks.kafkaSink.batchSize=10
```

#### D. Environment Experiment

We deployed a small local cluster for Hadoop and Spark on 11 nodes running Ubuntu 14.04 LTS and interconnected via one switch of 1Gb/s. The Hadoop cluster is built using Hadoop version 2.7.3. The Spark cluster is built using Spark version 2.0.0. One machine is designed as Master for both Spark and Hadoop, the others nodes are both the Hadoop slaves and Spark workers. The configuration is the same for all nodes:

- Intel(R) Core(TM) i5-3470 CPU 3.20GHz (4CPUs).
- 1Gb/s network connection.
- 300GB hard disk.
- 8GB Memory.

We built the different layers using Java version 8, Scala version 2.11.8, Flume version 1.7.0, Kafka version 2.11-0.10, Hive version 1.7.4, Sqoop version 1.4.6.

As Fig. 5 shows, our real time data pipeline architecture is composed of three main layers:

- *Data capture and integration layer*: Responsible for capturing data change from Moodle and ingesting data in the processing layer using Flume, Sqoop, and Kafka.
- *Data processing layer*: Consumes and processes live data and stores the result continuously in the persistence layer.
- *Persistence layer*: Hosts the data warehouse and responds to queries from different client applications like reporting tool.

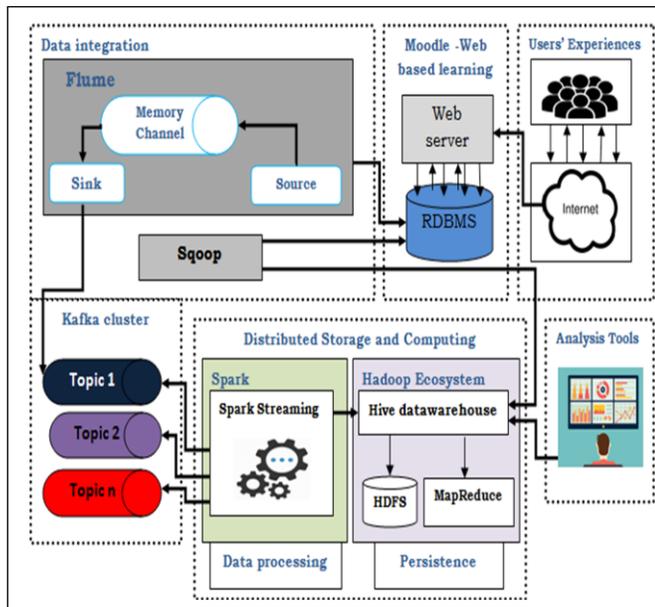


Fig. 5. Real time data pipeline architecture.

### E. Event Processing

An event is the latest row added in Mysql table, intercepted by the corresponding flume agent and respectively stored in Kafka topic. We write different streaming programs in Scala in order to ensure the following tasks:

- 1) Capture the incoming data from Kafka topic and create the DStream via the customized receivers.
- 2) Extract the value from the raw RDDs in the DStream, apply transformations such cleaning, parsing in objects and finally generate the new RDDs.
- 3) Convert the new RDDs to data frame then create a temporary view table to store the new events in a structured format which can be queried.
- 4) Use Spark SQL to extract statistics from the temporary view tables and tables already stored in data warehouse like profile data.
- 5) Persists the result continuously in Hive data warehouse.

KafkaUtils API is used to create the input stream in order to consume data from Kafka topic by using the createDirectStream method. Each event which comes from Flume is a line text that contains the headers data and the data in interest. From DStream we extract the value and clean message with different map operations.

After cleaning the message, we obtain a new RDD in a string comma separated values, then we create a RDD of row objects by inferring schema corresponding to a data type using Scala case class that encapsulates data as objects.

To process each RDD in real time we use foreachRDD method. The following sample code explains briefly the main steps:

**//parameters required to subscribe to a given Topic:  
log\_action1**

```
"key.deserializer" -> classOf[StringDeserializer],  
"value.deserializer" -> classOf[StringDeserializer],  
"group.id" -> "moodle-consumer-group",  
"auto.offset.reset" -> "earliest",  
"enable.auto.commit" -> "true",  
"auto.commit.interval.ms" -> "1000",  
"session.timeout.ms" -> "30000"  
)  
val kafkaTopics = "log_action1"  
val topicsSet = kafkaTopics.split(",").toSet  
//receive events from a Topic in plain text format  
val stream = KafkaUtils.createDirectStream[String,  
String](ssc, PreferConsistent, Subscribe[String,  
String](topicsSet, kafkaParams))  
//Extract the value from a stream and process each RDD  
with foreachRDD  
val lines = stream.map(_._value)  
lines.foreachRDD { rdd =>  
if (!rdd.isEmpty) {  
val sqc = new SQLContext(sc)  
import sqc.implicits._  
  
// Clean Convert RDD[String] to RDD[case class] to  
DataFrame  
val linesDataFrame =  
lines.map(_._replace(",","")).map(_._split(",")).  
map(p => logCaseExemple(p(0).toDouble, p(1).toDouble,  
p(2))) .toDF()  
  
// Creates a temporary view table using the DataFrame  
linesDataFrame.createOrReplaceTempView("view")  
  
//Insert continuous streams into hive table  
sqc.sql("insert into table logm_hive_table select * from  
view")  
  
// select the parsed messages from table using SQL and  
print it  
val linesDFquery = sqc.sql("SELECT courseid,  
count(distinct userid) from view where courseid > 1 group by  
courseid ") } }  
linesDFquery.show()  
  
// Start the computation on data stream  
ssc.start()  
ssc.awaitTermination()
```

### F. Data Visualization

Hive provides a service called hiveserver2 [30] based on Thrift RPC [31], which allows any client like Java, C++, php, and Javascript to interact with its data warehouse.

We build a web application connected to our data warehouse in order to retrieve live result and visualize a dashboard containing a set of indicators as student progress in courses, count course views, active courses, and student performance. Fig. 6 illustrates the visualization in near real time of the data extracted from the table named progress stored in data warehouse.

student	course_name	first_visit	number_visits	time_spent	course_activities	completed_activities	status	date_completed
ansam lahrizia	POO JAVA	10/23/2016	1	00:11:48	10	0	incomplete	
ghina naja	POO JAVA	10/12/2016	6	00:48:20	10	2	incomplete	
maria souda	POO JAVA	11/08/2016	1	00:22:00	10	3	incomplete	
souad yassine	POO JAVA	11/24/2016	42	02:30:00	10	4	incomplete	
amine bouzia	POO JAVA	10/23/2016	20	00:30:22	10	5	incomplete	
lamiae ouadghiri	POO JAVA	10/05/2016	33	01:57:00	10	7	incomplete	
chawki ismael	POO JAVA	11/02/2016	24	03:32:12	10	8	incomplete	

Fig. 6. Learners progress.

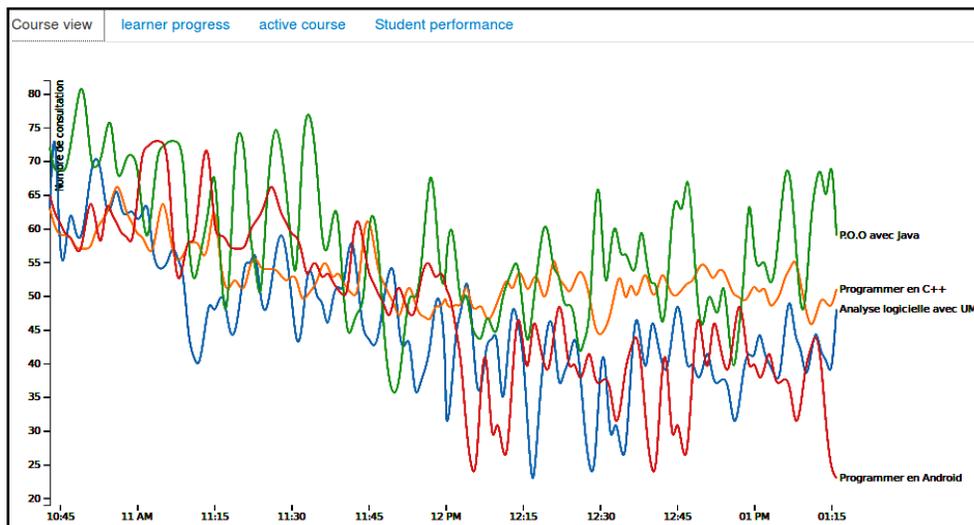


Fig. 7. Real time course view.

It summarizes information about students' progress in each course, such student name, course name, number of visits, the first visit, number of access, total of completed activities, status of progress in activities (completed or incomplete).

The dashboard offers the possibility to apply filters on the result. Fig. 7 illustrates the real time count view in all courses. The data is extracted from the table in data warehouse named count\_vcourse.

### G. Analysis of Students' Behavior using Clustering

Clustering is an unsupervised machine learning technique, used in data exploratory, knowledge discovery and is also the starting point of building a recommender system. Clustering algorithm attempts to find natural groups of similar items in data, and put these data points in the same cluster. Two standard methods are used in clustering [32] hierarchical clustering and partitioning clustering.

K-means is the best known partitioning algorithm and can be described as follows:

1) Choose random k points as initial cluster centers called centroids.

- 2) Assign each data point to their nearest centroid according to the Euclidean distance function.
- 3) Update the centroids for the clusters by calculating the mean value of the points assigned to the cluster.
- 4) Repeat phases 2 and 3 until the centroids do not change or the maximum number of iterations is reached.

A good K-means clustering model will split the objects in clusters by minimizing the total within-cluster variation or total within-cluster sum of square (known as WCSS) defined by the following formula:

$$WCSS(K) = \sum_{k=1}^K \sum_{x_i \in C_k} (x_i - \mu_k)^2$$

Where,  $x_i$  is a data point in cluster  $C_k$  and  $\mu_k$  is the mean value of the points assigned to the cluster  $C_k$ .

The dataset used in this section is extracted from data warehouse and contains 179 observations. Each observation is described by 9 attributes (see Table 1) related to the students' actions in the most active course which is Object Oriented Programming with Java.

Analysis is performed in Rstudio by using a SparkR library [33] which enables large scale data analysis in Spark engine from the R environment.

We did not include the attributes student\_id, student\_name in data preparation so the resulting dataset consists of 7 attributes. We have normalized all numerical data with z-score standardization method, in order to avoid the dominance of some features since they vary in range.

The appropriate cluster number is found as follow:

1) Execute k-means clustering algorithm for different numbers of k from 1 to 15 by using an implementation of k-means algorithm which is included in Spark MLlib (Spark Machine Learning library).

2) Compute the total within-cluster sum of square (WCSS) for each number of cluster and plot the curve of WCSS corresponding to values of k.

3) According to the Elbow method, the curve looks like an arm (see Fig. 8), the location of the “elbow” represents the optimal number of clusters.

As the goal of this analysis is to study the clusters of the students with similar browsing behavior we give in Fig. 9 the coordinates of cluster centroids. Because values are standardized, positive values represent the values that are above the overall mean for all students in dataset, and negative values represent the values that are below the mean.

In Fig. 10 below, the values represent variable means for each cluster in the original metric.

TABLE I. ATTRIBUTES DESCRIPTION OF THE DATASET

Attribute	Description
student_id	Student identifier
studentname	Student name
totalAccess	Number of times the student has visited the course
total_sectionsvisite	Number of views on sections
total_videovisited	Total of viewed videos
total_linksvisited	Total of visited links
time_spentcourse	Time spent in sections in minutes
time_spent_test	Time spent in tests in minutes
avgscore	Average score obtained in all tests (from 0 to 10)

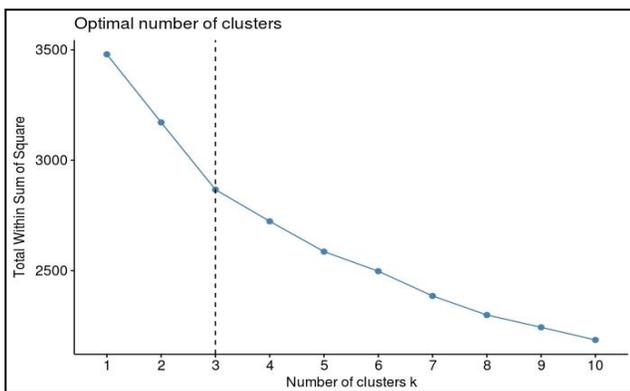


Fig. 8. Find optimal number of clusters with Elbow method.

Examining Fig. 9 and 10, we note that:

- The average of all actions of students in cluster 1 is below the global mean except the score.
- The average of the number of visits to the course, the number of views on sections and the time spent in course by the students in cluster 1 fall between those of the other clusters.
- The students of cluster 1 have consulted less videos and links, and have spent less time doing tests compared to students in other clusters.
- The cluster 2 represents the students at risk.
- The cluster 3 is the group of the average students.

	totalAccess	total_sectionsvisite	total_videovisited	total_linksvisited
1	-0.3392931	-0.04409049	-0.26585464	-0.4084128
2	-0.5894367	0.47547916	0.21567242	0.6305977
3	0.8436063	-0.39347589	0.04358658	-0.2057686
	time_spentcourse	time_spent_test	avgscore	
1	-0.1090387	-0.20270743	1.0606550	
2	-0.3204454	0.09550445	-0.5281638	
3	0.3905054	0.09605830	-0.4767004	

Fig. 9. Coordinates of the cluster centroids

cluster	totalAccess	total_sectionsvisite	total_videovisited	total_linksvisited
1	1 23.23214	63.33929	25.84821	30.84821
2	2 19.95575	79.93805	31.68142	40.30973
3	3 38.72581	52.17742	29.59677	32.69355
	time_spentcourse	time_spent_test	avgscore	
1	180.6518	91.95536	8.750000	
2	167.6549	96.45133	4.991150	
3	211.3629	96.45968	5.112903	

Fig. 10. Variable means in original metric

We can deduce that the cluster 1 is the group of students who have adopted a moderate behavior in all actions and achieved good results.

It is very early to confirm or deny that the strong presence in a platform guarantees a good result; therefore we can't generalize this result. To do this, we need more additional information, so we have to study how to feed our data warehouse with other data that are related to the students' profile, such as the academic past, personal data and other data interactions with the platform which are not available in the database of the E-Learning platform.

## VI. CONCLUSION

In this paper we addressed the challenge to implement an event-driven system around a web based learning platform. This system is in the form of a real time data pipeline capable to capture data change from RDBMS database source and extract valuable information. We adopt a big data concept to design, on inexpensive hardware, a flexible and distributed architecture composed of three layers: data capture, data processing and data persistence. We combine Apache Flume and Sqoop to collect fixed and live data. Apache Kafka is

responsible for organizing the data traffic. To process data in real time we use Spark Streaming library. Apache Hive is used to build our data warehouse hosted in a distributed storage system.

During this work which is based on a real experience we have identified new directions to extend the proposed work. The first is to study and investigate new methods to combine social networks data, past academic and personal data with actual data in data warehouse, to get more information about students. The second is to develop an adaptive learning system based on machine learning models like predictive and recommender system in order to apply these models to assist students during their interactions with the E-Learning platform.

#### REFERENCES

- [1] G. Siemens and D. Gasevic (2012). Guest Editorial –“Learning and Knowledge Analytics”. *Educational Technology & Society*, 15 (3), 1–2.
- [2] EDM: <http://www.educationaldatamining.org/>.
- [3] G. Siemens and R.S.J.d. Baker, “Learning Analytics and Educational Data Mining: Towards Communication and Collaboration”, LAK '12 Proceedings of the 2nd International Conference on Learning Analytics and Knowledge Pages 252-254.
- [4] S. Khan, X. Liu, K. A. Shakil, M. Alam, A survey on scholarly data: From big data perspective, *Information Processing & Management*, Volume 53, Issue 4, 2017, Pages 923-944, ISSN 0306-4573, <http://dx.doi.org/10.1016/j.ipm.2017.03.006>.
- [5] M. Ali-ud-din Khan, M. F. Uddin, N. Gupta, “Seven V’s of Big Data Understanding Big Data to extract Value”, Proceedings of 2014 Zone 1 Conference of the American Society for Engineering Education (ASEE Zone 1).
- [6] Tom White, “Hadoop - The Definitive Guide: Storage and Analysis at Internet Scale (2. ed.)”, Publisher: O’Reilly, ISBN: 978-1-449-38973-4.
- [7] P.K. Udipi, P. Malali and H. Noronha, “Big data integration for transition from e-learning to smart learning framework”, 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, 2016, pp. 1-4. doi: 10.1109/ICBDSC.2016.7460379.
- [8] B. Logica and R. Magdalena, “Using Big Data in the Academic Environment”, *Procedia Economics and Finance*, Volume 33, 2015, Pages 277-286, ISSN 2212-5671, [http://dx.doi.org/10.1016/S2212-5671\(15\)017128](http://dx.doi.org/10.1016/S2212-5671(15)017128). (<http://www.sciencedirect.com/science/article/pii/S2212567115017128>).
- [9] S.B. Aher and Lobo L.M.R.J., “Course Recommender System in E-learning”, *International Journal of Computer Science and Communication*, Vol. 3, No. 1, January-June 2012, pp. 159-164.
- [10] Y. Tabaa and A. Medouri, “LASyM: A Learning Analytics System for MOOCs”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 5, 2013.
- [11] Shyam R., Bharathi Ganesh HB, Sachin Kumar S, Prabaharan Poornachandran and Soman K P, “Apache Spark a Big Data Analytics Platform for Smart Grid”, *Procedia Technology* 21 (2015): 171-178. DOI: 10.1016/j.protey.2015.10.085.
- [12] Han Hu, Yonggang Wen, Tat-Seng Chua and Xuelong Li, “Toward Scalable Systems for Big Data Analytics: A Technology Tutorial”, *IEEE Access*, Vol. 2, 2014, pp.652-687, DOI: 10.1109/ACCESS.2014.2332453.
- [13] Apache Hadoop: <http://hadoop.apache.org/>.
- [14] S. Ghemawat, H. Gobioff and Shun-Tak Leung, “The Google File System”, *ACM SIGOPS operating systems review*, vol. 37, no. 5.p. 29-43. ACM, 2003.
- [15] J. Dean and S. Ghemawat, “MapReduce: Simplified Data Processing on Large Clusters”, In Proceedings of the 6th USENIX OSDI, pages 137–150,2004.
- [16] Apache Hive: <https://hive.apache.org/>.
- [17] G.P. Haryono and Y. Zhou, “Profiling apache HIVE query from run time logs”, *Big Data and Smart Computing (BigComp)*, 2016,IEEE DOI: 10.1109/BIGCOMP.2016.7425802.
- [18] A. Thusoo, J.S. Sarma, N. Jain, Z. Shao, P. Chakka, N. Zhang, S. Antony, H. Liu and R. Murth, “Hive - a petabyte scale data warehouse using Hadoop”, In Proceedings of International Conference on Data Engineering (ICDE), 2010, pp. 996- 1005.
- [19] Apache Spark: <http://spark.apache.org/>.
- [20] A. Shkapsky , M. Yang, M. Interlandi, H. Chiu, T. Condie and C. Zaniolo, “Big Data Analytics with Datalog Queries on Spark”, SIGMOD’16, June 26-July 01, 2016, San Francisco, CA, USA,2016 ACM, DOI: <http://dx.doi.org/10.1145/2882903.2915229>.
- [21] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, I. Stoica, “Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing”, in Proc. 9th USENIX Conf. Netw. Syst. Des. Implement., 2012, p. 2.
- [22] M. Zaharia, R.S. Xin, P. Wendell, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M.J. Franklin, A. Ghodsi, J. Gonzalez, S. Shenker, I. Stoica, “Apache Spark: a unified engine for big data processing”, *Magazine Communications of the ACM*, Vol. 59 No. 11, Pages 56-65.
- [23] Spark streaming: <http://spark.apache.org/streaming/>.
- [24] W. Wingerath, F. Gessert, S. Friedrich, and Norbert Ritter, “Real-time stream processing for Big Data”, *IT – Information Technology* 2016; 58(4): 186–194,DOI 10.1515/itit-2016-0002.
- [25] Deepak Vohra, “Practical Hadoop Ecosystem: A Definitive Guide to Hadoop-Related Frameworks and Tools”, Publisher: Apress, ISBN: 978-1-4842-2199-0, pp.287-300.
- [26] Apache Sqoop: <http://sqoop.apache.org/>.
- [27] Apache Kafka: <https://kafka.apache.org/>.
- [28] Apache Zookeeper: <https://zookeeper.apache.org/>.
- [29] Flume-ng-sql: <https://mvnrepository.com/artifact/org.keedio.flume.flume-ng-sources/flume-ng-sql-source/1.3.7>.
- [30] Hiveserver2: <https://cwiki.apache.org/confluence/display/Hive/HiveServer2+Clients>.
- [31] Apache Thrift: <http://thrift.apache.org/docs/concepts>.
- [32] Tuffery Stéphane, “Data Mining et statistique décisionnelle L’intelligence des données- 4 ème édition”, Publisher: Technip, ISBN: 9782710810179, pp. 265-317.
- [33] S. Venkataraman, Z. Yang, D. Liu, E. Liang, H. Falaki, X. Meng, R. Xin, A. Ghodsi, M. Franklin, I. Stoica, M. Zaharia, “SparkR: Scaling R Programs with Spark”, SIGMOD’16, June 26–July 1, 2016, San Francisco, CA, USA. ACM, 2016.

# Reducing Dimensionality in Text Mining using Conjugate Gradients and Hybrid Cholesky Decomposition

Jasem M. Alostad

The Public Authority for Applied Education and Training (PAAET), College of Basic Education  
P.O.Box 23167, Safat 13092, Kuwait

**Abstract**—Generally, data mining in larger datasets consists of certain limitations in identifying the relevant datasets for the given queries. The limitations include: lack of interaction in the required objective space, inability to handle the data sets or discrete variables in datasets, especially in the presence of missing variables and inability to classify the records as per the given query, and finally poor generation of explicit knowledge for a query increases the dimensionality of the data. Hence, this paper aims at resolving the problems with increasing data dimensionality in datasets using modified non-integer matrix factorization (NMF). Further, the increased dimensionality arising due to non-orthogonality of NMF is resolved with Cholesky decomposition (cdNMF). Initially, the structuring of datasets is carried out to form a well-defined geometric structure. Further, the complex conjugate values are extracted and conjugate gradient algorithm is applied to reduce the sparse matrix from the data vector. The cdNMF is used to extract the feature vector from the dataset and the data vector is linearly mapped from upper triangular matrix obtained from the Cholesky decomposition. The experiment is validated against accuracy and normalized mutual information (NMI) metrics over three text databases of varied patterns. Further, the results prove that the proposed technique fits well with larger instances in finding the documents as per the query, than NMF, neighborhood preserving; nonnegative matrix factorization (NPNMF), multiple manifolds non-negative matrix factorization (MMNMF), robust non-negative matrix factorization (RNMF), graph regularized non-negative matrix factorization (GNMF), hierarchical non-negative matrix factorization (HNMF) and cdNMF.

**Keywords**—Data mining; non-integer matrix factorization; Cholesky decomposition; conjugate gradient algorithm

## I. INTRODUCTION

Computing application in several fields generates numerous data over several instances. In order to extract knowledge from such instances, solutions are used conventionally with data mining tools. However, the large datasets with numerous instances poses severe challenges and that leads to improper processing of such huge data volume. The reduction of datasets or improved mining algorithm can overcome such challenges [1]. The reduction of improper values from the datasets provides a greater impact and this increases the performance of processing the large data [2]; hence, the improved mining approach is not useful in some cases [3].

The data reduction is the process of reducing the size or dimensionality of the data, however, the representation of the data should be retained. Selection of instance is one better way to reduce the data by reducing the total number of instances. In spite of many efforts to deal with such instances, data mining algorithm, however undergoes severe challenges due to non-applicability of datasets with large instances. Hence, the computational complexity of the system increases with larger instances [3], [4] and leads to problems in scaling, increased storage requirements and clustering accuracy. The other problems associated with larger data instances include, improper association or interaction in the feature space, lack of ability to handle the large datasets with discrete variables, inability to classify the data and poor knowledge generation for a given query, and finally poor computation due to missing variables.

Recently, there are significant developments in NMF for various clustering problems in data mining, defined above. The NMF process is used to factorize the input matrix into two matrices of non-negative variables in a lower rank order [5]-[8]. Several applications of NMF include: chemometrics, environmetrics, pattern recognition, text mining and summarization [9], multimedia data analysis [10], analysis of DNA gene expression [11], analyzing the financial data [12], and social network analysis [13]. Several algorithms are designed to overcome the problems associated with objective functions [14], classification [15], collaborative filtering [16] and computational methodologies.

Thought, NMF is used for data analysis, the recent trends has improved the discoverability and learning ability of NMF in the data mining to solve the problems associated with larger datasets. In order to avoid limitations associated with larger dataset, the following consideration are made in the present study: This proposed method uses NMF to study the feature vector of a text document and Cholesky decomposition is used to avoid the non-orthogonality problem in the NMF. Further, to avoid poor decomposition using Cholesky decomposition, conjugate gradient is used, which avoids the rapid multiplication by the gradients in the feature space.

Since, the NMF algorithm learns both the data and feature vector in the dataset feature space, the proposed method implies the following contributions:

- First, the metric using data matrix is estimated in the feature space using trained feature vector.
- Second, the Cholesky decomposition process is applied over the metric and the upper triangular matrix is identified.
- Third, upper triangular matrix is used a linear mapper for the associated data vector.
- Finally, conjugate gradient is applied to reduce sparse matrix through reduced multiplication and that avoids the NP-hard problem which significantly reduces the computational complexity [17], [18].

The outline of the paper is presented as follows: Section 2 discusses the related works. Section 3 provides the NMF model for clustering the larger datasets. Section 4 provides the modifications in NMF using Cholesky decomposition. Section 5 provides experimental verification of the proposed system over the associated datasets and section 6 concludes the paper with future work.

## II. RELATED WORKS

Several methods of NMF are discussed here, which include: Semi supervised constrained NMF [19], semi-supervised graph based discriminative NMF [20], Bayesian learning approach to reduce the generalization error in upper bound using NMF [21] and update rules [22], sparseness NMF, which provides better characterization of the features [23], sparse unmixing NMF [24], locally weighted sparse graph regularized NMF [25], graph-regularized NMF [26], graph dual regularization [27], multiple graph regularized NMF [28], graph regularized multilayer NMF [29], adaptive graph regularized NMF [30], hyper-graph regularized [31], graph regularization with sparse NMF [32], multi-view NMF [33], extended incremental NMF [34], incremental orthogonal projective NMF [35], correntropy induced metric NMF [36], multi-view NMF [37], patch based NMF [38], MMNMF [39], regularized NMF [40], FR conjugate gradient NMF [41]. However, these methods failed to address the problems associated with non-orthogonality due to the presence of non-negative elements in NMF.

## III. IMPROVED NMF METHOD

NMF is a non-negative low-rank approximation method associated with certain constraints that relates to the non-negative elements in the data and feature vectors. Here, non-orthogonality problem exist due to the presence of non-negative elements between the vectors and addition of linear combination, results in part representations. This interpretable and intuitive method for representing the text data elements is divided into two parts:

- 1) Data vector representation using Cholesky Decomposition (CD) with Conjugate Gradients (CG).
- 2) Feature vector representation using NMF.

The detail of these is shown in following sections:

### A. Fitness Function for NMF

In NMF, it is assumed that matrices contain non-negative elements, hence, factorization is approximated. Let input data matrix is  $X = (x_1, x_2, \dots, x_n)$ , which carries  $n$  elements of input data vectors and the data matrix is decomposed into two matrices,

$$X \approx FG^T \quad (1)$$

Where,  $\mathbf{X} \in \mathbb{R}^{p \times n}$ ,  $\mathbf{G} \in \mathbb{R}^{n \times k}$  and  $\mathbf{F} \in \mathbb{R}^{p \times k}$  and  $\mathbb{R}$  is the set of non-zero real numbers,  $\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n)$  and  $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n)$ . In general, the value of  $p < n$  and the rank of  $\mathbf{F}$  and  $\mathbf{G}$  matrices is less than  $X$  i.e.  $k \ll \min(p, n)$ . The rank  $\mathbf{F}$  and  $\mathbf{G}$  is generated using minimization fitness function and the sum of squared errors is used to evaluate the fitness function, which is represented as:

$$\min_{F, G \geq 0} J_{sse} = \|X - FG^T\|^2 \quad (2)$$

The matrix normalization is obtained using Frobenius norm and the values of  $\mathbf{F}$  and  $\mathbf{G}$  are non-negative with non-orthogonal column vectors in its Euclidean space. The non-deficiency cases for rank  $\mathbf{R}$  and  $\mathbf{G}$  is generated using **I-divergence** fitness function:

$$\min_{F, G \geq 0} J_{ID} = \sum_{i=1}^m \sum_{j=1}^n \left[ X_{ij} \log \frac{X_{ij}}{(FG^T)_{ij}} - X_{ij} + (FG^T)_{ij} \right] \quad (3)$$

Here, when  $I(x) = x \log x - x + 1 \geq 0$ , inequality holds i.e.  $x \geq 0$  and when  $x = 1$ , the equality holds. Hence, **I-divergence** using inequality condition is expressed as:  $I(u, v) = (u/v) \log(u/v) - u/v + 1$ .

### B. NMF Clustering

The initialization in NMF is an important process with clustering, similar to k-means clustering. However, the fitness function as a minimization function often undergoes local minimum problem [42], [43]. Due to such constraint, even if the minimization function is convex, the intrinsic alternating function is non-convex. If a random initialization is considered, then the factor matrices are initialized as random matrices and hence, it become ineffective due to slow convergence to attain the local minima. If clustering process is used in NMF, the initialization is obtained from fuzzy [44], divergence-k-means [45] and spherical k-means [46], [47]. However, the proposed method considers a simple strategy for document clustering, which is discussed below:

NMF is applied to cluster the documents and number of features vectors in the document of each dataset is set as total clusters in a document [48], [49]. Each cluster is assigned with individual instances and the representation  $\mathbf{g}$  is considered maximum, which is represented as:

$$c_g = \arg \max_c g_c \quad (4)$$

Where,  $\mathbf{g}_c$  is considered as the  $c$  th element in  $\mathbf{g}$ .

### C. NMF Representation Learning

The representative learning,  $\mathbf{G}$  is carried out by many supervised or unsupervised method using NMF, since it

reduces the dimensionality in an effective manner. Certain other techniques uses Euclidean space to conduct learning on  $\mathbf{G}$  [50]. However, the non-orthogonality problem during the representative learning process is not dealt and hence, the proposed study uses such problem to reduce the dimensionality in large datasets.

#### IV. CHOLESKY DECOMPOSITION

The main reasons for the non-orthogonality problem during the representative learning ( $\mathbf{G}$ ) is the formulation of distance (squared) between the paired instances ( $g_i, g_j$ ) as  $(g_i - g_j)^T (g_i - g_j)$ . The squared distance is implicitly assumes that  $g_i$  lies in Euclidean space. In general, the learning of  $(f_1, \dots, f_q)$  using NMF are considered non-orthogonal to each other and the use of squared Euclidean distance is not appropriate during the representative learning by  $\mathbf{G}$ . To solve this, generalized squared distance metric using Mahalanobis distance ( $\mathbf{M}$ ) measurement is used to solve the non-orthogonality of feature vector, which is represented as:

$$(g_i - g_j)^T \mathbf{M} (g_i - g_j) \quad (5)$$

The NMF property is exploited to decompose the data matrix,  $\mathbf{X}$  into,

- a)  $\mathbf{F}$  with column vectors  $(f_1, f_2, \dots, f_n)$  spans the feature space of the matrices, and
- b)  $\mathbf{G}$  provides the feature space representation.

With such decomposition property, the cdNMF,

- 1) Initially, the estimation of the NMF metric is carried out in feature space using the feature vectors (trained).
- 2) Then, the Cholesky decomposition is applied over the NMF metric, which finds the upper triangular matrix.
- 3) Finally, upper triangular matrix is used to map linearly the data vectors.

##### A. NMF Metric Estimation

In NMF, the data matrix ( $\mathbf{X}$ ) is approximated and it is represented in the feature space as  $\mathbf{G}$  and the feature representation in the data space is  $\mathbf{F}$ . The normalization [8] of  $f$  results in  $f^T f = 1$  as the metric  $\mathbf{M}$  is estimated as gram matrix  $\mathbf{F}\mathbf{G}\mathbf{F}$  of the feature vector.

$$\mathbf{M} = \mathbf{F}^T \mathbf{F}, \text{ s.t. } u_l^T u_l = 1, \forall l = 1, \dots, q \quad (6)$$

The metric estimation do not use label information for estimating  $\mathbf{M}$  and the data vector is approximated over the feature space through  $u_1, \dots, u_q$  and it is seen that  $\mathbf{M} = \mathbf{F}^T \mathbf{F}$  can be used to estimate the feature space metric.

##### B. Cholesky Decomposition over NMF Metric

The estimation of metric using (5) guarantees  $\mathbf{M}$  as symmetric positive semi-definite matrix. Linear algebra guarantees  $\mathbf{M}$ , which decomposes the upper triangular matrix  $\mathbf{T}$  using Cholesky decomposition:

$$\mathbf{M} = \mathbf{T}^T \mathbf{T} \quad (7)$$

By substituting (7) into (5), the Cholesky function to represent the upper triangular matrix  $\mathbf{T}$  is given as:

$$\mathbf{G} \rightarrow \mathbf{T}\mathbf{G} \quad (8)$$

##### C. Conjugate Gradients (CG)

Assuming, upper triangular elements to be sparse, hence linear representation of the data vectors is not considered valid. The use of CG for removing the sparse value in the matrices is found with the set of linear equations. The CG is applied on upper triangular matrix to remove the sparse value. The proposed method is used to utilize the trained representation using cdNMF, without any modifications in the algorithms over the learning representation. The elimination of sparse matrix is avoided by eliminating the rapid multiplication and clustering such data leads to increased convergence rate with faster association of elements in the dataset. Here,  $\mathbf{M} = (\mathbf{T}\mathbf{G}^T \mathbf{T}\mathbf{G})^{-1}$  is the pre-conditioner to enhance the multiplication process, in case of incomplete Cholesky decomposition, where  $\mathbf{M} = \mathbf{T}\mathbf{G}^T \mathbf{T}\mathbf{G}$  defines the incomplete Cholesky decomposition.

##### Algorithm 1 cdNMF

**cdNMF** ( $\mathbf{X}$ , NMF,  $q$ , parameters)

- 1: **Find**  $\mathbf{X} \in \mathbb{R}^{p \times n}$ , NMF,  $q$  and parameter(NMF)
- 2:  $\mathbf{F}, \mathbf{G} :=$  run NMF on  $\mathbf{X}$  with parameter and  $q$  // metric estimation
- 3:  $\mathbf{M} := \mathbf{F}^T \mathbf{F}$
- 4:  $\mathbf{T} := \text{CD}(\mathbf{M})$  s.t.  $\mathbf{M} = \mathbf{T}^T \mathbf{T}$
- 5: **Apply** CF once the linear coordinates changes,  $x = \mathbf{T}\mathbf{G}y$  and  $\det \mathbf{T}\mathbf{G} \neq 0$
- 6: **Use** CG for solving  $\mathbf{T}\mathbf{G}^T \mathbf{A}\mathbf{T}\mathbf{G}y = \mathbf{T}\mathbf{G}^T b$
- 7: **Set**  $x' = \mathbf{T}\mathbf{G}^{-1}y'$
- 8: **Set** the preconditioner  $\mathbf{M} = (\mathbf{T}\mathbf{G}^T \mathbf{T}\mathbf{G})^{-1}$
- 9: **Multiply**  $\mathbf{T}\mathbf{G}$  by  $\mathbf{T}\mathbf{G}^{-1}$
- 10: **Compute**  $x' = \mathbf{T}\mathbf{G}^{-1}y'$
- 11: **Return**  $\mathbf{M}, x', \mathbf{T}\mathbf{G}$

This algorithm helps in reducing increased multiplication process and increases the convergence rate. The computation of  $x' = \mathbf{T}\mathbf{G}^{-1}y'$  is carried out only at the end of multiplying  $\mathbf{T}\mathbf{G}$  by  $\mathbf{T}\mathbf{G}^{-1}$  and the computation process is multiplied with  $\mathbf{M}$ .

#### V. EXPERIMENTAL RESULTS

In the proposed system, the cdNMF is used to cluster the documents and compared with other algorithms to prove its effectiveness. The cdNMF system for evaluating the datasets is compared with conventional algorithms and that include: NMF [51], GNMF [5], NPNMF [6], MMNMF [7] and RNMF [8].

##### A. Text Mining Datasets

The proposed cdMNF with conjugate gradient is evaluated on text datasets: 20 Newsgroups data (Table 1), Reuters 21578 data (Table 2) and R52 data (Table 3).

TABLE I. ATTRIBUTES OF NEWSGROUPS DATA

Class	Number of train documents	Number of test documents	Total Number of documents
soc.religion.christian	598	398	996
talk.politics.guns	545	364	909
talk.politics.mideast	564	376	940
talk.politics.misc	465	310	775
talk.religion.misc	377	251	628

TABLE II. ATTRIBUTES OF REUTERS 21578 DATA

Reuters 21578				
Topics	Number of training documents	Number of test documents	Number of other documents	Total documents
0	1828	280	8103	10211
1	6552	2581	361	9494
2	890	309	135	1334
3	191	64	55	310
4	62	32	10	104

TABLE III. ATTRIBUTES OF R52 DATA

R52			
Class	Number of training documents	Number of test documents	Total documents
Crude	253	121	374
Earn	2840	1083	3923
Interest	190	81	271
money-supply	123	28	151
Trade	251	75	326

Each document is represented as standard vector model [1] that contains occurrence of classes and terms in a document. Each document is represented as single line in the file and represented using a word or document class with TAB character, delimiting spaces and the terms. A total of 5 classes are used from each dataset with a set of training documents, test documents and other documents. A cluster is created with 5 classes of 20 Newsgroups data, Reuters 21578 data and R52 data.

Hence, three clusters are used in this study that includes a set of 4248, 21453 and 5045 documents, respectively for 20 Newsgroups, Reuters 21578 data and R52 data. The clusters with sub-clusters are classes are used to create the samples and a total of 100 documents from each sub-clusters of all the classes form the sample. Likewise, 20 such samples are created from the text datasets.

Here, the each text sample is conducted with pre-processing operations that include: trunc5 stemmer [52] and POS Tagger [53] and removal of stop words and finally it selects a total of 30000 words with mutual information in a larger perspective. The selection of sub-clusters for the sample formation is shown in Table 4.

### B. Clustering Metrics

The performance of the clustering metrics is evaluated using two metrics Accuracy (*acc*) and Normalized Mutual Information (*NMI*). The parameter, *acc* is used to estimate the overall performance of the cluster, which is defined in the form of a fraction metric,  $acc = \eta_t / \eta_{ov}$ , where,  $\eta_t$  is the correctly clustered documents sample and  $\eta_{ov}$  is the overall amount of samples. The Mutual Information (MI) finds the interdependency between the variables and if the text variables are equal, then MI is zero and it is defined as:

$$MI(x, y) = \sum_{\hat{x} \in \mathcal{X}} \sum_{\hat{y} \in \mathcal{Y}} p(x, y) \log \left( \frac{p(\hat{x}, \hat{y})}{p(\hat{x})p(\hat{y})} \right)$$

TABLE IV. SELECTION OF 20 SAMPLES FROM THE DATASETS

Samples	20 News Group	Reuters 21578	R52
#1	7	6	7
#2	6	7	7
#3	7	7	6
#4	8	7	5
#5	8	5	7
#6	7	5	8
#7	7	8	5
#8	5	8	7
#9	5	5	10
#10	5	10	5
#11	10	5	5
#12	10	0	10
#13	0	10	10
#14	10	10	0
#15	15	0	5
#16	0	15	5
#17	15	5	0
#18	0	0	20
#19	0	20	0
#20	20	0	0

Where,  $p(\hat{x}, \hat{y})$  is the joint Probability Distribution Function (PDF) of  $x$  and  $y$ ,  $p(\hat{x})$  and  $p(\hat{y})$  are the marginal PDF of  $x$  and  $y$ . The MI provides information related to the amount of uncertainty measured between documents  $x$  and  $y$  and one documents reduces the uncertainty of the other documents. Entire information is shared between the documents if the value of MI is zero and vice versa. The Normalized MI or NMI is denoted as:

$$NMI(x, y) = \frac{MI(x, y)}{\max(E(x), E(y))}$$

Where,  $E(x)$  and  $E(y)$  are the entropy of the document  $x$  and  $y$ .

### C. Evaluation and Comparisons

NMF is the baseline algorithm, GNMF uses KNN graph with regularization term for preserving the structure of geometry, NPNMF uses local linear embedding and graph approach in NMF uses trained regularization term, MMNMF uses an eleven graph for exploring the multiple manifold data structure, RNMF adds noise in NMF and HNMF encodes the geometry into matrix factorization using hyper graph. These systems are tested against accuracy and Normalized Mutual Information (NMI) over the sample datasets.

The results of *acc* from Table 5 show that the cdNMF performs well than conventional schemes. Here, the performance of cdNMF increases gradually from samples 1 to 20. It is inferred that if the documents of similar dataset are more, the accuracy is more and it reduces when the 20 sample documents are equally distributed from similar clusters. The overall accuracy of cdNMF is slightly higher than HNMF and

GNMF but it has higher *acc* rate than MM-NMF, NPNMF and NMF. This is the same for NMI, shown in Table 6.

TABLE V. ACCURACY OF DIFFERENT METHOD OF 20 SAMPLE DATA SETS

Samples	NMF	NPNMF	MM-NMF	RNMF	GNMF	HNMF	cdNMF
#1	0.4582	0.4801	0.4681	0.5142	0.6925	0.6841	0.7315
#2	0.4648	0.4872	0.5015	0.5345	0.7124	0.7241	0.7452
#3	0.4677	0.4877	0.5248	0.5913	0.7182	0.7518	0.7824
#4	0.4678	0.4908	0.5312	0.6011	0.7225	0.7622	0.7895
#5	0.4755	0.4912	0.5324	0.6025	0.7315	0.7693	0.7917
#6	0.4757	0.4979	0.5399	0.6104	0.7318	0.7721	0.7924
#7	0.5084	0.5012	0.5401	0.6201	0.7324	0.7741	0.7954
#8	0.5134	0.5032	0.5428	0.6215	0.7328	0.7779	0.8001
#9	0.5156	0.5089	0.5542	0.6225	0.7355	0.7815	0.8017
#10	0.5228	0.5206	0.5581	0.6241	0.7419	0.7826	0.8156
#11	0.5352	0.5273	0.5595	0.6255	0.7421	0.7892	0.8249
#12	0.5384	0.5478	0.5716	0.6258	0.7552	0.7912	0.8456
#13	0.5432	0.5501	0.5792	0.6293	0.7621	0.7924	0.8592
#14	0.5488	0.5591	0.5932	0.6309	0.7624	0.7927	0.8604
#15	0.5583	0.5858	0.6083	0.6385	0.7815	0.7935	0.8621
#16	0.5881	0.6078	0.6145	0.6515	0.7858	0.8019	0.8665
#17	0.5892	0.6378	0.6489	0.6922	0.8009	0.8245	0.8912
#18	0.5995	0.6584	0.6692	0.6945	0.8012	0.834	0.8942
#19	0.6087	0.6697	0.6745	0.7152	0.8241	0.8402	0.9018
#20	0.6724	0.6795	0.6845	0.7564	0.8512	0.8576	0.9156
<b>Average</b>	<b>0.5326</b>	<b>0.5496</b>	<b>0.5748</b>	<b>0.6301</b>	<b>0.7559</b>	<b>0.7848</b>	<b>0.8284</b>

TABLE VI. NMI OF DIFFERENT METHOD OF 20 SAMPLE DATA SETS

Samples	NMF	NPNMF	MM-NMF	RNMF	GNMF	HNMF	cdNMF
#1	0.4789	0.5203	0.6113	0.753	0.7433	0.7854	0.7754
#2	0.4825	0.5504	0.6384	0.7586	0.7542	0.7884	0.7951
#3	0.5231	0.5648	0.6569	0.7769	0.7823	0.7885	0.7952
#4	0.5321	0.6212	0.6805	0.7872	0.7911	0.7946	0.8025
#5	0.5549	0.633	0.69	0.7875	0.8009	0.7952	0.8107
#6	0.6122	0.6749	0.7203	0.8124	0.8014	0.8095	0.812
#7	0.6199	0.6825	0.7254	0.8147	0.8306	0.8149	0.833
#8	0.6317	0.6842	0.7274	0.8164	0.8325	0.8412	0.8342
#9	0.6823	0.6924	0.7373	0.822	0.8336	0.8423	0.8466
#10	0.6842	0.7043	0.7504	0.8245	0.8365	0.8435	0.8587
#11	0.6902	0.7047	0.7582	0.8342	0.8424	0.8502	0.8629
#12	0.6948	0.7063	0.7592	0.84	0.8489	0.8567	0.8743
#13	0.7001	0.7137	0.7687	0.8412	0.8598	0.8676	0.8822
#14	0.7042	0.7158	0.7691	0.8489	0.8674	0.8781	0.8918
#15	0.7094	0.724	0.77	0.8695	0.8858	0.9001	0.8939
#16	0.7142	0.7288	0.7742	0.8731	0.8904	0.9079	0.9008
#17	0.7355	0.743	0.7745	0.9088	0.8998	0.9139	0.9153
#18	0.7412	0.7436	0.7769	0.9138	0.9335	0.9418	0.9441
#19	0.7419	0.7477	0.7787	0.9205	0.9356	0.9504	0.9514
#20	0.7496	0.765	0.8057	0.9309	0.9392	0.9505	0.9515
<b>Average</b>	<b>0.6491</b>	<b>0.6810</b>	<b>0.7337</b>	<b>0.8367</b>	<b>0.8454</b>	<b>0.8560</b>	<b>0.8616</b>

The average values of NMI results claim that the interdependence of documents belonging to similar cluster during testing is also high. It is seen further the documents are

equal in number in sample dataset and the interdependency is less for other conventional algorithms, however, cdNMF performs well.

TABLE VII. AVERAGE ACCURACY OF PROPOSED METHOD VS. EXISTING METHOD USING THREE DATA SETS

Cluster	NMF	NPNMF	MM-NMF	RNMF	GNMF	HNMF	cdNMF
20 News Group	0.8945	0.9342	0.8731	0.9261	0.9779	0.9822	0.9932
Reuters 21578	0.8591	0.8998	0.8779	0.8595	0.9286	0.9486	0.9524
R52	0.9023	0.9201	0.8769	0.8315	0.9599	0.9738	0.9869

TABLE VIII. AVERAGE NMI OF PROPOSED METHOD VS. EXISTING METHOD USING THREE DATA SETS

Cluster	NMF	NPNMF	MM-NMF	RNMF	GNMF	HNMF	cdNMF
20 News Group	0.7602	0.6216	0.749	0.9382	0.9412	0.9551	0.9592
Reuters 21578	0.7605	0.7281	0.709	0.9186	0.9386	0.9491	0.9585
R52	0.7721	0.7054	0.6421	0.9282	0.939	0.9575	0.9621

The average values of the *acc* and the NMI test results for individual dataset is shown in Tables 7 and 8. It is seen that proposed cdNMF performs well with better accuracy to cluster the documents than conventional ones. Finally, the comparison with baseline NMF proves that the proposed cdNMF has better *acc* and NMI rate for the individual datasets.

## VI. CONCLUSIONS

In this paper, we present a new matrix factorization method called Cholesky Decomposition based non-negative matrix factorization. The Cholesky decomposition collects the data vector, specifically it avoids the non-orthogonality of the non-negative matrix factorization due to its local representation. Also, the presence of non-negative constraints is avoided finally with upper triangular matrix representation for mapping the data vectors. Further, the sparse matrix is eliminated using conjugate gradients, which takes hold of the complex conjugate values from the data vectors. Finally, better accuracy and normalized mutual information is obtained during the experimental validation and it enables better learning of the text data elements with reduced redundancy.

In future, we would like to improve the proposed approach on a graph based NMF framework that could generate better patterns to improve the learning representations of NMI for text mining.

## REFERENCES

- [1] <http://www.cs.umb.edu/~smimarog/textmining/datasets/>.
- [2] Sang, Y. and Yi, Z., "Instance selection by using polar grids," ICACTE. IEEE 3rd International Conference on Advanced Computer Theory and Engineering, vol. 3, pp. V3-344, August 2010.
- [3] García-Osorio, C., de Haro-García, A. and García-Pedrajas, N., "Democratic instance selection: a linear complexity instance selection algorithm based on classifier ensemble concepts," Artificial Intelligence, vol.174, no.5, pp.410-441, 2010.

- [4] Cano, J.R., Herrera, F. and Lozano, M., "Evolutionary stratified training set selection for extracting classification rules with trade off precision-interpretability," *Data & Knowledge Engineering*, vol.60, no.1, pp.90-108, 2007.
- [5] Liao, Q. and Zhang, Q., "Efficient Rank-one Residue Approximation Method for Graph Regularized Non-negative Matrix Factorization," In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer Berlin Heidelberg, pp. 242-255, September 2013.
- [6] Gu, Q. and Zhou, J., "Neighborhood Preserving Nonnegative Matrix Factorization," In *BMVC*, pp.1-10, 2009.
- [7] Shen, B. and Si, L., "Non-Negative Matrix Factorization Clustering on Multiple Manifolds," In *AAAI*, pp. 575-580, July 2010.
- [8] Zhang, L., Chen, Z., Zheng, M. and He, X., "Robust non-negative matrix factorization," *Frontiers of Electrical and Electronic Engineering in China*, vol.6, no.2, pp.192-200, 2011.
- [9] Wang, D., Li, T., Zhu, S. and Ding, C., "Multi-document summarization via sentence-level semantic analysis and symmetric matrix factorization," In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 307-314, July 2008.
- [10] Cooper, M. and Foote, J., "Summarizing video using non-negative similarity matrix factorization," In *proceedings of 2002 IEEE Workshop on Multimedia Signal Processing*, pp. 25-28, December, 2002.
- [11] Brunet, J.P., Tamayo, P., Golub, T.R. and Mesirov, J.P., "Metagenes and molecular pattern discovery using matrix factorization," *Proceedings of the national academy of sciences*, vol.101, no.12, pp.4164-4169, 2004.
- [12] de Fréin, R., Drakakis, K., Rickard, S. and Cichocki, A., "Analysis of financial data using non-negative matrix factorization," In *International Mathematical Forum, Journals of Hikari Ltd*, vol.3, no.38, pp.1853-1870, 2008.
- [13] Chi, Y., Zhu, S., Gong, Y. and Zhang, Y., "Probabilistic polyadic factorization and its application to personalized recommendation," In *Proceedings of the 17th ACM conference on Information and knowledge management*, pp. 941-950, October 2008.
- [14] Dhillon, I.S. and Sra, S., "Generalized nonnegative matrix approximations with Bregman divergences". In *NIPS*, vol.18, December 2005.
- [15] Sha, F., Saul, L.K. and Lee, D.D., "Multiplicative updates for nonnegative quadratic programming in support vector machines," In *NIPS*, pp.1041-1048, December 2002.
- [16] Wang, J., Zhou, J., Wonka, P. and Ye, J., "Advances in neural information processing systems," In *Neural information processing systems foundation*, 2013.
- [17] Ding, C.H., Li, T. and Jordan, M.I., "Convex and semi-nonnegative matrix factorizations," *IEEE transactions on pattern analysis and machine intelligence*, vol.32, no.1, pp.45-55, 2010.
- [18] Ding, C., Li, T. and Jordan, M.I., "Nonnegative matrix factorization for combinatorial optimization: Spectral clustering, graph matching, and clique finding," *ICDM'08. Eighth IEEE International Conference on Data Mining*, pp.183-192, December 2008.
- [19] Yu, D., Chen, N., Jiang, F., Fu, B. and Qin, A., "Constrained NMF-based semi-supervised learning for social media spammer detection," *Knowledge-Based Systems*, vol.125, pp.64-73, 2017.
- [20] Li, H., Zhang, J., Shi, G. and Liu, J., "Graph-based Discriminative Nonnegative Matrix Factorization with label information," *Neurocomputing*, 2017.
- [21] Hayashi, N. and Watanabe, S., "Upper Bound of Bayesian Generalization Error in Non-Negative Matrix Factorization," *Neurocomputing*, 2017.
- [22] Schachtner, R., Poeppel, G., Tomé, A.M. and Lang, E.W., "A bayesian approach to the lee-seung update rules for nmf," *Pattern Recognition Letters*, vol.45, pp.251-256, 2014.
- [23] Du, J.X., Zhai, C.M. and Ye, Y.Q., "Face aging simulation and recognition based on NMF algorithm with sparseness constraints," *Neurocomputing*, vol.116, pp.250-259, 2013.
- [24] Xu, Y., Deng, S., Li, X. and He, Y., "A sparse unmixing model based on NMF and its application in Raman image," *Neurocomputing*, vol.207, pp.120-130, 2016..
- [25] Feng, Y., Xiao, J., Zhou, K. and Zhuang, Y., "A locally weighted sparse graph regularized Non-Negative Matrix Factorization method," *Neurocomputing*, vol.169, pp.68-76, 2015.
- [26] Liao, Q. and Zhang, Q., "Local coordinate based graph-regularized NMF for image representation," *Signal Processing*, vol.124, pp.103-114, 2016.
- [27] Shang, F., Jiao, L.C. and Wang, F., "Graph dual regularization non-negative matrix factorization for co-clustering," *Pattern Recognition*, vol.45, no.6, pp.2237-2250, 2012.
- [28] Wang, J.J.Y., Bensmail, H. and Gao, X., "Multiple graph regularized nonnegative matrix factorization," *Pattern Recognition*, vol.46, no.10, pp.2840-2847, 2013.
- [29] Li, X., Shen, X., Shu, Z., Ye, Q. and Zhao, C., "Graph regularized multilayer concept factorization for data representation," *Neurocomputing*, vol.238, pp.139-151, 2017.
- [30] Wang, J.J.Y., Huang, J.Z., Sun, Y. and Gao, X., "Feature selection and multi-kernel learning for adaptive graph regularized nonnegative matrix factorization," *Expert Systems with Applications*, vol.42, no.3, pp.1278-1286, 2015.
- [31] Zeng, K., Yu, J., Li, C., You, J. and Jin, T., "Image clustering by hypergraph regularized non-negative matrix factorization," *Neurocomputing*, vol.138, pp.209-217, 2014.
- [32] Sun, F., Xu, M., Hu, X. and Jiang, X., "Graph regularized and sparse nonnegative matrix factorization with hard constraints for data representation," *Neurocomputing*, vol.173, pp.233-244, 2016.
- [33] Ou, W., Yu, S., Li, G., Lu, J., Zhang, K. and Xie, G., "Multi-view non-negative matrix factorization by patch alignment framework with view consistency," *Neurocomputing*, vol.204, pp.116-124, 2016.
- [34] Qian, C., Zhuang, Y. and Xu, Z., "Visual tracking with structural appearance model based on extended incremental non-negative matrix factorization," *Neurocomputing*, vol.136, pp.327-336, 2014.
- [35] Wang, D. and Lu, H., "On-line learning parts-based representation via incremental orthogonal projective non-negative matrix factorization," *Signal Processing*, vol.93, no.6, pp.1608-1623, 2013.
- [36] Wang, Y., Wu, S., Mao, B., Zhang, X. and Luo, Z., "Correntropy induced metric based graph regularized non-negative matrix factorization," *Neurocomputing*, vol.204, pp.172-182, 2016.
- [37] Rad, R. and Jamzad, M., "Image annotation using multi-view non-negative matrix factorization with different number of basis vectors," *Journal of Visual Communication and Image Representation*, vol.46, pp.1-12, 2017.
- [38] Arabnejad, E., Moghaddam, R.F. and Cheriet, M., "PSI: Patch-based script identification using non-negative matrix factorization," *Pattern Recognition*, vol.67, pp.328-339, 2017.
- [39] Zong, L., Zhang, X., Zhao, L., Yu, H. and Zhao, Q., "Multi-view clustering via multi-manifold regularized non-negative matrix factorization," *Neural Networks*, vol.88, pp.74-89, 2017.
- [40] Chung, H., Plourde, E. and Champagne, B., "Regularized non-negative matrix factorization with Gaussian mixtures and masking model for speech enhancement," *Speech Communication*, vol.87, pp.18-30, 2017.
- [41] Li, X., Zhang, W. and Dong, X., "A class of modified FR conjugate gradient method and applications to non-negative matrix factorization," *Computers & Mathematics with Applications*, vol.73, no.2, pp.270-276, 2017.
- [42] Han, J., Pei, J. and Kamber, M., "*Data mining: concepts and techniques*", Elsevier, 2011.
- [43] Tan, P.N., "*Introduction to data mining*," Pearson Education India, 2006.
- [44] Zheng, Z., Yang, J. and Zhu, Y., "Initialization enhancer for non-negative matrix factorization," *Engineering Applications of Artificial Intelligence*, vol.20, no.1, pp.101-110, 2007.
- [45] Xue, Y., Tong, C.S., Chen, Y. and Chen, W.S., "Clustering-based initialization for non-negative matrix factorization," *Applied Mathematics and Computation*, vol.205, no.2, pp.525-536, 2008.

- [46] Ding, C., Li, T., Peng, W. and Park, H., "Orthogonal nonnegative matrix t-factorizations for clustering," In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, pp.126-135, August 2006.
- [47] Wild, S., Curry, J. and Dougherty, A., "Improving non-negative matrix factorizations through structured initialization," Pattern recognition, vol.37, no.11, pp.2217-2232, .
- [48] Xu, W., Liu, X. and Gong, Y., "Document clustering based on non-negative matrix factorization," In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, pp.267-273, July 2003.
- [49] Ding, C., Li, T., Peng, W. and Park, H., "Orthogonal nonnegative matrix t-factorizations for clustering," In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 126-135, August 2006.
- [50] Kamvar, K., Sepandar, S., Klein, K., Dan, D., Manning, M. and Christopher, C., "*Spectral learning*," In International Joint Conference of Artificial Intelligence, Stanford InfoLab, April 2003.
- [51] Cai, D., He, X., Wu, X. and Han, J., "Non-negative matrix factorization on manifold," ICDM'08. Eighth IEEE International Conference on Data Mining 2008, pp.63-72, December 2008.
- [52] Flores, F.N. and Moreira, V.P., "Assessing the impact of Stemming Accuracy on Information Retrieval – A multilingual perspective," Information Processing & Management, vol.52, no.5, pp.840-854, 2016.
- [53] Ferro, M.V., Bilbao, V.M.D and Pena, F.J.R., "Modeling of learning curves with applications to POS tagging," Computer Speech & Language, vol.41, pp.1-28, 2017.

# An Efficient Distributed Traffic Events Generator for Smart Highways

Abdelaziz Daaif

SSDIA ENSET Mohammedia, Hassan II University of  
Casablanca, Morocco

Omar Bouattane

SSDIA ENSET Mohammedia, Hassan II University of  
Casablanca, Morocco

Mohamed Yousfi

SSDIA ENSET Mohammedia, Hassan II University of  
Casablanca, Morocco

Oum El Kheir Abra

Computer Science Department  
CRMEF Rabat-Salé Kénitra

**Abstract**—This paper deals with a spatiotemporal traffic events generator for real highway networks. The goal is to use the event generator to test real-time and batch traffic analysis applications. In this context, we represent a highway network as an oriented graph based on the geographic data of the different sensors locations. The traffic is generated based on a socio-cultural calendar using a virtual clock to speed up the simulation process. In order to enable our generator to support the global worldwide highway networks, we propose a dynamic sized distributed architecture based on multi-agent systems. In this platform, we distinguish the physical model based on sensors from the logical model based on an oriented graph. The architecture of the simulator and the results of some of its implementations applied to the Moroccan highway network are presented.

**Keywords**—Event generator; smart highway; simulation; multi-agent systems; distributed computing

## I. INTRODUCTION

Intelligent highways must be instrumented by a set of sensors that detect the passage of vehicles at various strategic points of the infrastructure. Sensors generate immutable events that are collected for real-time or delayed processing depending on the needs of the applications. The availability of reliable real-time measurements or estimations of traffic conditions is a prerequisite for successful traffic control on these highways. The availability and generation of these large masses of data becomes increasingly easy and reliable through the introduction of a number of new automation and communication systems in new vehicles. The main aim of these systems is to improve the safety and convenience of driving, but they are also of great help in alleviating traffic congestion [1].

To achieve improvements in the efficiency of traffic flows on highways, it is essential to develop new methodologies for modeling, estimating and controlling traffic. The literature is very rich in terms of approaches related to modeling and traffic flow control [2]-[5].

To develop and test such Big-data applications before the installation of the sensor infrastructure and the integration of any useful information source, it is important to carry out a

simulation step, which generates the Traffic, and all related events. Such a Big-data platform must merge and harmonize heterogeneous and dynamic data flows. It must also take into account the qualitative and quantitative aspects most relevant to defining the main data, namely:

- Volume, from the always increased data collected.
- Speed, growth of data acquisition.
- Variety, based on the heterogeneity of the data formats and the protocols used.
- Take into account the quality of the data.
- Ability to cope with existing data standards to ensure harmonization of data.
- Provide a robust and scalable storage system.

Not to mention the interoperability of the data considered, as well as the production of missing data in the absence of sensors or information sources, remains an important challenge. For example, in the case of conventional traffic, sensors installed in specific road locations must provide the necessary measures. When the density of the sensors is sufficiently high (e.g. every 500 m), the measurements collected are generally sufficient for monitoring and traffic control; Whereas for a low sensor density, appropriate estimators should be used to reproduce the traffic condition at the required spatial resolution (usually 500 m). The works in [6]-[9] represent some examples dealing with estimation of motorway traffic by the use of conventional detector data.

In a real context, the implementation and maintenance of all instruments dedicated to a road or highway require high costs. To overcome this weakness, various research projects [10]-[14] address the use of other less costly data sources, such as the mobile phone or GPS (Global Positioning System) to estimate road traffic variables.

Over the past decades, technology has changed the way people live, interact and work. The revolution produced by smartphones, the Internet and sensors, results in the daily collection of large volumes of data. For example, Intelligent Transport Systems (ITS) have become flooded with data from

road sensors, mobile device detectors, cameras, radio frequency identification readers, microphones, social media streams and other sources [15].

All ITS actors behave as suppliers and consumers of data, and must react to these large masses of data in their decision-making processes. Their big challenge is not how to collect, but how to process and model large volumes of unstructured data for later analyzes that cannot be effectively addressed through traditional approaches. Thus, it is necessary to develop innovative services and applications capable of processing and inferring information in real time to better support decision-making, but also to anticipate complex situations related to traffic before they occur and to take proactive measures.

In order to meet the challenges outlined above, data must be collected, cleaned, processed and stored efficiently. The ETL process, which means Extract -Transform - Load, is a concept in which data is loaded from a source to a unified data repository. The developers of the ETL platform have extended their solutions to the “Big ETL” platform [16] to provide large data extraction, transformation and loading between large data platforms and traditional data management platforms.

This work proposes a simulator that generates multitudes of events and data to be processed by other modules that draw on the functionalities of ETL and other dedicated platforms such as ITS services, capable of predicting flows Real-time traffic, or to detect traffic-related events [17]-[20].

We validate the performance of estimation schemes developed using simulations using a traffic flow model well known as the ground truth for the state of traffic.

This article is organized as follows: We will begin by giving an overview on highway network, its constituent elements and a smart highway. We then model a highway network by an oriented graph and show the transformations to be carried out until obtaining all the possible paths. We then describe the models used and the architecture of the simulator. Before concluding, we show some results obtained using the first implementation of the simulator applied to the Moroccan motorway network.

## II. SMART HIGHWAYS MODEL

### A. Highway Network Components

A highway network (Fig. 1) consists of several highways that can be interconnected by exchangers. Each highway is viewed as bidirectional graph. It is described in a single direction by a list of elements representing nodes. For the sake of simplicity Table 1 gives a description of an example of highway part from a given city 1 to city 2, giving the properties of each node.

Any highway consists of a symmetric set of elements describing it in one direction; it always starts from an entry followed by several intermediate elements and ends in an exit. The second direction is drawn by inverting the input and the output.

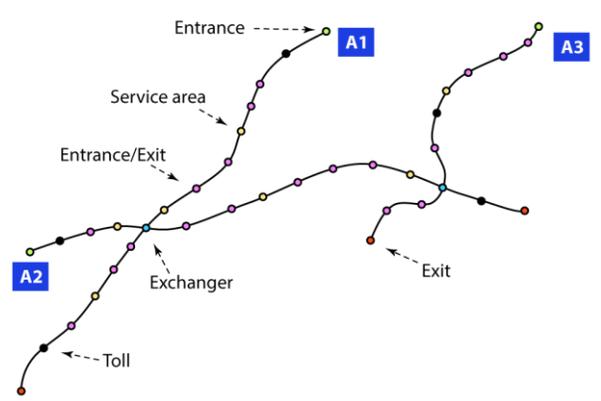


Fig. 1. Highway network.

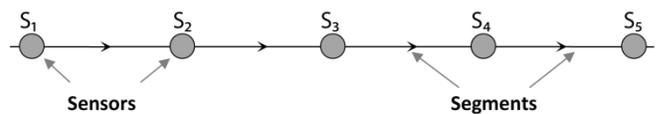


Fig. 2. Segments and sensors.

TABLE I. LIST OF HIGHWAY ELEMENTS

A7 From city 1 to city 2				
	Type	Km	Location	
	Entrance	0	Location 1	
	Entrance/Exit	16	Location2	
	Exchanger	24	Location3	A8
	Service Area	44	Location4	
	Entrance/ Exit	64	Location5	
	Toll	92	Location6	
	Exit	108	Location7	

### B. Smart Highways

In an intelligent highway, all the elements must be instrumented by vehicle traffic sensors (Fig. 2). The sensors delimit segments for which the number of vehicles can be timely determined. Other sensors can be interposed between the elements at strategic points to increase the number of monitored segments.

When a vehicle passes through a sensor, the latter generates a time-stamped immutable event containing information about the vehicle. By processing the events, it is possible to know at any time the number of vehicles in the corresponding segment.

The data collected will be used in real time: firstly by the network management applications, secondly by the network users’ applications such as drivers, smart vehicles, etc.

### C. Highway Network Graph

From the description lists of the highways, an initial oriented graph IOG is constructed. The elements of the lists are represented by IOGV vertices or nodes. The edges IOGE of the graph represent the succession of these elements. The representation of the IOG graph is given in XML file.

In order to determine all the possible paths from all the entrances of the network to all the possible exits, the IOG must be transformed into a new oriented graph TOG describing the entire network in both directions. The TOG nodes will be represented by TOGV and the edges by TOGE.

From TOG, the Dijkstra Shortest Path Algorithm (DSPA) is performed to determine the list of all possible paths.

As mentioned above, the IOG description is given in an XML file consisting of a collection of vertices and a collection of edges. A vertex is represented by an XML Element “vertex” having the following Attributes:

- name: Sensor identifier (ID)
- type: Element type (Enumeration)
- label: Name of the highway (string)
- locality: The locality name of the sensor position (string)
- long: Longitude (double)
- lat: Latitude (double)
- factor: Attendance factor

The edges are represented by the XML Element “edge” having the following Attributes:

- source: Source node (IDREF)
- target: Destination node (IDREF)
- speed: Segment limit speed (double)
- distance: Distance between the two nodes (double)
- lanes : Number of lanes (int)

The “type” attribute of the “vertex” Element can take one of the following values: {I (Entrance), IO (Entrance/Exit), X (Exchange), R (Service Area), T), S (Sensor), O (Exit)}.

The TOG is obtained by performing an elementary transformation at each vertex of the IOG.

To avoid boundary effects and make all these transformations independent, we have inserted white vertices by splitting each of the edges of the IOG (Fig. 3).

Depending on the type of vertex, an elementary transformation will be provided. Table 2 presents a summary of these transformations.

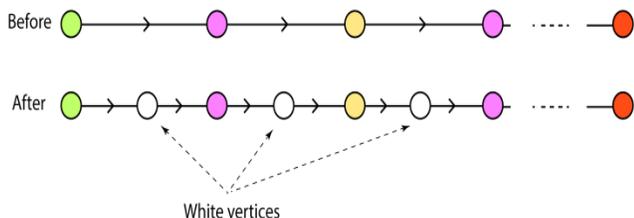


Fig. 3. Isolation of nodes before transformation.

TABLE II. ELEMENTARY TRANSFORMATIONS

Type	IOG Element	WTOG Element
	id-1-1	id-1-1 id-2-1
	id-1-2	id-1-2-4 id-2-2-4 id-1-2-3 id-2-2-3 id-2-2-1 id-2-2-2
	id-1-3	id-2-3-2 id-1-3-2 id-2-3-1 id-1-3-1 id-1-3-3 id-2-3-3
	id-1-4	id-1-4 id-2-4
	id-1-5	id-1-5 id-2-5
	id-1-6	id-1-6 id-2-6
	id-1-7	id-1-7 id-2-7

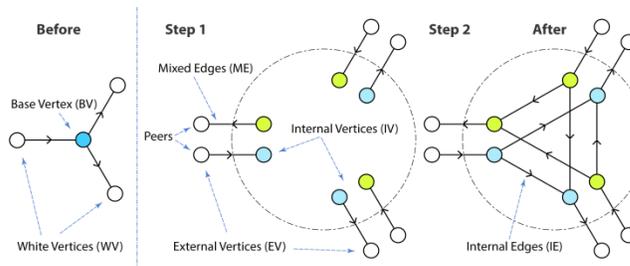


Fig. 4. Transforming an exchanger.

Except for the exchangers, all the elementary transformations generate two independent sub-graphs for the two directions. These sub-graphs are inserted into a temporary intermediate graph WTOG:

$$WTOGV = \{v_1, v_2, v_3, \dots, v_n\}$$

$$WTOGE = \{e_1, e_2, e_3, \dots, e_m\} \text{ where } e_i = (s_i, t_i)$$

The transformation of an exchanger shown in Fig. 4 is performed in two stages:

1) Initialization:

BV = Basis Vertex // Elementary Vertex to be transformed  
 NNW = {  $wv_1, wv_2, \dots, wv_n$  } // White Vertices neighbor of BV

```

EV = {} // External Vertices
ME = {} // Mixed Edges
IV = {} // Internal Vertices
IE = {} // Internal Edges

```

**Stage 1:**

```

For each white vertex  $wv_i$  from NWV
  Create  $ev1_i$  (Clone of  $wv_i$ ) and insert it into EV,
  Create its peer vertex  $ev2_i$  and insert it into EV
  //  $ev2_i$  represents the sensor for the opposite direction,
  Create  $iv1_i$  (Clone of BV) and insert it into IV,
  Create its peer vertex  $iv2_i$  and insert it into IV,
  //  $iv2_i$  represents the sensor for the opposite direction
  IF  $wv_i$  is successor of BV
    Create the edge  $(iv1_i, wv1_i)$  and insert it into ME
    Create the edge  $(wv2_i, iv2_i)$  and insert it into ME
  Else
    Create the edge  $(wv1_i, iv1_i)$  and insert it into ME
    Create the edge  $(iv2_i, wv2_i)$  and insert it into ME
  End IF
End For

```

**Stage 2:**

```

For each edge  $me_i = (s_i, t_i)$  in ME
  IF  $t_i \in IV$ 
    For each edge  $me_j = (s_j, t_j)$  in ME
      IF  $i \neq j$  and  $t_i$  not a peer of  $s_j$  and  $s_j \in IV$ 
        Create edge  $(t_i, s_j)$  and insert it into IE
      End IF
    End For
  End IF
End For

```

2) Sub-graph result:

```

WTOGV = UNION(EV, IV)
WTOGE = UNION(ME, IE)

```

Obtaining the final graph TOG is done by linking the WTOG. This operation consists of removing all the white vertices and restoring the links between the transformed vertices (see Fig. 5).

Finally, we used DSPA to determine all possible paths in the highway network.

**Algorithm:**

**Inputs :**

```

Entries = Vertices of type « I » from TOGV
Exits = Vertices of type « O » from TOGV
Paths = {}

```

**Begin**

```

For each  $v_i$  from Entries
  For each  $v_j$  from Exits
    path := Dijkstra( $v_i, v_j$ )
    IF path is not null
      Insert path into Paths
    End IF
  End For
End For

```

**End For**

**End**

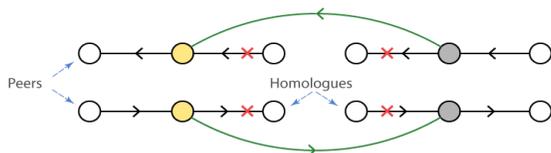


Fig. 5. TOG linking.

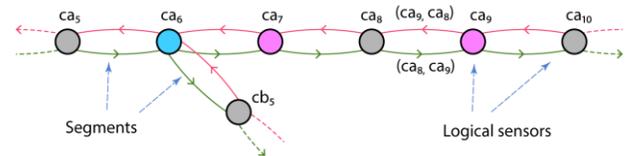


Fig. 6. Logical sensors.

**D. Physical and Logical Sensors**

The paths determined in the previous section of the TOG are normalized so that the nodes forming them belong to the IOG graph (only vertices appearing in the highways description lists). To express the opposite direction of the network, we have added for each edge  $e (s_i, t_i)$  of the graph IOG, an opposite edge  $e (t_i, s_i)$ .

The logical graph corresponds to the graph IOG (Fig. 6). The simulation will use IOG and the list of possible paths to generate traffic.

When passing a vehicle, a logical sensor generates an event that must distinguish the origin and destination of the vehicle. The sensor  $ca_8$  of Fig. 6 has possible origins  $ca_7$  and  $ca_9$  and for possible destinations  $ca_9$  and  $ca_7$ . Therefore, it must be decomposed into two physical sensors  $ca_{1-8}$  for one direction and  $ca_{2-8}$  for the opposite direction. The entrance/exit and exchangers have structures that are more complex. Fig. 7 describes the composition of an entrance/exit.

In Fig. 7 the logical sensor  $ca_7$  is broken down into six physical sensors. Each physical sensor is viewed as a triplet consisting of, the predecessor logical sensor, the current logical sensor and the following logical sensor.

For the logical sensor  $ca_7$ , the physical sensors will have the triplets of Table 3:

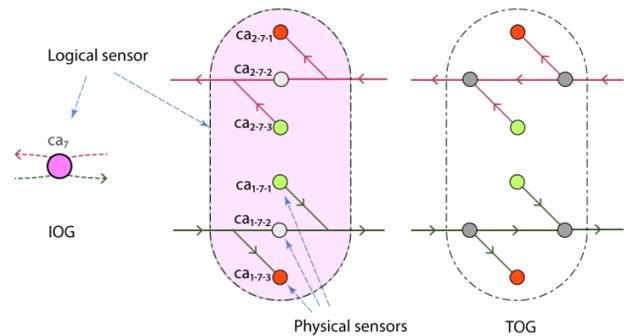


Fig. 7. Physical sensors of an entrance/exit node.

TABLE III. PHYSICAL SENSORS IDENTIFICATION

Logical Sensor	Direction	Physical Sensor	Physical sensor identification
Ca <sub>7</sub>	1	Ca <sub>1-7-1</sub>	(null, ca <sub>7</sub> , ca <sub>8</sub> )
		Ca <sub>1-7-2</sub>	(ca <sub>6</sub> , ca <sub>7</sub> , ca <sub>8</sub> )
		Ca <sub>1-7-3</sub>	(ca <sub>6</sub> , ca <sub>7</sub> , null)
	2	Ca <sub>2-7-1</sub>	(ca <sub>8</sub> , ca <sub>7</sub> , null)
		Ca <sub>2-7-2</sub>	(ca <sub>8</sub> , ca <sub>7</sub> , ca <sub>6</sub> )
		Ca <sub>2-7-3</sub>	(null, ca <sub>7</sub> , ca <sub>6</sub> )

E. Event Model

Depending on the type of sensor, various information can be detected when passing a vehicle. The most common are the speed, the length and the weight of the vehicle. Each time a vehicle passes, an event is generated and a message is sent either directly or via a gateway to an ingestion server. The message must contain at least three basic information; the sensor identifier (previous, current, next), date (Timestamp) and vehicle speed. Here is the general format of the message:

- Identifier of predecessor sensor
- Identifier of the current sensor
- Successor sensor identifier
- Timestamp
- Speed
- Further information about the vehicle depending on the type of sensor.

In the simulation, each vehicle is granted a unique identifier. This makes it possible to track its movement through the highway network.

The simulator is equipped with a virtual clock which allows us to speed up the simulation process. The following section describes the architecture of the proposed simulator.

III. SIMULATOR ARCHITECTURE

A. Overview

The simulator consists of four modules (Fig. 8):

1) *Supervisor*: This module communicates with the coordinator using web sockets to configure and monitor the simulation.

2) *Controller*: This module centralizes the configuration of the highway network infrastructure through the “highway.xml” file.

3) *Worker*: This module runs on several nodes, it starts on each node as many schedulers as processors. Each scheduler uses a random generator to choose a path from the possible paths list, starts a vehicle and assigns it to that path. The vehicle takes the selected path and travels it autonomously by regulating its speed according to the densities of the segments traveled. Each time a vehicle crosses the boundary of a segment an event is generated and sent to the gateway.

4) *Gateway*: This module receives all traffic events generated by all sensors on the network and ingest them to a destination according to the gateway implementation.

B. Simulator Operation

The simulation cluster consists of a node executing the “Controller”, several nodes running “Workers” and one or more nodes performing the “Gateway” module. The controller module must be started first, followed by workers and gateways. The Supervisor module runs on an external node from the cluster.

Fig. 9 shows a “Worker” running on a node with two processors in a simulation cluster of four “Workers”.

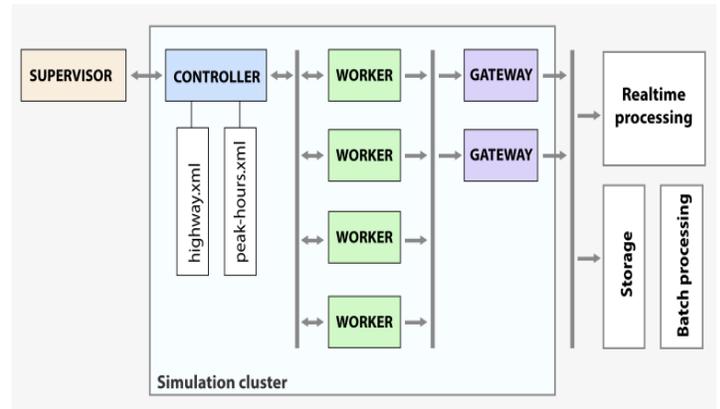


Fig. 8. Simulator architecture.

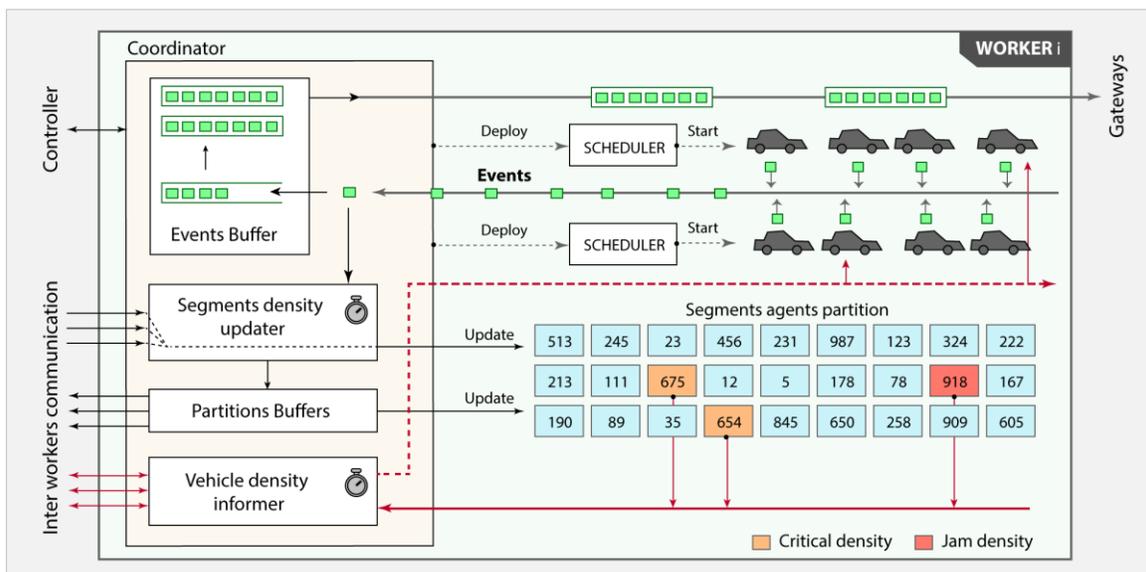


Fig. 9. Worker module architecture.



```

IF r ≥ s AND r < s + pi
    Return pathi
End IF
s = s + pi
End For
End
    
```

#### D. Temporal Traffic Variation

Highway traffic is conditioned by the life and occupations of drivers over time. Most people travel a day rather than a night. During the working days of the week, people travel to work in the morning and return at the end of the day. There is also a significant increase in travel at the beginning and end of holidays. These variations are defined in the XML file “peak-hours.xml”. This file represents the variations related to a nominal state having the value unit. We considered only the variations per day of the week as in the curve of Fig. 11 and 12.

At regular time interval T, the scheduler starts q(iT) vehicles. Where, i is a positive integer. If the duration of the simulation is D and assuming that D is multiple of T, then the total number of vehicles N<sub>v</sub> that are started during D is:

$$N_v = \sum_{i=1}^{\frac{D}{T}} q(iT) = \frac{D}{T} \cdot qa \quad (4)$$

Where, qa is the average of q(t).

If we consider that the duration of the simulation is sufficiently long, we will have:

$$N_v = \sum_{j=1}^M N_j \quad \text{and} \quad N_j = N_v p_j \quad (5)$$

Where, M is the number of possible paths in the graph, N<sub>j</sub> is the number of times that Path<sub>j</sub> has been taken, and p<sub>j</sub> is the probability associated with the choice of this route.

The total distance KVD (6) traveled by all the vehicles that will be engaged during the duration D is the sum of the products of the distances d<sub>j</sub> and the number of times N<sub>j</sub> that the path j has been chosen and traveled by a vehicle:

$$KVD = \sum_{j=1}^M N_j \cdot d_j = N_v \cdot \sum_{j=1}^M d_j \cdot p_j \quad (6)$$

$$dA = \frac{KVD}{N_v} = \sum_{j=1}^M d_j \cdot p_j \quad (7)$$

Where, dA is the average of all chosen paths in the duration D.

The nominal period T is obtained from relations (4) and (7):

$$T = \frac{dA \cdot D \cdot qa}{KVD} \quad (8)$$

The KVD indicator is often given for a period of one year and is marked K<sub>VY</sub> (Vehicles-Kilometers per Year):

$$T = \frac{dA \cdot Year \cdot qa}{K_{VY}} \quad (9)$$

#### E. Simulation Speed

When the objective of the simulation is to generate the data of a long duration to do batch processing, it is important to be able to increase the speed of the simulation process (see Fig. 13).

The time factor (TF) is the ratio between the target duration of the highway traffic and the actual duration of the simulation (10). It expresses the speed of the simulation.

$$TF = \frac{(t'_1 - t'_0)}{(t_1 - t_0)} = \frac{(t' - t'_0)}{(t - t_0)} \quad (10)$$

$$t' = (t - t_0) \cdot TF + t'_0$$

Where, t<sub>0</sub> and t<sub>1</sub> are the starting and ending dates respectively of the simulation.

To increase the speed of the simulation the generator of a scheduler will use the generation period T<sub>g0</sub>:

$$T_{g0} = \frac{T}{TF} \quad (11)$$

Where, TF ≥ 1

With TF = 365, the events of one year are generated in 24 hours of simulation.

The events produced by the simulation are time-stamped with the value of t'.

To be able to simulate large highway networks, a single generator will not suffice. The load should therefore be distributed over several processing nodes.

```

<peak-hours>
...
<peak day="3" from="07:00" to="21:00" factor="1" />
<peak day="3" from="06:00" to="09:00" factor="2" />
<peak day="3" from="12:00" to="13:30" factor="1" />
<peak day="3" from="16:00" to="19:00" factor="2" />
...
    
```

Fig. 11. Peak-hours.xml file.

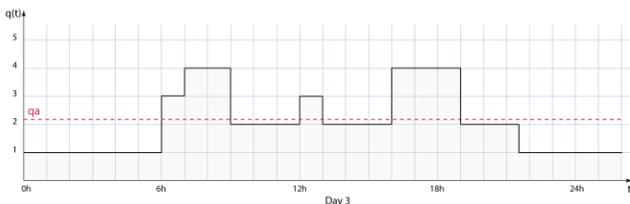
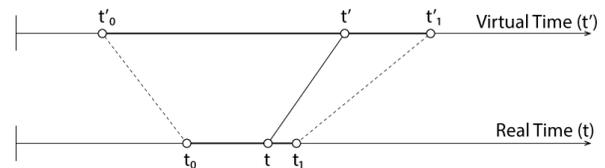


Fig. 12. Traffic variation over one day.



Virtual time = Event time      t'<sub>1</sub> - t'<sub>0</sub> = Target duration  
 Real time = Processing time      t<sub>1</sub> - t<sub>0</sub> = Processing duration

Fig. 13. Event time and processing time.

## F. Distribution and Load Balancing of Traffic Generators

### 1) Schedulers distribution

In a cluster of N workers nodes each having M processors, NxM traffic generators can be executed. The load is automatically distributed proportionally using the vehicle commitment period  $T_g$  :

$$T_g = N.M. \frac{T}{TF} \quad (12)$$

Whenever a vehicle leaves one segment to enter another, it must inform the agents responsible for these two segments. The number of these agents is equivalent to the number of segments of the highway network.

These agents must also be deployed equitably on the N workers nodes.

### 2) Partitioning and Distributing Segment Agents

The agent is responsible for the segments (Fig. 9), must be partitioned equitably and distributed on all workers in the cluster. A path i is a sequence of one or more segments. A segment belongs to one or more paths. The number of messages arriving at a given segment is related to its frequentation. Knowing the paths probabilities, we can determine the level of attendance at each segment. To do this, we give for each segment (Edge) a weight (ws) which will have the sum of all the probabilities of the paths to which this segment belongs.

The partitioning is done by quasi-balancing the sums of the weights of the segments of each partition.

Initialization:

Amounts = ArrayOfDouble(k)

Partitions = ArrayOfSegments(k)

Segments = {s<sub>1</sub>, s<sub>2</sub>, ..., s<sub>n</sub>}

WS = {ws<sub>1</sub>, ws<sub>2</sub>, ..., ws<sub>n</sub>}

Begin :

For each Segment s<sub>j</sub> in Segments

// get the index of the minimum Amounts

index = getMinIndexOf(Amounts)

Insert s<sub>j</sub> in Partitions[index]

End for

End

## G. Events Density over Time

When the time factor TF increases, the number of events generated also increases in a given time. It is then important to know in advance this events density (DE) with respect to time (Events/s) in order to size the simulator modules.

$$DE = TF \cdot \frac{Nv}{D} \cdot \sum_{j=1}^M p_j \cdot Nc_j \quad (13)$$

Where, Nc<sub>j</sub> is the number of sensors in the path j.

## H. Vehicle Itinerary

Before entering the highway, the vehicle registers to be informed about the state of the density of the first segment. After a certain delay, the vehicle calculates its speed, determines the duration of the crossing of the segment and starts a timer to warn it at the end of the segment.

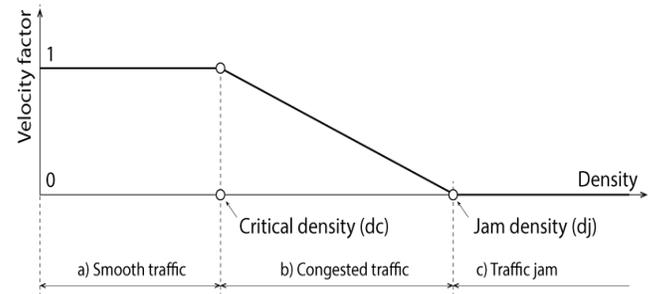


Fig. 14. Relationship between velocity factor and segment density.

The nominal speed  $V_n$  (14) is calculated using the speed limit  $V_l$  in that segment and a random component representing the temperament of the driver.  $V_l$  is indicated in the highway.xml file:

$$V_n = V_l \cdot \left( \frac{1 + \alpha \cdot r}{2} \right) \quad (14)$$

Where, r is the temperament factor of the driver, it is a random number with  $0 \leq r < 1$ ,  $\alpha$  is the speed excess factor with  $\alpha \geq 1$  and  $V_l$  is the speed limit of the next segment.

The speed of the vehicle (15) is obtained by weighting the nominal speed by the velocity factor  $F_v$  (16) which depends on the density of the segment according to Fig. 14.

When a vehicle arrives on a segment, it registers itself to be informed about the density of the next segment. Messages are sent only to concerned vehicles and only when the density of a given segment reaches or exceeds the critical value (Fig. 14).

$$F_v = \begin{cases} 1, & d < dc \\ \frac{d - dj}{dc - dj}, & dc \leq d < dj \\ 0, & d \geq dj \end{cases} \quad (15)$$

## IV. APPLICATION: MOROCCAN HIGHWAY NETWORK

### A. Simulator Implementation Approach

In order to scale and simulate large highway networks, the simulator's architecture must be distributed and based on loosely coupled components. The reactive-programming (RP) proposed in the "manifesto of reactive programming", provides a general framework for this kind of applications. The RP systems are responsive, resilient, elastic and message driven. In this article, we proposed an architecture using multi-agent systems in which agents are autonomous and collaborate with each other by exchanging messages. The scheduler, vehicle, agents responsible for highway segments are the examples of agents used in this application. When a vehicle is created by a scheduler, it becomes autonomous and is the only one to have access to its state. To make decisions, it needs to be informed about the state of its environment. To do so it proceeds by exchanging messages with the agents to get information. Whenever a vehicle arrives at the end of a segment, a message is sent to a gateway which itself is an agent. In the case of a large network, multiple balanced gateways can be used. Each worker hosts a partition of

representative agents of the segments. These agents receive regular messages when vehicles enter or leave their segment. When the density reaches a critical threshold, the agents must inform all vehicles that are about to arrive in that segment.

From a functional point of view, the application must manage a large number of autonomous agents and must be intensive messaging oriented. Since the application is distributed, the data network exchange is also intensive.

Our first implementation of the simulator uses the toolkit Eclipse Vert.x which facilitates the implementation of the RP.

**B. Platform Materials**

The proposed simulator is tested using a real example of Moroccan highway network. The seven highways of the Moroccan network (Fig. 15) are entered in the highway.xml file (Fig. 16). The function  $q(t)$  expresses the temporal traffic variations in highway (Fig. 17). Table 4 shows some highway network features. We performed a test with the topology shown in Table 5 using the parameters in given in Table 6.

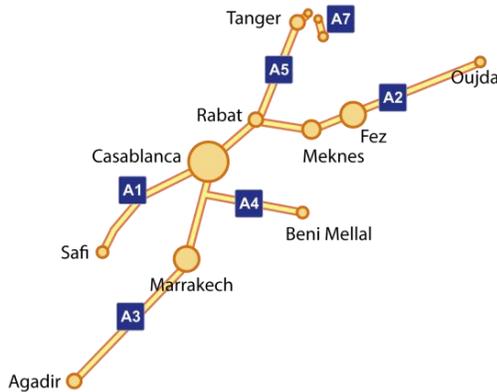


Fig. 15. Map of the Moroccan highway network.

```
<?xml version="1.0" encoding="UTF-8" ?>
<?xml-stylesheet type="text/xsl" href="highway.xsl" ?>
<graph>
  <vertex name="A1-1-01" km="9" label="A1" type="O" lat="33.846320"
  long="-6.898465" locality="Tamesna" factor="0.7" direction="1" />
  <!-- Other 145 vertices -->
  <edge source="A1-1-01" target="A1-1-02" speed="120" distance="5" lanes="2"
  />
  <!-- Other 142 edges -->
```

Fig. 16. Highway.xml.

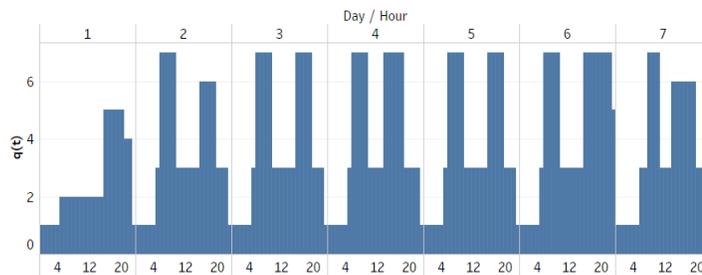


Fig. 17. Values of  $q(t)$  extracted from peak-hours.xml.

TABLE IV. MOROCCAN HIGHWAY NETWORK FEATURES

Number of highways	7
Network total length	1 736 Km
Number of vertices (Nodes)	146
Number of edges (Segments)	286
Service areas	29
Number of entrances	184
Number of exits	184
Number of exchangers	5
Number of paths	2646
The shortest path length	2 Km
The longest path length	709 Km
The average path length	199.7475 Km

TABLE V. TEST PLATFORM TOPOLOGY

Module	Nodes	Features
Worker	2 Nodes	OS Linux 64bit 8 cores 3GHZ RAM 16GB NIC10GB
Controller + supervisor	1 Node	OS Linux 64bit 4 cores RAM 16GB NIC 10GB
Gateway	1 Node	OS Linux 64bit 4 cores RAM 16GB NIC 10GB

TABLE VI. TEST PARAMETERS

Parameter	Value
Simulation duration	24 hours
Coefficients $\alpha, \beta$ et $\gamma$	0.3, 0.3, 0.4
KVD: Kilometer Vehicle per Day	24 000 000 Km
TF : Time factor (1 year)	365
dc : Critical density	20 Vehicles/Km
jd : Jam density	200 Vehicles/Km

**C. Simulation Results**

Table 7 shows the predictable results of the simulation calculated only with the parameters of the simulation. The processing of the data generated by the simulation confirms the expected results. Traffic analysis shows that the spatial variations of traffic expressed by the path factor (2) have been respected as shown in Fig. 18. Fig. 19 shows that the set point of 24 million vehicles kilometers per day has been reached. The variation profile of the daily traffic corresponds to the set point given by the peak-hours.xml file (Fig. 20). Fig. 21 shows the number of traveled vehicles in the highway network each month of 2017 year.

TABLE VII. PREDICTABLE RESULTS

Paramètre	Valeur
$dA$ : The average route distance $dA = \sum_{j=1}^M d_j \cdot p_j$	160.784549
$Qa$ : The average of $q(iT)$ determined from peak-hours.xml file.	1.907
Total number of engaged vehicles : $Nv = \frac{KVY}{dA}$	54 482 847
$T$ , Nominal period of starting of vehicles $T = \frac{dA \cdot D \cdot qa}{KVD}$	1.103818 s
$Tg$ , Effective period of starting of vehicles $Tg = N \cdot M \cdot \frac{T}{TF}$	0.048386 s
DE : Events density (Events/s) $DE = TF \cdot \frac{KVD}{dA \cdot D} \cdot \sum_{j=1}^M p_j \cdot Ns_j$	9485.2 Events/s

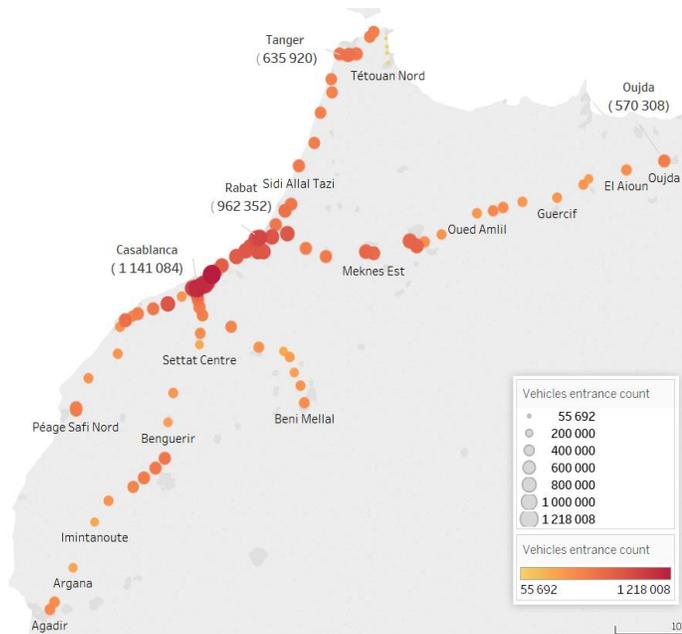


Fig. 18. Spatial variation of vehicle entries.

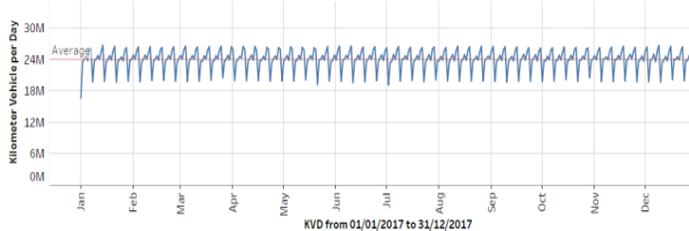


Fig. 19. Kilometer Vehicle per Day (Average = 24 014 791).

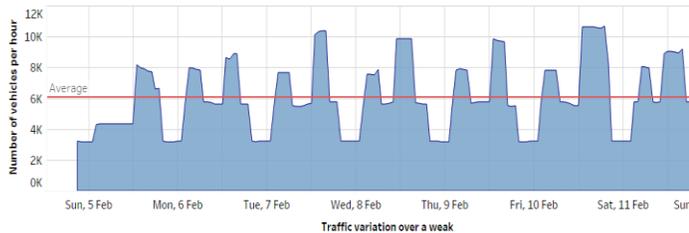


Fig. 20. Variation over a week from Wed 1 Feb to Tue 7 Feb 2017.



Fig. 21. Number of engaged vehicles per month over 2017.

## V. CONCLUSION

In this paper, we proposed a distributed simulator model of intelligent highways based on multi-agent systems and a virtual calendar. The simulator allows near realistic traffic to be generated based on the distribution of the population juxtaposing the network and the lengths of the journeys to be traveled. The simulator takes into account the variations of highway traffic over time which are described in an XML file. At regular time intervals and according to the current date, a traffic generator generates a flow of vehicles and affects them random paths. Each vehicle is an autonomous agent; it regulates its speed according to the state of the traffic. To validate this model, we implemented it using reactive programming and we tested it for the case of the Moroccan highway network. We carried out a simulation of the traffic of a year in 24 hours with a constraint of 25 million of Vehicles Kilometer per Day, producing an events density of 9485.2 Events/s. The amount of data produced by this simulation is of the order of 102 GB.

As a perspective, real-time and batch processing models of data generated by traffic will be developed and tested. The simulator can then be extended by allowing vehicle interaction with Big-data applications to improve the use and safety of highway networks. On the other hand, the spatial and temporal components of our generator are independent. The generator should link these two components as is the case in reality; a socio-cultural event is associated with both geography and time.

## REFERENCES

- [1] C. Diakaki, M. Papageorgiou, I. Papamichail, and I. K. Nikolos, "Overview and analysis of vehicle automation and communication systems from a motorway traffic management perspective," *Transp. Res. A, Policy Pract.*, vol. 75, pp. 147–165, May 2015.
- [2] C. de Fabritiis, R. Ragona, and G. Valenti, "Traffic estimation and prediction based on real time floating car data," in *Proc. IEEE Conf. Intell. Transp. Syst.*, Beijing, China, 2008, pp. 197–203.
- [3] J. I. Ge and G. Orosz, "Dynamics of connected vehicle systems with delayed acceleration feedback," *Transp. Res. C, Emerging Technol.*, vol. 46, pp. 46–64, Sep. 2014.
- [4] A. Kesting, M. Treiber, M. Schonhof, and D. Helbing, "Adaptive cruise control design for active congestion avoidance," *Transp. Res. C, Emerging Technol.*, vol. 16, no. 6, pp. 668–683, Dec. 2008.
- [5] S.-C. Lo and C.-H. Hsu, "Cellular automata simulation for mixed manual and automated control traffic," *Math. Comput. Modell.*, vol. 51, no. 7/8, pp. 1000–1007, Apr. 2010.
- [6] L. Alvarez-Icaza, L. Munoz, X. Sun, and R. Horowitz, "Adaptive observer for traffic density estimation," in *Proc. IEEE ACC*, Boston, MA, USA, 2004, pp. 2705–2710.
- [7] I.-C. Morarescu and C. Canudas de Wit, "Highway traffic model-based density estimation," in *Proc. IEEE ACC*, San Francisco, CA, USA, 2011, pp. 2012–2017.
- [8] F. Morbidi, L.-L. Ojeda, C. Canudas deWit, and I. Bellicot, "A new robust approach for highway traffic density estimation," in *Proc. IEEE ECC*, 2014, pp. 2575–2580.
- [9] L. Munoz, X. Sun, R. Horowitz, and L. Alvarez, "Traffic density estimation with the cell transmission model," in *Proc. IEEE ACC*, 2003, pp. 3750–3755.
- [10] A. Anand, G. Ramadurai, and L. Vanajakshi, "Data fusion-based traffic density estimation and prediction," *J. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 367–378, 2014.
- [11] V. Astarita, R. L. Bertini, S. d'Elia, and G. Guido, "Motorway traffic parameter estimation from mobile phone counts," *Eur. J. Oper. Res.*, vol. 175, pp. 1435–1446, 2006.

- [12] J. C. Herrera et al., "Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment," *Transp. Res. C, Emerging Technol.*, vol. 18, no. 4, pp. 568–583, Aug. 2010.
- [13] W. Deng, H. Lei, and X. Zhou, "Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach," *Transp. Res. B, Methodol.*, vol. 57, pp. 132–157, Nov. 2013.
- [14] T. Seo, T. Kusakabe, and Y. Asakura, "Estimation of flow and density using probe vehicles with spacing measurement equipment," *Transp. Res. C*, vol. 53, pp. 134–150, Apr. 2015.
- [15] J. Fiosina, M. Fiosins and J. Müller, "Big Data Processing and Mining for Next Generation Intelligent Transportation Systems," *Jurnal Teknologi*, vol. 63, no. 3, pp. 21-38, 2013.
- [16] J. Caserta and E. Cordo, "Big ETL: The Next 'Big' Thing," 9 February 2015. <http://data-informed.com/big-etl-next-bigthing/>.
- [17] Q. Ou, R. L. Bertini, J. W. C. van Lint, and S. P. Hoogendoorn, "A theoretical framework for traffic speed estimation by fusing low-resolution probe vehicle data," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 747–756, Sep. 2011.
- [18] M. Papageorgiou and A. Messmer, "METANET: A macroscopic simulation program for motorway networks," *Traffic Eng. Control*, vol. 31, no. 9, pp. 466–470, 1990.
- [19] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data," *Transp. Res. C, Emerging Technol.*, vol. 19, no. 6, pp. 1243–1262, Dec. 2011.
- [20] "Next Generation SIMulation (NGSIM)," U.S. Dept. Transp., Washington, DC, USA, 2005. Available: [www.ngsimcommunity.org](http://www.ngsimcommunity.org).

# The Role of Strategic Information Systems (SIS) in Supporting and Achieving the Competitive Advantages (CA): An Empirical Study on Saudi Banking Sector

Nisreen F. Alshubaily

Al Imam Mohammad Ibn Saud Islamic University (IMSIU)  
Riyadh, Saudi Arabia

Abdullah A. Altameem

Al Imam Mohammad Ibn Saud Islamic University (IMSIU)  
Riyadh, Saudi Arabia

**Abstract**—The purpose of this research paper is to identify the significant role of Strategic Information Systems (SIS) in supporting the Competitive Advantage (CA). It also explains its role on the dimensions that increase the competitive advantage which are the operational efficiency, information quality and innovation. In order to achieve the goal of this study and to collect the primary data, the researchers designed a survey, in the form of an electronic questionnaire. This survey instrument consisted of 20 questions. It was distributed to members of the study sample, which contains the managers at all levels, and the employees in the Saudi banking sector. The number of the participants included in the survey was 147. The results of this study revealed that there is a significant role of strategic information systems on increasing operational efficiency, improving the quality of information and promoting innovation. This in turn enabled the organizations to achieve higher levels of competitive advantages. The strategic information systems have deep consequences for organizations that adopt them; managers could achieve great and sustainable competitive advantages from such systems if carefully considered and developed. On the other hand, this study was conducted in the banking sector in KSA context. So, more research is needed in other sectors and in the context of other countries; to confirm and generalize the results. Finally, the paper's primary value lies in its ability to provide the evidence that the strategic information systems play a significant role in supporting and achieving the competitive advantages in Saudi Arabia, particularly in the banking sector. Since there was a lack of such research in the Saudi context, this paper can provide a theoretical basis for future researchers as well as practical implications for managers.

**Keywords**—Strategic information systems (SIS); competitive advantage (CA); operational efficiency; information quality; innovation

## I. INTRODUCTION

In the light of today's global economy, the organizations face several challenges such as globalization, privatization, stiff competition and more demanding customer expectations, coupled with daily advancement in information and communication technologies. In this environment, the top managers should understand and realize that the IS/IT is not merely a resource to support day-to-day operations. They should also realize that the clever use of IS/IT can

significantly change an organization's long term strategic position in national and global markets. Therefore, it becomes increasingly imperative that the managers create new and different strategies including the change of top management for long-term planning and strategic decision-making versus the operational decision-making. Subsequently, if the organizations wish to remain successful and to be competitive, the managers need to consider Information Systems (ISs) as a tool utilized to gain competitive advantages, in order to overcome the other competitive organizations. So, the information systems that help seize opportunities of gain competitive advantages are often called Strategic Information Systems (SIS).

The strategic information system can be defined as an information system that creates or enhances the company's competitive advantage or changes the industry structure by fundamentally changing how business is conducted. It is conventional information systems used in innovative ways [1]. In [2], [3], the authors confirm that it can be any kind of information systems (such as TPS, MIS, DSS, EIS, OAS, ERP, etc.) that helps an organization: 1) gain a competitive advantage; 2) reduce a competitive disadvantage; and/or 3) meet other strategic organization objectives. Hence, any IS has the ability to change the goals, processes, products, or environmental relationships to help an organization gain a competitive advantage or reduce a competitive disadvantage is a strategic IS [1], [4].

In addition, the SIS involves using information technology to develop products, services, and capabilities that give a company strategic advantages over the competitive forces it faces in the global marketplace [2]. The advances in information provision have led organizations to attempt to develop IS or IT strategies align with their business strategies to achieve many benefits [1]. Such as helping the organization to reduce overall costs, get fewer errors and greater accuracy when performing operations, produce high-quality products and services, accelerate communication and data sharing, improve performance and productivity, and make management more efficient and effective. Moreover, it gives the managers the ability to adjust, control and monitor all business processes which accordingly will accelerate the processes of the decision-making [3], [5]-[7].

In this context, the purpose of this paper is to highlight the role of Strategic Information Systems (SIS) in supporting and achieving Competitive Advantage (CA); in order to obtain the higher level of operational efficiency, information quality and innovation on the Saudi banking sector.

## II. LITERATURE REVIEW

The concept of Strategic Information Systems (SIS) was introduced for the first time in the field of information systems in the early 1980s by Dr. Charles Wiseman [8]. The strategic information systems have been established as a core activity in the governance and management of information technology in organizations. Moreover, they have become a very challenging subject for scientists and practitioners in the recent years [9], [10]. Strategic information systems are essential to help organizations succeed in today's highly competitive global business environment.

The organizations in the current IT age need to use information systems effectively which require an understanding of the organization, management, and information technology that form the systems [11]. It also requires an understanding that the mission of the information systems itself is changed and evolved from a focus on efficiency and effectiveness to a focus on organization performance as the foundation for competitiveness in a rapidly changing environment [12]. Therefore, the top managers should understand that the information systems alone cannot provide an enduring business advantage. In order to obtain the competitive advantage; it is important to develop appropriate strategies that help to use the IS/IT based systems effectively and provide means to manage them [11]. To achieve this goal, an increasing number of the organizations are turning to develop Information Systems Strategies (ISS) by applying one methodology or approach of Strategic Information Systems Planning (SISP) [13], which helps them to convert their conventional information systems to Strategic Information Systems (SIS).

Strategic Information Systems are systems that help organizations alter their business strategies, plans or structure. They are also used to hasten the reaction time of the environmental changes and aid the organization to achieve a competitive advantage over its competitors. Strategic information systems are the traditional or conventional information systems used in innovative ways [1]. The essential purpose of the strategic information systems is to help organizations to do things better [14]. They also aim to develop and maintain the IS/IT systems that support the business operations in an effective way [11].

They are as instrumental in achieving the organization's competitive objectives or other strategic objectives [3]. They help an organization gain a competitive advantage through its contribution to achieve the strategic goals and/or its ability to significantly increase the performance and productivity [4]. The main objective of SIS according to [1] is "to define the explicit connection between an organization's business plans and IS plans to better achieve the goals and objectives and to provide closer management control for the critical IS/IT systems".

However, the successful use of information systems in order to achieve a competitive advantage is difficult. It requires precise coordination of technology, organization, and management [15]. Although strategic information systems can support or shape organizational strategies and its success promises considerable and great benefits, but the failure to develop and implement them is common [16]. The implementation of strategic systems often requires extensive organizational change and a transition from one socio-technical level to another. Such changes are called strategic transitions and are often difficult and painful to achieve [15]. In addition, there are several attempts that have been made to identify the opportunities that help in developing the strategic systems as competitive weapons. Nevertheless, the most instances of strategic information systems success are retrospective, and most organizations have no definite plan to develop effective strategic information systems. This leads to questions concerning how much IS managers really understand about the development of strategic systems [17], [18].

Laudon in [15] confirms that the top management must understand that not all strategic information systems are profitable, they can be expensive to build and easily copied by other firms so that strategic advantage is not always sustainable. Strategic information systems have to be built on the strengths of the company that cannot be easily imitated. It has been determined that lasting, sustainable competitive advantage can be gained from strategic information systems only if an organization possesses other resources as well. Such resources include 1) a well-developed and flexible information technology platform or a database to obtain the advantages; and 2) continual investment to maintain those advantages [19]. Therefore, some of the recommended conditions/actions should be followed before the development and implementation of strategic information systems in the Organization. These proposed conditions/actions ensure the successful deployment of strategic information systems, which are as follows:

- Active support of senior organization management — not just MIS management — in the discovery of strategic opportunities and in the implementation process.
- Integration of planning for the strategic use of information systems into the overall organization strategic planning process.
- Direct reporting by those responsible for strategic use of information systems to the business managers of the area to be affected by the new system.
- Placement of control mechanisms in the hands of these business managers.
- Readiness for strategic use of information systems, implying the successful use of the MIS and technological platform already in place and experience with technological innovation [19].

In the light of the above, organizations should seek to apply one or more strategic information systems to obtain

many benefits and strategic reasons. The literature classified the benefits of strategic IS under three classifications, which are the alignment and competitiveness as well as strategic analysis. The benefits under strategic analysis will support the organization to raise its efficiency, effectiveness and performance to the highest level. On the other hand, the benefits under both alignment and competitiveness will support the organization to achieve the sustainable advantages. The following are the main features and benefits of strategic information systems under each classification: 1) Strategic analysis (support decision-making process, increase organizational efficiency and effectiveness, support different organizational levels, increase productivity of employees, support coordination of work, increase quality, reduce costs, support reactions to change and create new strategic opportunities). 2) Competitiveness (develop/produce new product/services, obtain competitive advantages, increase organizational competitiveness, support innovation and improve market share). 3) Alignment focus (integrate IS strategic plan into business strategic plan, consolidate the operations by integrating distributed systems, create standards, improve knowledge and improve resource creativity and flexibility) [3], [5]-[7].

### III. DIMENSIONS THAT FORM THE ROLE OF SIS FOR ACHIEVING CA

It is critical that the organizations understand how to build the competitive advantages from the strategic IS. They also should understand that this process involves understanding the needs of the stakeholders, and devising strategies to IS to effectively utilize the resources available (or which can be obtained). This in turn will increase the organizational performance that is sustainable and successful over the long term [20], [21]. In addition, many researchers confirmed that the competitive advantage is at the core of an organization's success or failure [4]. Ketchen et al. in [22] defined the competitive advantage as "the extent to which an organization has the competency to create a defensible position over its competitors as a result of critical management decisions based on established strategies which differentiate itself from its rivals" [22]. Moreover, these strategies should take into account the target market, the business' strengths and weaknesses, the business' goals and objectives, the product and service and the strategies of the competition; to be able to achieve the competitive advantages [20]. The feedback from some of the literature reviews concluded that the utilization of strategic information systems has a positive association with the competitive advantage by different dimensions such as increasing operational efficiency, improving information quality and encouraging innovation. The objective of this section is to clarify the dimensions that form the role of strategic information systems, which all previous studies agreed to consider these dimensions as a competitive advantage for delivering a higher level of organizational performance.

#### A. Operational Efficiency

Based on the literature review, the operational efficiency refers to the effective use of human and material resources to increase the output of products and services, reduce costs and

maximize profits. It also reflects the capability of an organization to deliver products or services to its customers in the most cost-effective manner while still ensuring the high quality of its products and services. In addition, it means producing more products and services with no greater use of resources or maintaining the same level of production using fewer resources. Increasing or improving operational efficiency is a key component of the company's revenue growth and the competitive advantages achievement.

In order to achieve operational efficiency, the company needs to minimize the redundancy and waste while taking advantage of resources that contribute significantly to its success and benefiting from the best manpower, technologies and business processes. It is also achieved by streamlining the company's core processes to respond more effectively to the continually changing market forces [23], [24]. The process of reducing internal costs resulting from operational efficiency enables the company to achieve higher profit margins or be more successful in highly competitive markets [24] and this can be done by using several strategies and techniques such as SIS [25]. Philip in [26], believes that the operational efficiency is the strategic IS goal, especially when competitive advantage is the primary objective.

Therefore, the organizations should bear in their mind that the strategic IS for operational efficiency can be as important and productive as planning for competitive advantages [26].

#### B. Information Quality

Based on the literature review, information quality refers to the quality of outputs that information systems produced, which can be in the form of reports or online screens. The organizations should focus on assuring the reliability of data to increase the quality of the systems [27]. Therefore, the high quality of the systems leads to high quality of the information and services. The high quality of information and services helps organizations to manage their business processes, increase decision-making efficiency, improve organizational performance and perform their jobs efficiently and effectively [27], [28]. In addition, the information provided by ISs, which is inconsistent with the needs of users, will be subjected to heavy maintenance costs and disrupt the operations, resulting in an increase in the overall costs at the organization [27].

Moreover, the information systems processing is similar to production processing in manufacturing organizations. If the product (information) is not delivered on time (timeliness) and the product (information) does not conform to the needs (relevance) of customers (users), then the customers (users) will be dissatisfied and the firm will lose business [29]. So, the information quality has been typically regarded as a key strategic component of competitive advantage. It also helps to enhance the service and product quality in organizations [27], [28]. The organizations without the ability to assess the quality of their information, they cannot assess the status of their organizational performance and monitor their improvement [30].

This can be done by using strategic information systems, which will provide significant benefits by improving the quality of information that enables the decision-maker to

further improve the quality of his/her decisions outcomes in order to obtain competitive advantages [31].

### C. Innovation

The organizations frequently adopt innovations to gain capabilities and competitive advantages [32]. If organizations want to be competitive, they must also be innovative [33]. Rogers in [34], defined the processes of innovation as the development and implementation of introducing new ideas or new technologies, that lead to achieve the sustainable competitive advantages to the organization. The innovation capability of an organization depends closely on its intellectual and/or organizational knowledge assets and on its ability to employ these assets [35]. Adopting the innovations in organizations in order to achieve higher levels of competitiveness, will lead to produce lower-cost products with better quality compared to those competitors [36]. Moreover, the innovations are not just a process of developing new information systems or technologies to produce new products or services, but in many cases, are a process of creating new models and strategies for information systems or technologies to do business better in the face of change [33].

In addition, the innovations are a crucial driver for improving organizational performance and achieving sustainable competitive advantages [36]. Therefore, an organization needs to enable innovation to take place through employing IT experts with SIS skills and managerial IT skills to obtain the competitive advantages [27]. Finally, the corporate leaders and IT managers should view the SIS as part of their company's strategic innovation, which has been considered as an important tool for achieving the competitive advantages.

## IV. RESEARCH CONCEPTUAL MODEL AND HYPOTHESES

Based on the overall results, derived from other studies that are closely related to this research, the conceptual model is proposed. This model will be used for identifying and analyzing the nature of the problem, detailing exactly what is going to be researched and examining the arguments of others. It also will be used as a road map for empirical data collection and analysis, and to establish a comprehensive overview of applying the strategic information system as a competitive advantage tool, especially on Saudi Banking Sector. Fig. 1 shows the components of the research conceptual model including the role of Strategic Information Systems (SIS) which is the independent variable. In addition, the figure shows the Competitive Advantage (CA) and its dimensions which are operational efficiency, information quality and innovation as the study dependent variables.

Based on the literature review and research model, the main hypothesis of the research is given below:

**H1:** There is a positive association between the role of Strategic Information Systems (SIS) and Competitive Advantage (CA).

This branched into three dimensions that increase the competitive advantage:

**H1<sub>a</sub>:** There is a positive association between the role of Strategic Information Systems (SIS) and operational efficiency.

**H1<sub>b</sub>:** There is a positive association between the role of Strategic Information Systems (SIS) and information quality.

**H1<sub>c</sub>:** There is a positive association between the role of Strategic Information Systems (SIS) and innovation.

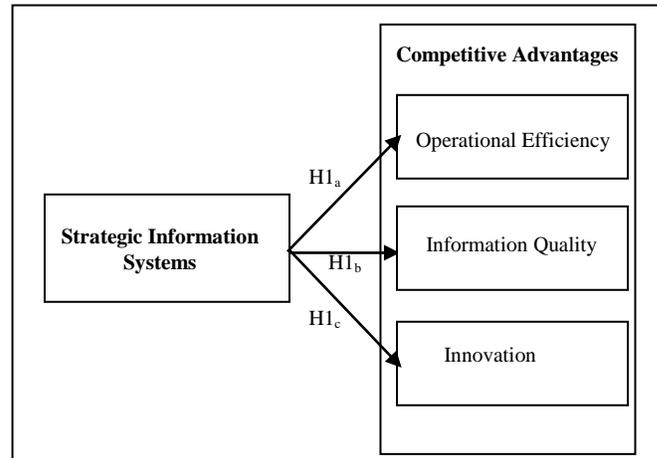


Fig. 1. The proposed conceptual model.

## V. RESEARCH DESIGN AND METHOD

### A. Sampling

The target population for this study consisted of the senior managerial level and the employees of various banking, regardless their size and location in Saudi Arabia. Researchers chose Saudi bank sector because it relies mainly on information systems when conducting their procedures. Therefore, it is important to measure the role of SIS in this sector.

In addition, random sampling was used as sampling technique so that each population member had an equal chance of being selected.

### B. Data Collection Procedure

In order to collect the primary data for this empirical study and to test the research model and verify the hypotheses, the researchers developed survey; in the form of an electronic questionnaire. It is an electronic form containing a series of questions, and designed to extract specific information from respondents [37]. This method of data collection was considered appropriate because it is inexpensive and has the ability to collect a large amount of information from a large number of Bank employees in a short period of time. In addition, the survey is a popular research technique in obtaining quantitative data.

The survey instrument in this study consisted of 20 questions categorized into five main parts. The first part included questions related to demographic information of the respondents (e.g. gender, age, educational level, etc.). The other parts included the questions related to the measurement of the role of Strategic Information Systems (SIS) in supporting and achieving Competitive Advantage (CA)

through the three dimensions which are operational efficiency, information quality and innovation.

C. Data Analysis Procedure

The fundamental purpose of the data analysis procedure is to achieve the research objectives and to find answers to the identified research questions and hypotheses. Davis suggested in [38] four primary steps that should be followed to conduct the data analysis, which are as following:

- Step 1: Selecting an analytical tool
- Step 2: Preparing the data for analysis
- Step 3: Identifying specific statistical techniques
- Step 4: Finally, presenting the analysis results [38].

In this study, the researchers followed these four steps to carry out the data analysis. Statistical Packages for the Social Sciences (SPSS) version (24) was used to analyze the collated data. However, after the selection of the analytical software and before the process of analysis, the raw data was prepared and cleaned to ensure the data has got no missing value or outliers. The SPSS was used to encode the data, and the examination was carried out to clean the data. This study did not find any missing values and hence proceeded to the next step. Thereafter, to meet the purposes of this study and to support the findings, the following statistical techniques were used and applied:

- 1) Cronbach’s alpha coefficient to test the reliability and Spearman correlation analysis to test the validity,
- 2) Descriptive statistics technique which consists of frequency tables, measures of mean and measures of dispersion, to indicate respondents’ degree of agreement or disagreement,
- 3) Pearson correlation analysis method to test the bivariate relationships between measured and latent variables in order to evaluate the hypotheses, and
- 4) Regression analysis to test the hypotheses.

The final step was the presentation and discussion of the analysis results which available in the following sections.

D. Measurement instrument

In this study, the researchers used a five-point Likert scale as a measuring instrument to indicate respondents' degree of agreement or disagreement with each of the last (15) statements, that ranged from “strongly agree (scored as 5) to “strongly disagree” (scored as 1). It is noted that the length of the interval (mane range) used here is (4/5), or about 0.80. It has been calculated on the basis that the five-scale Likert has kept among 4 distances [39]. It has been calculated on the basis that the five-scale Likert has kept among 4 distances [39]. Therefore, the scales and interval (mane range) have been translated in Table 1.

TABLE I. LIKERT – SCALE AND GUIDELINES INTERVAL FOR EVALUATION AND INTERPRETATION

Verbal Interpretation	Scale	Mean Range
Strongly Disagree	1	1.00 – 1.79
Disagree	2	1.80 – 2.59

Neutral	3	2.60 – 3.39
Agree	4	3.40 – 4.19
Strongly Agree	5	4.20-5.00

VI. DATA ANALYSIS AND RESULTS

A. Demographic profile

The demographic profile of the respondents is illustrated in Table 2, which contains the frequency and percentage of demographic characteristics of the respondents studied.

TABLE II. THE MAIN CHARACTERISTICS OF THE STUDY’S SAMPLE

Demographic Questions	Category	Frequency	Percentage
Gender	Male	85	57.8%
	Female	62	42.2%
Age	Less than 25	12	8.2%
	25 to 35	43	29.3%
	36 to 46	52	35.4%
	More than 46	40	27.2%
Education level	Diploma or less	0	0%
	Bachelor degree	75	51.1%
	Master’s degree or more	72	48.9%
Occupational Status	Regional Manager	3	2%
	General Manager	14	9.5%
	Branch Manager	18	12.2%
	Deputy Branch Manager	20	13.6%
	Department Manager	26	17.7%
	Employee	65	44.2%
	Others	1	0.7%
Duration in position	Less than 3 years	32	21.8%
	3 to 6 years	48	32.7%
	7 to 10	38	25.9%
	More than 10 years	29	19.7%

The collected data set in the study was 147 responses. The first question reveals that 57.8% of the samples were males compared to 42.2% females. In addition, 35.4% had age between 36 to 46 years. The educational levels of the respondents revealed that 51.1% had a bachelor degree, while 48.9% had a master’s degree or more. Most of the respondents were from the employees represented by 44.2%. Approximately, 32.7% of the respondents had between 3 to 6 years of experience in the position.

B. Reliability Test

The reliability refers to the probability that the same results would be obtained if the study were to be repeated at a later time, or with different samples [40]. The reliability in this study was calculated through measuring Cronbach’s coefficient alpha. Which is the most common indicator used for assessing the internal consistency of the data post-gathering to reflect the homogeneity of the scale [41]-[43]. This coefficient had values which range between 0 and 1. Hinton et al. in [44], suggest four indicators as a level of

internal reliability (see Table 3). As a rule, if the value of Cronbach's alpha is closer and gets to 1.0, it is better and the scale is considered reliable and acceptable [43], [45], [46]. In accordance with this recommendation, Cronbach's coefficient alpha will be acceptable when its value equal to or greater than 0.7.

In this study, the value of Cronbach's alpha of the 15-items was 0.953, which means that the values higher than the minimum standard of recommended level and resulted in excellent-range scores (see Table 4).

TABLE III. LEVEL OF INTERNAL RELIABILITY

Range	Reliability Result
0.90 and above	Excellent Reliability
0.70 - 0.90	High Reliability
0.50 – 0.70	High Moderate Reliability
0.50 and below	Low Reliability

TABLE IV. CRONBACH'S ALPHA OF THE 15-ITEMS

No. of Items	Reliability	Reliability Result
15	0.953	Excellent Reliability

TABLE V. CRONBACH'S ALPHA OF THE 4 MEASURING DIMENSIONS

Measuring Dimensions	No. of Items	Reliability	Result
SIS	6	0.887	High
Operational Efficiency	3	0.839	High
Information Quality	3	0.843	High
Innovation	3	0.849	High

### C. Validity Test

In addition, the researchers conducted reliability analysis by grouping the items to the five measuring dimensions which they derived from the study conceptual model. The values of Cronbach's alpha of these measuring dimensions varied between 0.887 to 0.849, which means that the alpha values of all measuring dimensions are higher than the minimum standard of recommended level and resulted in high-range scores (see Table 5). This indicates that the measure of reliability for this research revealed an appropriate level of internal consistency.

The validity in quantitative research refers to determine whether the research survey truly measures what it is intended to measure. Moreover, it also measures the truthfulness of the research findings [47]. The type of validity test carried out in this study is construct validity. One of the methods used to establish the construct validity is the correlation analysis by calculating the value of Spearman's rho coefficient. This method tests the relationships between each item and the overall scale of its dimension that it belonged. The interpretation of the validity levels has been translated as follows in Table 6. In general, if the value of correlations coefficients is closer gets to 1.0, the better and it is indicated that the scale is valid and it is measuring its intended concept [48], [49]. Table 7 shows the Spearman's rho correlation coefficient and validity results of the research instrument.

The values of Spearman's rho correlation indicate that the measure of validity for this research revealed an appropriate level of internal consistency. The results confirmed that the questionnaire was a valid measurement tool.

TABLE VI. LEVEL OF INTERNAL RELIABILITY

Range	Validity Result
0.90 and above	Very-High Validity
0.70 - 0.90	High Validity
0.50 – 0.70	Moderate Validity
0.50 and below	Low Validity

TABLE VII. SPEARMAN'S RHO CORRELATION VALUES TO MEASURE VALIDITY

Item no.	Validity	Validity Result
<b>SIS (Independent Variable)</b>		
1	0.835**	High Validity
2	0.813**	High Validity
3	0.750**	High Validity
4	0.821**	High Validity
5	0.768**	High Validity
6	0.819**	High Validity
<b>Operational Efficiency (OE)</b>		
7	0.880**	High Validity
8	0.892**	High Validity
9	0.837**	High Validity
<b>Information Quality (IQ)</b>		
10	0.891**	High Validity
11	0.884**	High Validity
12	0.845**	High Validity
<b>Innovation (INN)</b>		
13	0.874**	High Validity
14	0.864**	High Validity
15	0.893**	High Validity
<b>CA (Dependent Variable)</b>		
OE	0.918**	Very High Validity
IQ	0.936**	Very High Validity
INN	0.922**	Very High Validity

\*\* . Correlation is significant at the 0.01 level (2-tailed).

### D. Descriptive Statistics

Descriptive statistics used to convert the raw data into useful information which can be interpreted to explain a group of dimensions [50].

The descriptive statistics for all items related to the five measuring dimensions included in the study are present in Tables 8 to 12. These tables contain the percentage, frequency distribution, mean and standard deviation scores. They also show the satisfaction of respondents for each item and dimension based on the interpretation of 5-Likert scale (see Table 1).

According to the analysis in Table 8, the result of the average weighted mean of the first measuring dimension is equal to 4.11. This is equivalent to "agree" by five-scale Likert, which means that the respondents have thought that the strategic information systems have an effective role to improve the organizational performance.

According to the analysis in Table 9, the result of the average weighted mean of the first dimension of CA is equal to 4.18. This is equivalent to “agree” by five-scale *Likert*, which means that the respondents have thought that the strategic information systems play an important role in increasing and enhancing the operational efficiency of the work within their organizations. This in turn will help to support and achieve the competitive advantage. Based on this argument, this study asserts that  $H1_a$  is true, which states that there is a positive association between the role of strategic information systems and operational efficiency.

According to the analysis in Table 10, the result of the average weighted mean of the second dimension of CA is equal to 4.24. This is equivalent to “strongly agree” by five-scale *Likert*, which means that the respondents have thought that the strategic information systems play an important role in increasing and enhancing information quality within their organizations. This in turn will help to support and achieve the competitive advantage. Based on this argument, this study asserts that the  $H1_b$  is true, which states that there is a positive association between the role of strategic information systems and information quality.

According to the analysis in Table 11, the result of the average weighted mean of the third dimension of CA is equal to 4.28. This is equivalent to “strongly agree” by five-scale *Likert*, which means that the respondents have thought that the strategic information systems play an important role in increasing and enhancing the innovation of the work within their organizations. This in turn will help to support and achieve the competitive advantage. Based on this argument, this study asserts that the  $H1_c$  is true, which states that there is a positive association between the role of strategic information systems and innovation.

According to the analysis in Table 12, the result of the average weighted mean of the dependent variable (CA) is equal to 4.23. This is equivalent to “strongly agree” by five-scale *Likert*, which means that the respondents have thought that the strategic information systems play an important role in supporting and achieving the competitive advantage. Based on this argument, this study asserts that the main hypothesis is true, which states that there is a positive association between the role of strategic information systems and competitive advantage.

TABLE VIII. DESCRIPTIVE STATISTICS, FOR THE INDEPENDENT VARIABLE [SIS] OF THE STUDY

Independent variable (SIS)		1	2	3	4	5	Total	Mean	S.D	Result
The Bank’s senior management is committed to developing the necessary strategies for its programs, activities and projects related to IS	Frequency	3	4	13	90	37	147	4.05	0.80	Agree
	Percentage	2%	2.7%	8.8%	61.2%	25.3%	100%			
The Bank is keen to align the objectives and strategic plans of IS with overall objectives and strategic plans	Frequency	1	3	16	75	52	147	4.18	0.76	Agree
	Percentage	0.7%	2%	10.9%	51%	35.4%	100%			
The Bank’s senior management is committed to developing and reviewing the strategic plans of IS to identify the necessary changes	Frequency	1	2	25	83	36	147	4.03	0.73	Agree
	Percentage	0.7%	1.4%	17%	56.4%	24.5%	100%			
The Bank’s employees are involved in the development and improvement of IS	Frequency	4	8	12	68	55	147	4.10	0.96	Agree
	Percentage	2.7%	5.4%	8.2%	46.3%	37.4%	100%			
An adequate budget is allocated to build and develop strategic plans of information systems	Frequency	2	4	16	76	49	147	4.13	0.81	Agree
	Percentage	1.4%	2.7%	10.9%	51.7%	33.3%	100%			
The SIS that used in the Bank is a key source to support the processes of decision making	Frequency	1	6	15	68	57	147	4.18	0.83	Agree
	Percentage	0.7%	4.1%	10.1%	46.3%	38.8%	100%			
<b>Average Weighted Mean</b>								<b>4.11</b>	<b>0.65</b>	<b>Agree</b>

TABLE IX. DESCRIPTIVE STATISTICS, FOR FIRST DIMENSION (OPERATIONAL EFFICIENCY)

Dimension 1 (Operational Efficiency)		1	2	3	4	5	Total	Mean	S.D	Result
The ISs that used in Bank help in increasing the coordination and integration of different operations and units	Frequency	1	7	13	78	48	147	4.12	0.81	Agree
	Percentage	0.7%	4.7%	8.8%	53.1%	32.7%	100%			
The ISs that used in Bank help in increasing the operational efficiency and production	Frequency	1	7	17	71	51	147	4.12	0.84	Agree
	Percentage	0.7%	4.7%	11.6%	48.3%	34.7%	100%			
The ISs that used in Bank have ability of to provide information to all units despite the increasing volume and diversity of the various operations	Frequency	1	5	9	71	61	147	4.27	0.78	Strongly Agree
	Percentage	0.7%	3.4%	6.1%	48.3%	41.5%	100%			
<b>Average Weighted Mean</b>								<b>4.18</b>	<b>0.70</b>	<b>Agree</b>

TABLE X. DESCRIPTIVE STATISTICS, FOR SECOND DIMENSION (INFORMATION QUALITY)

Dimension 2 (Information Quality)		1	2	3	4	5	Total	Mean	S.D	Result
The current organizational structure of the Bank helps to the rapid exchange of information and to make better use of it.	Frequency	1	8	9	69	60	147	4.22	0.84	Strongly Agree
	Percentage	0.7%	5.4%	6.1%	46.9%	40.9%	100%			
The information systems that used in Bank make it easy to link administrative units together to get information at any time and properly	Frequency	1	8	9	72	57	147	4.20	0.83	Strongly Agree
	Percentage	0.7%	5.4%	6.1%	49%	38.8%	100%			
The bank provides an integrated knowledge base to provide all administrative units with environmental changes surrounding them	Frequency	1	2	11	71	62	147	4.30	0.73	Strongly Agree
	Percentage	0.7%	1.4%	7.5%	48.3%	42.1%	100%			
<b>Average Weighted Mean</b>								<b>4.24</b>	<b>0.69</b>	<b>Strongly Agree</b>

TABLE XI. DESCRIPTIVE STATISTICS, FOR THIRD DIMENSION (INNOVATION).

Dimension 3 (Innovation)		1	2	3	4	5	Total	Mean	S.D	Result
The ISs that used in Bank help to create opportunities for the creativities and initiatives	Frequency	1	3	13	62	68	147	4.31	0.77	Strongly Agree
	Percentage	0.7%	2%	8.8%	42.2%	46.3%	100%			
The ISs that used in Bank help to accelerate the development of services and diversify them to serve the needs of consumers	Frequency	1	4	11	74	57	147	4.24	0.76	Strongly Agree
	Percentage	0.7%	2.7%	7.5%	50.3%	38.8%	100%			
The ISs that used in Bank facilitate the processes of research and development	Frequency	1	6	11	62	67	147	4.28	0.83	Strongly Agree
	Percentage	0.7%	4.1%	7.5%	42.1%	45.6%	100%			
<b>Average Weighted Mean</b>								<b>4.28</b>	<b>0.68</b>	<b>Strongly Agree</b>

TABLE XII. DESCRIPTIVE STATISTICS, FOR THE DEPENDENT VARIABLE

Dependent variable (CA)	Mean	S.D	Result
<b>Dimension 1:</b> Operational Efficiency	4.18	0.70	Agree
<b>Dimension 2:</b> Information Quality	4.24	0.69	Strongly Agree
<b>Dimension 3:</b> Innovations	4.28	0.68	Strongly Agree
<b>Average Weighted Mean</b>	<b>4.23</b>	<b>0.65</b>	<b>Strongly Agree</b>

TABLE XIII. THE QUANTITATIVE INTERPRETATION OF THE DEGREE OF LINEAR RELATIONSHIP (ALTARES, 2005)

Value of r	Interpretation
±1.00	Perfect Positive (Negative) Correlation
±0.91 - ±0.99	Very High Positive (Negative) Correlation
±0.71 - ±0.90	High Positive (Negative) Correlation
±0.51 - ±0.70	Moderately Positive (Negative) Correlation
±0.31 - ±0.50	Low Positive (Negative) Correlation
±0.01 - ±0.30	Negligible Positive (Negative) Correlation
0.00	No Correlation

E. Hypotheses Test

In order to evaluate the hypotheses of this study, the researchers conducted two tests, which are the correlation analysis and simple linear regression analysis.

- Correlation analysis

To understand the positive relationship which appeared in descriptive analysis better, the researchers conducted correlation analysis by calculating the value of Pearson's Coefficient (r). This analysis helps to examine the strength and direction of the linear relationship between independent variable and dependent variables [51]. According to the results of this analysis (see Table 14) and the quantitative interpretation of the degree of linear relationship (see Table 13), positive and high-level relationship was found among the role of SIS, operational efficiency, information quality and innovation which in turn will help organizations to achieve a higher level of competitive advantages and thus will help them to optimize their organizational performance.

TABLE XIV. PERSON CORRELATION VALUES OF RELATIONSHIP BETWEEN VARIABLES

	CA	Operational Efficiency	Information Quality	Innovation
<b>Role of SIS</b>	0.864**	0.801**	0.819**	0.778**

\*\* Correlation is significant at the 0.01 level (2-tailed).

- Regression analysis

After the correlation analysis results were analyzed, the simple linear regression analysis was performed with competitive advantage (CA), operational efficiency, information quality and innovation as the dependent variables and strategic information systems (SIS) as the independent

variable. Tables 15 to 18 contain the values of ( $R^2$ , F, T,  $\beta$  and .Sig).

As shown in Table 15, the significant of “F” and “ $\beta$ ” values were (0.000) for the first dimension of the CA, which are less than the value of the confidence level (0.05). Furthermore, the value of  $R^2$  also found that the strategic information systems explain (64.2%) of the variance of operational efficiency. In addition, the value of  $\beta$  was 0.801, this means there is a positive strong effect of SIS in Saudi Banking sector on operational efficiency, which in turn will help to achieve a higher level of competitive advantages.

As shown in Table 16, the significant of “F” and “ $\beta$ ” values were 0.000 for the second dimension, which are less than the value of the confidence level (0.05). Furthermore, the value of  $R^2$  also found that the strategic information systems explain 67.1% of the variance of information quality. In addition, the value of  $\beta$  was 0.819, this means there is a positive strong effect of SIS in Saudi Banking sector on information quality, which in turn will help to achieve a higher level of competitive advantages.

As shown in Table 17, the significant of “F” and “ $\beta$ ” values were 0.000 for the third dimension of CA, which are less than the value of the confidence level (0.05). Furthermore, the value of  $R^2$  also found that the strategic information systems explain 60.5% of the variance of innovation. In addition, the value of  $\beta$  was 0.778, this means there is a positive strong effect of SIS in Saudi Banking sector on innovation, which in turn will help to achieve a higher level of competitive advantages.

As shown in Table 18, the significant of “F” and “ $\beta$ ” values were 0.000 for the dependent variable, which are less than the value of the confidence level (0.05). Furthermore, the value of  $R^2$  also found that the strategic information systems explain 74.6% of the variance of competitive advantage. In addition, the value of  $\beta$  was 0.864, this means there is a positive strong effect of SIS in Saudi Banking sector, which in turn will help to achieve a higher level of competitive advantages.

Therefore, the regression analysis showed that the four research hypotheses results were statistically significant. It also, proved that they are supported. The research results are summarized in Table 19.

This means that:

**H1:** There is a significant positive association between the role of Strategic Information Systems (SIS) and Competitive Advantage (CA).

**H1<sub>a</sub>:** There is a significant positive association between the role of Strategic Information Systems (SIS) and operational efficiency.

**H1<sub>b</sub>:** There is a significant positive association between the role of Strategic Information Systems (SIS) and information quality.

**H1<sub>c</sub>:** There is a significant positive association between the role of Strategic Information Systems (SIS) and innovation.

TABLE XV. SIMPLE LINEAR REGRESSION BETWEEN SIS AND OPERATIONAL EFFICIENCY

Dependent Variables	R <sup>2</sup>	F	T	$\beta$
Operational Efficiency	.642	260.4	16.1	.801
		.Sig.= .000		.Sig.= .000

Predictors: (Constant), SIS.

TABLE XVI. SIMPLE LINEAR REGRESSION BETWEEN SIS AND INFORMATION QUALITY

Dependent Variables	R <sup>2</sup>	F	T	$\beta$
Information Quality	.671	295.5	17.2	.819
		.Sig.= .000		.Sig.= .000

Predictors: (Constant), SIS.

TABLE XVII. SIMPLE LINEAR REGRESSION BETWEEN SIS AND INNOVATION

Dependent Variables	R <sup>2</sup>	F	T	B
Innovation	.605	222.3	14.9	.778
		.Sig.= .000		.Sig.= .000

Predictors: (Constant), SIS.

TABLE XVIII. SIMPLE LINEAR REGRESSION BETWEEN SIS AND CA

Dependent Variables	R <sup>2</sup>	F	T	$\beta$
CA	.746	426.9	20.7	.864
		.Sig.= .000		.Sig.= .000

Predictors: (Constant), SIS.

TABLE XIX. HYPOTHESES TESTING RESULTS

HN	Description	Results
H1	There is a positive association between the role of SIS and CA	Supported
H1 <sub>a</sub>	There is a positive association between the role of SIS and operational efficiency.	Supported
H1 <sub>b</sub>	There is a positive association between the role of SIS and information quality	Supported
H1 <sub>c</sub>	There is a positive association between the role of SIS and innovation	Supported

## VII. DISCUSSION

It is clear from the literature review and the results of the study, that there is growing in the term of strategic information systems. This growth emerges with increased the competition intensity, which has made many business organizations adopt information systems with strategic impact. This type of systems called strategic information systems; and it has the ability to store, retrieve and analyze enormous data and information, that are invested to deal with the rapidly changing environment. Moreover, many researches confirmed that the SIS has a greater impact on the operational efficiency, information quality and innovation to gain competitive advantage, which are the dimension of the proposed conceptual model. On the other hand, the results of this study revealed that the weighted mean of these dimensions were 4.18, 4.24, and 4.28, respectively with the average weighted mean for the competitive advantage (4.23). This means that the respondents have thought that the strategic information systems are tools to collect data and provide information to support the strategic decision-making process in order to

achieve higher levels of competitive advantages. Moreover, the general statistical results of the correlation analysis and regression analysis of this study revealed that there is a positive and high-level relationship among the role of SIS with information quality (.Sig=.000, R=.819\*\*), operational efficiency (.Sig=.000, R=.801\*\*) and innovation (.Sig=.000, R=.778\*\*), respectively. This in turn will help organizations to achieve a higher level of competitive advantages (.Sig=.000, R=.864\*). Therefore, the officials of the departments in the Saudi banking sector should play an important role in directing senior managements and motivating them to use the strategic information systems to upgrade their knowledge skills and experience in this field.

Subsequently, it can be summarized that the information systems used in Saudi banking sector considered as strategic information systems. Moreover, they have been a significant role as competitive advantages tool to increase information quality, operational efficiency and innovation, respectively. This in turn indicates that the hypotheses proposed in this research are acceptable.

### VIII. PRACTICAL IMPLICATIONS

The literature review of the strategic information systems has emphasized the fact that the organizations who implement the SIS projects will benefit from the competitive advantages they will receive and consequently will encourage the other competition organizations to adopt one or more of these strategic systems. Empirically, this research has shown that if the organizations develop the SIS projects, managers could achieve great competitive advantage; such increase in the operational efficiency, improve the information quality and encourage the innovation; only if carefully considered and developed. Then possibly this practice might spread the development of the SIS projects among all Saudi organizations. Furthermore, the main practical lesson that has emerged from the analysis presented in this study is:

- The full potential of strategic information systems is unlikely to be realized without an understanding of the significant challenges of developing the SIS and without understanding that the strategic impact is not always realized, even when the systems were strategically aligned and the implementation of IT was a success [18]. Therefore, more focus is needed to develop strategies based on the strengths of the organization that cannot be easily imitated, this will lead to minimize the negative impact of these challenges and achieve the sustainable competitive advantages.

### IX. RESEARCH LIMITATION AND FUTURE RESEARCH

This study was conducted in the banking sector in KSA context, and hence it is hard to decide whether the proposed conceptual model is applicable in other sectors and whether it could be applied in the context of other countries. This is due to several reasons, firstly: the lack of researches in other sectors that confirm the results; and secondly: the great differences in culture, values, beliefs, standards and knowledge between countries. Therefore, re-studying in other sectors or other countries will help to confirm and generalize

the results. However, this study provides interesting opportunities for future SIS research, which need to be investigated and explored further. In addition, there are some recommended directions that could be further initiated. These are as follows:

- Future researchers should consider more comprehensive studies to identify the major consequences of strategic information systems as a tool to achieve the competitive advantages in a public sector, where the measuring of the competitive advantages in the public sector is difficult.
- Future researchers should consider more comprehensive studies to identify the major consequences of strategic information systems as a tool to achieve the competitive advantages with electronic businesses, such as e-government or e-commerce.

This may include studying more dimensions than these proposed in the conceptual model of this study.

### X. CONCLUSION

Based on the findings of the study, the researchers concluded a set of recommendations that represent the importance of activating the role of the strategic information systems by the organizations in general and the Saudi banking sector in particular. The researchers confirmed that the organizations must be interested in SIS because of their impact on the continuity, growth and survival of them in the context of competition. The researchers also emphasized that the organizations should work on the introduction of SIS to improve the efficiency of operations and enhance the quality of information at the lowest cost and fastest time to enable the company to achieve the competitive advantages. In addition, the researchers emphasized the importance of developing specialized administrative units in aspects of strategic information systems that are entrusted with broader and more comprehensive functions than the departments of Management Information Systems (MIS). The researchers recommend that the organizations must involve their staff in training courses to 1) increase their skills and experience; 2) enable them to deal with strategic information systems which are characterized by the rapid development; and 3) enable them to meet the new work requirements. The researchers also strongly recommend that the organizations must pay sustained attention to the dimensions that can achieve, build and maintain the competitive advantages through the activation of strategic information systems roles. Especially that the organizations today operate in the international market, which is characterized by the sharp competition.

Ultimately, the researchers confirmed that the well-structured strategic information systems would create a flexible framework to support the organizational capabilities and to improve the organizational performance. In addition, the SIS offers the opportunity to renovate the current operating environment and create a reliable, scalable and flexible platform to obtain the sustainable competitive advantages, get the brand differentiation, optimize the risk management and improve the decision-making efficiency. It enables the organizations in general and the Saudi banking

sector in particular to gain and sustain competitive advantage in the marketplace by improving the operational efficiency, information quality and innovation of products and services.

However, it requires consideration of five key factors to realize and success of strategic information systems, which are: 1) active support of senior organization management — not just MIS management — in the discovery of strategic opportunities and in the implementation process; 2) integration of planning for the strategic use of information systems into the overall organization strategic planning process; 3) direct reporting by those responsible for strategic use of information systems to the business managers of the area to be affected by the new system; 4) placement of control mechanisms in the hands of these business managers; and 5) readiness for strategic use of information systems, implying successful use of the MIS and technological platform already in place and experience with technological innovation [19]. By focusing on these five key factors, organizations can quickly increase operational efficiency, improve information quality and support innovation. These organizations will also devote themselves for supporting and achieving the sustainable competitive advantage in the coming years.

#### REFERENCES

- [1] M. Rubel, M. Abul Kashem and N. Sultana, "Strategic Information System and Organizational Structure a Case on Meena Bazar Super Shop in Bangladesh." *International Journal of Management (IJM)*, vol.5, no.4, pp. 50-65, 2014.
- [2] H. Gupta, *Management information system*. New Delhi: International Book House, 2011.
- [3] G. Marakas and J. O'Brien *Introduction to information systems*. New York, NY: McGraw-Hill/Irwin, 2013.
- [4] M. Hemmatfar, M. Salehi and M. Bayat, "Competitive Advantages and Strategic Information Systems". *IJBM*, vol.5, no.7, pp.158-169, 2010.
- [5] A. Basahel and I. Zahir, "Examining the Strategic Benefits of Information Systems: A Global Case Study," In *European, Mediterranean & Middle Eastern Conference on Information Systems (EMCIS)*. Abu Dhabi, UAE, 2010, pp. 1-17.
- [6] E. Ankrah, "The Impact of Information System Strategy on Bank Performance in Ghana," Ph.D thesis, University of Ghana, Legon, Feb.2014.
- [7] V. Olufemi, *Strategic Information System Planning (SISP) in Accenture*, London : University of Roehampton, 2015.
- [8] T. Liang and M. Tang., "VAR analysis," *ACM SIGMIS Database*, vol.23, no.1, pp.27-35, 1992.
- [9] F. Al-Aboud, "Strategic Information Systems Planning: A Brief Review," *International Journal of Computer Science and Network Security*, vol.11, no.5, pp.179-183, 2011.
- [10] S. Maharaj and I. Brown, "The impact of shared domain knowledge on strategic information systems planning and alignment," *S Afr j inf manag*, vol.17, no.1, 2015.
- [11] A. Issa-Salwe, M. Ahmed, K. Aloufi and M. Kabir, "Strategic Information Systems Alignment: Alignment of IS/IT with Business Strategy," *Journal of Information Processing Systems*, vol.6, no.1, pp.121-128, 2010.
- [12] C. Gaines, D. Hoover, W. Foxx, T. Matuszek and R. Morrison, "Information systems as a strategic," *Journal of Management and Marketing Research*, 2012.
- [13] R. Hoque, E. Hossin and W. Khan, "Strategic Information Systems Planning (SISP) Practices In Health Care Sectors Of Bangladesh," *European Scientific Journal*, vol.12, no.6, pp.307-312, 2016.
- [14] M. Shirazi and J. Soroor, "An intelligent agent-based architecture for strategic information system applications," *Knowledge-Based Systems*, vol.20, no.8, pp.726-735, 2007.
- [15] K. Laudon and J. Laudon, *Essentials of management information systems*, Upper Saddle River, NJ: Prentice Hall, 2005.
- [16] M. Martinsons and A. Leung, "Strategic information systems: a success factors model," *Int. J. Services Technology and Management*, vol.3, no.4, pp.398-416, 2002.
- [17] R. Kini, "Strategic Information Systems," *Information Systems Management*, vol.10, no.4, pp.42-45, 1993.
- [18] V. Arvidsson, J. Holmström and K. Lyytinen, "Information systems use as strategy practice: A multi-dimensional view of strategic information system implementation and use," *The Journal of Strategic Information Systems*, vol.23, no.1, pp.45-61, 2014.
- [19] V. Zwass, *Foundations of information systems*. Boston, Mass.: Irwin/McGraw-Hill, 1998.
- [20] C. Ehmke and MS, "Strategies for competitive advantage," *Department of Agricultural and Applied Economics, University of Wyoming*, 2008.
- [21] Z. Awino, "Business Strategy, Internal Resources, National Culture And Competitive Advantage: A Critical Review," In *1st DBA-Africa Management Review International Conference*, 2015, pp.66-81.
- [22] D. Ketchen, W. Rebarick, G. Hult and D. Meyer, "Best value supply chains: A key competitive weapon for the 21st century," *Business Horizons*, vol.51, no.3, pp.235-243, 2008.
- [23] P. Jindal, "Operational efficiency of selected public-sector banks in India". Ph.D Thesis, Maharishi Markandeshwar University, Nov. 2013.
- [24] N. Hussein, "Procurement Performance and Operational Efficiency in Telecommunication Industry in Kenya," Master thesis. University of Nairobi, Nov. 2014.
- [25] T. Hovelja, A. Rožanec and R. Rupnik, "Measuring the Success of the Strategic Information System Planning in Enterprises in Slovenia" *Management*, vol.15, pp. 25-46, 2010.
- [26] G. Philip, "IS Strategic Planning for Operational Efficiency," *Information Systems Management*, vol.24, no.3, pp.247-264, 2007.
- [27] N. Gorla, T. Somers and B. Wong, "Organizational impact of system quality, information quality, and service quality," *The Journal of Strategic Information Systems*, vol.19, no.3, pp.207-228, 2010.
- [28] Y. Al-Mamary, A. Shamsuddin and N. Aziati, "The Relationship between System Quality, Information Quality, and Organizational Performance," *International Journal of Knowledge and Research in Management & E-Commerce*, vol.4, no.3, pp.7-10, 2014.
- [29] B. Lam, "An Empirical Study of Factors Impacting School Management Based on Schools in Hoa Binh Province of Vietnam," Master thesis, Shu-Te University, June 2011.
- [30] y. Lee, D. Strong, B. Kahn and R. Wang, "AIMQ: a methodology for information quality assessment," *Information & Management*, vol.40, no.2, pp.133-146, 2002.
- [31] A. Madapusi, "ERP Information Quality and Information Presentation Effect on Decision Making," *SWDSI 2008 Proceedings*, pp.628-633, 2008.
- [32] K. Huang, R. Dyerson, L. Wu and G. Harindranath, "From Temporary Competitive Advantage to Sustainable Competitive Advantage," *Brit J Manage*, vol.26, no.4, pp.617-636, 2015.
- [33] A. Afuah, *Strategic innovation: New Game Strategies for Competitive Advantage*. New York: Routledge, 2009.
- [34] M. Rogers, "The Definition and measurement of innovation," *Melbourne Institute Working Paper*, vol.10, no.98, 1998.
- [35] C. Martín-de, M, Delgado-Verde, J. Navas-López and J Cruz-González, "The moderating role of innovation culture in the relationship between knowledge assets and product innovation," *Technological Forecasting and Social Change*, vol.80, no.2, pp.351-363, 2013.
- [36] H. Urbancova, "Competitive Advantage Achievement through Innovation and Knowledge," *JOC*, vol.5, no.1, pp.82-96, 2013.
- [37] D. Nulty, "The adequacy of response rates to online and paper surveys: what can be done?," *Assessment & Evaluation in Higher Education*, vol.33, no.3, pp.301-314, 2008.
- [38] D. Davis, *Business research for decision making*, 1st e., Thomson/South-Western: Mason, OH, 2005.
- [39] R. Likert, "A technique for the measurement of attitudes," *Journal of Systems Management*. Vol.22, no.140, pp. 44-53, 1932.

- [40] B. Ticehurst and A. Veal, *Business research methods: a managerial approach*, Longman: French Forest, NSW, 2000.
- [41] M. Litwin, *How to Measure Survey Reliability and Validity*, 1st ed., Thousand Oaks: SAGE Publications, Incorporated, 1995.
- [42] E. Carmines, and R. Zeller, *Reliability and validity assessment*, 1st ed., Newbury Park, Calif. [u.a.]: Sage Publ, 2008.
- [43] M. Tavakol, and R. Dennick, "Making sense of Cronbach's alpha," *International Journal Of Medical Education*, vol.2, pp. 53-55, 2011.
- [44] P. Hinton, C. Brownlow, L. McMurvay and B. Cozens, *SPSS Explained. East Sussex*, England: Routledge Inc, 2004.
- [45] J. Pallant, *The SPSS survival manual: A step-by-step guide to data analysis using SPSS for Windows (version 10)*, St Leonards, NSW: Allen & Unwin., 2001.
- [46] D. Straub, M. Boudreau and D. Gefen, Validation Guidelines for IS Positivist Research. *Communications of the AIS*, vol.13, no.1, pp. 380-427, 2004.
- [47] C. Kothari and G. Garg, *Research methodology*, 1st ed., New Delhi: New Age International (P) Limited, 2016.
- [48] J. Hair, *Multivariate data analysis*, 1st ed., Upper Saddle River [u.a.]: Prentice Hall, 2010.
- [49] C. Robson and K. McCartan, *Real world research*. Chichester: Wiley, 2016.
- [50] A. Bryman and E. Bell, *Business research methods*, 4th ed., Oxford: Oxford University Press, 2012.
- [51] J. Pallant., *SPSS survival manual*, 1st ed., Maidenhead, Berkshire, England: McGraw-Hill Education, 2016.

# The Optimization of Query Processing in Seabase Cloud Databases based on CCEVP Model

Abdulkadir ÖZDEMİR

Faculty of Social Sciences, Ataturk University  
Erzurum, Turkey

Hasan Asil

Faculty of Computer, Azarshahr Branch,  
Islamic Azad University, Azarshahr, Iran

**Abstract**—A cloud database is a database usually installed on cloud computing software platforms. There are several methods for query processing in cloud databases. This study tried to optimize query processing in the SeaBase cloud database and reduce query processing time. This method used adaptability for optimization. This method was designed for cloud-based databases. The algorithm is composed of three components: 1) multi cloud query separator; 2) query similarity detector based on the execution plan; and 3) replacement policy. This method is implemented as a system for a fully object-oriented simulation. The system is added to the SeaBase as an agent. The evaluation result show that this method reduced response time by 1.9 percent.

**Keywords**—SeaBase; optimization; query processing; database; adaption

## I. INTRODUCTION

Database management systems are software packages that can be used to create and maintain one or more databases. However, with the rise of cloud computing, database management systems have become a new kind of service with unique advantages. In these services, DBMS is a part of a larger service which is likely to be more effective in terms of results and assigned tasks [1].

A cloud database is a database usually installed on cloud computing software platforms. Using a virtual machine, users can independently launch databases on cloud, or they can purchase an account to access database services maintained by cloud database providers [2].

SeaBase is an implementation based on cloud computing. Based on CCEVP model, it can convert different data types into one. In fact, SeaBase is a relational cloud database which can merge a pair of databases together [3]. Like SQL server, DB2, Sybase, MySQL, and other similar databases, SeaBase was designed in order to integrate data taken from several heterogeneous databases and provide users with them in a unified way [4]. Fig. 1 indicates the structure of SeaBase based on CCEVP model.

The CCEVP model uses three layers: 1) physical; 2) virtual; and 3) effective. The physical layer is a set of multisource physical tables from similar or dissimilar databases. The virtual layer is a set of relationship schemas determined by the SeaBase users. The effective layer allows users to create a unified vision to the SeaBase.

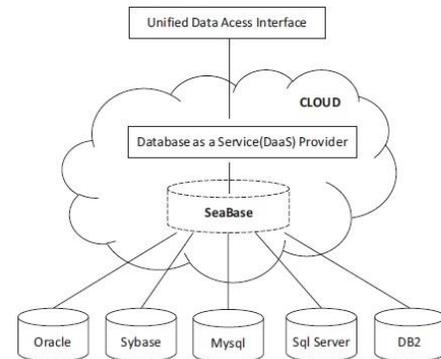


Fig. 1. CCEVP Model (Cloud Computing-based Effective-Virtual-Physical) [1].

## II. OPTIMIZING QUERY PROCESSING IN CLOUD DATABASE EASE OF USE

There are several methods for query processing in cloud databases. Many of these methods have offered new technologies to optimize query processing in the database [3]. Some of these methods use replication for query processing and accelerate the process by data sampling. Some methods use traditional methods for query processing in the database. Some methods attempted to optimize the execution plans, which are known as Selinger methods [4], [5].

Most of these methods use a special procedure (Selinger-Style) for optimization, and generally after query processing and query execution, the plan optimized for query processing is eliminated. But today, in addition to these methods, other methods are also being offered to optimize query processing in the database [6], [7]. But the question is that whether the produced optimized plan (or the frequent queries sent to the SeaBase) can be used for executing subsequent queries. As we know, in the application-related databases, usually the queries sent to the database have high adaptability power because in these systems, the queries sent to the database have the same structure and are adapted as soon as possible. This method can be used to optimize the SeaBase.

## III. THE PROPOSED METHOD

The aim of this study was to optimize query processing in the database model. Matching techniques were used to

optimize processing. This method was proposed in order to use the optimal schemas generated for the execution of next queries or the ones frequently sent to the database. An agent was added to this model for implementation [8].

Combining the technologies of optimizing queries and agents in this algorithm, a multi-agent system was proposed. Using information collection technology based on processing users' queries, this system tries to provide users with an adaptable environment based on the query types and how they are made [9].

Given the fact that requests are sent to the system by applications or users, the majority of queries have the same structure. They are repeated over time. Therefore, this paper was intended to propose a method by which the database could identify the queries of the same type over time. Using a specific schema of execution, it was also intended to identify the highly frequent queries sent to the database and answer them. Put another way, the database was to be matched so that the queries would be processed with the prepared execution schemas at a lower cost. In fact, it was meant to decrease the number of steps required for processing highly frequent queries sent to the database.

In this algorithm, an agent was added to SeaBase so that a matched cloud database would be generated. Decreasing the number of steps required for processing the highly frequently queries sent to the cloud database, this algorithm tried to increase the optimization of query processing in cloud databases.

This research uses the previous heterogeneous distributed database query processing method and develops it for SeaBase [4], [8].

The algorithm uses a method for optimizing query processing in the heterogeneous distributed databases. This method was designed for cloud-based databases. The algorithm is composed of three components:

- Multi Cloud query Separator
- Query similarity detector based on the execution plan
- Replacement policy

The main objective of this approach is to identify the most frequent instructions sent to the cloud database and store their execution plans in the system so that in case of a request to the database, the same execution plan is used for query execution. The separator part separates instructions and the instructions whose execution plan has not high cost. The query similarity detector is used to identify similar instructions. The replacement policy detects the most frequent instructions sent to the SeaBase and stores their execution plans. The following algorithm shows the structure of this method (Fig. 2).

#### A. Separator of distributed instructions

Distributed instructions are instructions which include several sub-queries and receive information from several DBMSs.

#### Query processing optimization in SeaBase

1. Begin
2. Examine query by separator(can Separate distributed query)
3. Produce query execution plan if query is one of exceptions
4. If it's not an exception check it's execution plan availability in system by similarity recognizer
  - 4.1 If execution plan exists select it
  - 4.2 otherwise,
  - 4.3 send it in order to producing execution plan
5. executing plan for replying to query
6. Check whether it's the time for substituting or not?
  - 6.1 If so, do substitution
7. END

Fig. 2. Suggested algorithm.

The purpose of this function is to identify functions that based on assessment, need more time for execution or receive information from several databases. Different queries are sent to the SeaBase, and the database needs cost to respond to queries depending on their type and structure. As mentioned, three layers are used to execute the instructions sent to the database: physical, virtual and effective. Based on the layer, the separator detects instructions that require the use of several databases. The separator aims to identify the instructions that use multiple databases and need a link.

#### B. The query similarity detector based on execution plan

Any query for execution in the cloud database requires the same steps used in a non-cloud database. Any query for execution must have a specific plan. In applications, requests are usually sent to the database with a specific format and different parameters. The purpose of this section is to identify queries with similar plans.

To make adaptive the query processing in the database, we need a part in the proposed system that can compare the sent queries and identify similar queries. For example, consider the following two queries.

```
SELECT *
FROM   tblKala INNER JOIN
tblHavaleKala ON tblKala.KalaiD = tblHavaleKala.KalaID
INNER JOIN
tblHavale ON tblHavaleKala.HavaleID =
tblHavale.HavaleIDwhere kalaid=20

SELECT*
FROM   tblKala INNER JOIN tblHavaleKala ON
tblKala.KalaiD = tblHavaleKala.KalaID INNER JOIN
tblHavale ON tblHavaleKala.HavaleID =
tblHavale.HavaleIDwhere kalaid=31
```

As can be seen, these two instructions request information on products 20 and 31 from the database. But the two instructions are the same and can be executed with the same plan. This part of the system should be able to detect such instructions.

### C. Replacement policy

In the algorithm, a set of frequent execution plans must be kept in the agent and in case of request, the same query request should be used. The replacement policy is used to create and update the set. An important part of this research is to determine the replacement action is done how, when and with what policy.

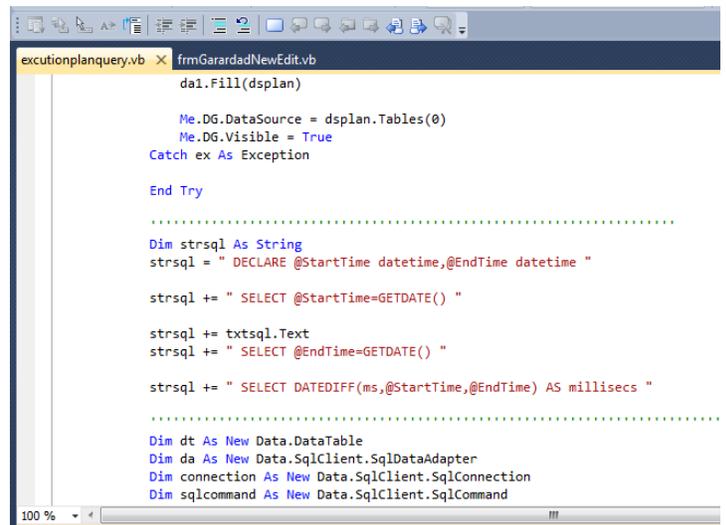
As know adding an agent to SeaBase, which always adapts queries, constrains cost to system. For adapting system the method examines sent queries to database for a while and the execution plan of similar frequent queries is substituted in database. (It's considered that only queries receipted by separator will be sent to this part). The time between two adaptation said that this time is calculated by value of adapted queries and dynamically. It means that whatever score goes up, adaptable queries will stay in the system for longer time and if adapting enjoys low score they will stay in the system for shorter time. It was found that on the long run, the increasing score of adapted queries have increase this time. The method does adapting on queries sent to database in busy hours. The method saves queries in busy hours and in quiet hours it will does adapting on these queries when it's time to adapt.

Now about the ways of adapting, at first create a bank of sent queries then if sent query was similar to one of available queries in database we increase the weight of query and also if sent query was not available in the bank the method adds it to databank and continues adapting. After adapting the method saves queries with high scores. Ways of saving queries in database follows a distinct format and standard in order to constrain less cost when the queries are examined.

## IV. ASSESSMENT AND PRACTICAL RESULTS

Several methods are currently used to measure the performance of the database system. One of the most common methods among the above methods is runtime in the system. Runtime is the time from the sending moment to the system response. This study tries to identify the most frequent queries sent to the database and keep their execution plans for executing subsequent queries. In fact, this method tries to make the query processing in the database adaptive.

For assessment, this method is implemented as a system for a fully object-oriented simulation. The system is added to the SeaBase as an agent. Then the results of execution using this method are compared with the SeaBase without this agent. Furthermore, we need the desired data based on relationship dependence. For this purpose, the SQL Toolbelt database and simulator is used to create data and determine the table dependence. The.NET and the SQL API functions are used to implement the algorithm and make comparisons. The following Fig. 3 shows some of the code in this system:



```
da1.Fill(dsplan)

Me.DG.DataSource = dsplan.Tables(0)
Me.DG.Visible = True
Catch ex As Exception

End Try

.....

Dim strSQL As String
strSQL = " DECLARE @StartTime datetime,@EndTime datetime "

strSQL += " SELECT @StartTime=GETDATE() "

strSQL += txtsql.Text
strSQL += " SELECT @EndTime=GETDATE() "

strSQL += " SELECT DATEDIFF(ms,@StartTime,@EndTime) AS millsecs "

.....

Dim dt As New Data.DataTable
Dim da As New Data.SqlClient.SqlDataAdapter
Dim connection As New Data.SqlClient.SqlConnection
Dim sqlCommand As New Data.SqlClient.SqlCommand
```

Fig. 3. Part of the simulation.

After simulation of the system, the following results will be provided.

- The query runtime cost in a normal manner.
- This cost is equal to the time required for the SeaBase query processing and respond to the user. This cost is assessed without adding the agent to the system.
- The cost of the proposed algorithm execution.
- After adding the agent to the SeaBase, the adaptability cost and the query execution cost must be added up and evaluated. The algorithm execution cost is the adaptability cost.
- The execution cost of the adapted query as an execution plan.
- This cost is the execution cost of query with the help of agent. It is worth mentioning that with regard to the adaptability of some queries, the cost of some queries is normal and some less.

After obtaining the above results, the second and third costs are added up and compared with the first cost.

In the algorithm, the times required for executing the queries sent to the database are compared in adaptive and non-adaptive databases. Fig. 4 shows the time required to respond to the adaptive and non-adaptive queries per day. A cloud database with adaptive queries is called adaptive cloud base.

This diagram shows the total time required for executing adaptive queries in the database as well as the total time for executing adaptive queries in the non-adaptive mode. It should be noted that in this figure, the adaptability cost is not currently added to the above calculations because the system is not adaptive at any time and will do this action only at certain times of low traffic.

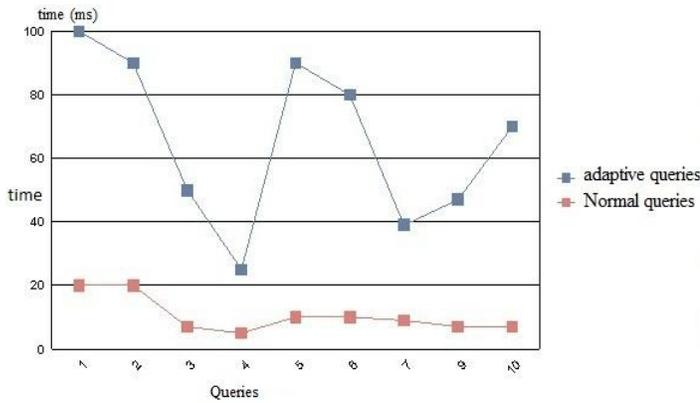


Fig. 4. Report of response time per day for adaptive queries.

However, these costs will be taken into account in the next assessments.

Fig. 5 shows the reduced time of executing adaptive queries. These queries in the SeaBase are queries which have become adaptive. Obviously, due to making high-traffic queries adaptive, this method reduces the server workload at times of high traffic.

The first row of Table 1 represents the total time for responding to adaptive queries and reduced time of response time for all adaptive queries sent to the database. It also shows the cost of making queries adaptive.

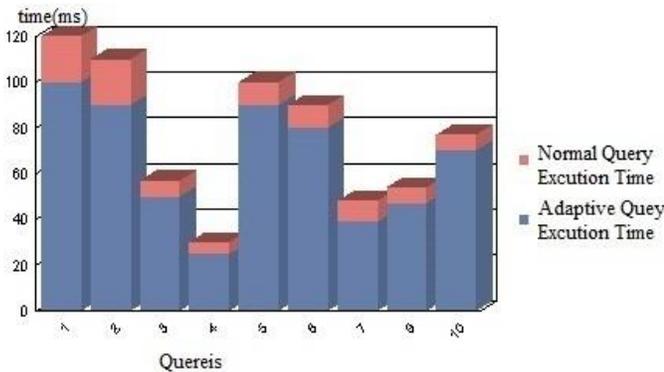


Fig. 5. Report of reduced cost of adapted query.

TABLE I. TOTAL SYSTEM EVOLUTION

Row	Type of queries	Decrease response time	Total Execution time
1	Distributed queries(Join)	15%	100%
2	All queries	1.9%	100%

In this system, when sending queries to the database, the query separator separates some queries and blocks their way to the database. The second row represents time and cost for all queries sent to the database plus adaptability cost.

As shown in Table 1, the system reduced response time by 1.9 percent.

## V. CONCLUSION

The increase in data volume in many applications and the need for their calculations are the database challenges. Cloud computing and the use of SeaBase databases are a solution to integrate a variety of DBMSs and integrated access to tables in databases. This study tried to optimize query processing in the SeaBase cloud database and reduce query processing time. This method used adaptability for optimization. The purpose of this method is to make adaptive the execution plans of high-traffic queries sent to the SeaBase. For adaptability, this method uses three parts: separator, similarity detector and replacement policy. This method is added to the database as an agent. The results show that the system optimizes query processing in the database and reduces response time by one percent. Based on the replacement policy, this method also reduces workload. In the future, response time can further decrease by changing the replacement policy.

## REFERENCE

- [1] Shusheng Guo ; State Key Lab. of Comput. Syst. & Archit., China ; Zimu Yuan ; Li Zha ; Zhiwei Xu, "SeaBase: An Implementation of Cloud Database", 10th International Conference on Semantics, Knowledge and Grids (SKG), Beijing, 2014
- [2] Waleed Al Shehri, CLOUD DATABASE DATABASE AS A SERVICE, International Journal of Database Management Systems ( IJDM ) Vol.5, No.2, April 2013
- [3] Clayton Maciel Costa, Adaptive Query Processing in Cloud Database Systems, IEEE International Conference on Cloud and Green Computing, 2013
- [4] Ennaz zafarani , Mohammad\_Reza Feizi\_Derakhshi , Hasan Asil , Amir asil "Presenting a New Method for Optimizing Join Queries Processing in Heterogeneous Distributed Databases" , WKDD2010, Phuket , Thailand , 9-10 January, 2010.
- [5] Mohammad\_Reza Feizi\_Derakhshi , Hasan Asil , Amir Asil,ennaz zafarani "Optimizing Query Processing in Practical Software Database by Adapting" WKDD2010, Phuket , Thailand , 9-10 January, 2010.
- [6] Mohammad\_Reza Feizi\_Derakhshi, Hasan Asil, Amir Asil "Proposing a New Method for Query Processing Adaption in Data Base " WCSET 2009: World Congress on Science, Engineering and Technology Dubai, United Arab Emirates VOLUME 37, January 28-30, 2009 ISSN 2070-3740
- [7] Amol Deshpande, Zachary Ives, and Vijayshankar Raman. Adaptive query processing. Foundations and Trends in Databases, 1(1), 2007.
- [8] Mohammad\_Reza Feizi\_Derakhshi , Hasan Asil , Amir Asil,ennaz zafarani "Practical Software Query Optimizing by Adapting Why and How?" Australian Journal of Basic and Applied Sciences, jourdan , January, 2010
- [9] Agent Working Group, "Agent Technology Green Paper". OMG Documentagent/00-09-01 Version 1.0, 2000.

# A Copula Statistic for Measuring Nonlinear Dependence with Application to Feature Selection in Machine Learning

Mohsen Ben Hassine

Department of Computer Science,  
University of El Manar, Tunis,  
Tunisia

Lamine Mili

Bradley Department of Electrical and  
Computer Engineering, Northern  
Virginia Center, Virginia Tech, Falls  
Church, VA 22043, USA

Kiran Karra

Bradley Department of Electrical and  
Computer Engineering, VTRC-A,  
Arlington, VA 22203, USA

**Abstract**—Feature selection in machine learning aims to find out the best subset of variables from the input that reduces the computation requirement and improves the predictor performance. This paper introduces a new index based on empirical copulas, termed as the Copula Statistic (CoS) to assess the strength of statistical dependence and for testing statistical independence. It shows that this test exhibits higher statistical power than other indices. Finally, applying the CoS features selection in machine learning problems, which allow a demonstration of the good performance of the CoS.

**Keywords**—Copula; multivariate dependence; nonlinear systems; feature selection; machine learning

## I. INTRODUCTION

Measures of statistical dependence among random variables and signals are paramount in many scientific applications: engineering, signal processing, finance, biology and machine learning to cite a few. They allow one to find clusters of data points and signals, test for independence to make decisions, and explore causal relationships. The classic measure of dependence is provided by the correlation coefficient, which was introduced in 1895 by Karl Pearson. Since it relies on moments, it assumes statistical linear dependence. However, in biology, ecology and finance, and other fields, applications involving nonlinear multivariate dependence prevail. For such applications, the correlation coefficient is unreliable. Hence, researchers have initiated many proposals in order to address this deficiency [1]-[5]. Reshef *et al.* [6], [7] introduced the Maximal Information Coefficient (MIC) and later the Total Information Coefficient (TIC), Lopes-Paz *et al.* [8] proposed the Randomized Dependence Coefficient (RDC), and Ding *et al.* [9], [10] put forth the Copula Correlation Coefficient (Ccor). Additionally, Székely *et al.* [11] proposed the distance correlation (*dCor*). These metrics are able to measure monotonic and non-monotonic dependencies between random variables, but each has strengths and shortcomings [12]-[18]. Feature selection in machine learning is a typical battlefield to appraise the quality and the reliability of a dependence index, it is to find out the best subset of variables (from the input) that reduces the computation requirement and feed up the predictor algorithm for optimal performance [19], [20].

In this paper, a new index based on copulas, termed the Copula Statistic (*CoS*), for measuring the strength of nonlinear statistical dependence and for testing for statistical independence is introduced. The *CoS* ranges from zero to one and attains its lower and upper limit for the independence and the functional dependence case, respectively.

Monte Carlo simulations are carried out to estimate bias and standard deviation curves of the *CoS*, to assess its power when testing for independence. The simulations reveal that for large sample sizes, the *CoS* is approximately normal and approaches Pearson's  $\rho_P$  for the Gaussian copula and Spearman's  $\rho_S$  for many copulas. The *CoS* is shown to exhibit strong statistical power in various functional dependencies as compared to many other indices. Finally, the *CoS* is applied to feature selection problem to unveil bivariate dependence.

The paper is organized as follows. Section II proves two new and essential theorems on copulas used to derive the *CoS* index. Section III introduces a relative distance function and proves several of its properties. Section IV defines the *CoS* and provides an algorithm that implements it. Section V investigates the statistical properties of the *CoS* and treats the case of bivariate dependence. Section VI compares the performance of the *CoS* with the *dCor*, *RDC*, *Ccor*, and the *MICe* in measuring bivariate functional and non-functional dependencies between synthetic datasets. It also shows how the *CoS* can unveil statistical dependence in real datasets of breast tumor and proceed with an in-depth analysis in order to find out the best feature subset for this problem.

## II. BIVARIATE COPULA

In the following, attention is restricted to two-dimensional copulas to develop a statistical index, the *CoS*, in the bivariate dependence case. To define the *CoS* of two continuous random variables  $X$  and  $Y$  with copula  $C(u, v)$ , three definitions of bivariate dependencies are provided, from weaker to stronger versions, as introduced by Lehmann [21]. Then, three theorems are stated which help build the foundation for the *CoS*.

**Definition 1:** Two random variables,  $X$  and  $Y$ , are said to be concordant (or discordant) if they tend to simultaneously take large (or small) values.

A more formal definition is as follows. Let  $X$  and  $Y$  be two random variables taking two pairs of values,  $(x_i, y_i)$  and  $(x_j, y_j)$ .  $X$  and  $Y$  are said to be concordant if  $(x_i - x_j)(y_i - y_j) > 0$ ; they are said to be discordant if the inequality is reversed.

**Definition 2:** Two random variables,  $X$  and  $Y$ , defined on the domain  $\mathfrak{D} = \text{Range}(X) \times \text{Range}(Y)$  are said to be Positively Quadrant Dependent (PQD) if

$$P(X \leq x, Y \leq y) \geq P(X \leq x) P(Y \leq y),$$

that is,  $C(u, v) \geq \Pi(u, v)$  and Negative Quadrant Dependent (NQD) if

$$P(X \leq x, Y \leq y) < P(X \leq x) P(Y \leq y),$$

that is,  $(u, v) \leq (u, v)$  for all  $(x, y) \in \mathfrak{D}$ .

**Definition 3:** Two random variables,  $X$  and  $Y$ , are said to be comonotonic (respectively countermonotonic) if  $Y = f(X)$  almost surely and  $f(\cdot)$  is an increasing (respectively a decreasing) function.

In short, two random variables are monotonic if they are either comonotonic or countermonotonic.

**Theorem 1:** (Fréchet [13]: Let  $X$  and  $Y$  be two continuous random variables. Then,

a)  $X$  and  $Y$  are comonotonic if and only if the associated copula is equal to its Fréchet-Hoeffding upper bound, that is,

$$C(u, v) = M(u, v) = \text{Min}(u, v);$$

b)  $X$  and  $Y$  are countermonotonic if and only if the associated copula is equal to its Fréchet-Hoeffding lower bound, that is,  $C(u, v) = W(u, v) = \text{Max}(u+v-1, 0)$

c)  $X$  and  $Y$  are independent if and only if the associated copula is equal to the product copula, that is,  $C(u, v) = \Pi(u, v) = uv$ .

In the following theorems and corollaries, it is assumed that  $X$  and  $Y$  are continuous random variables and related via a function  $f(\cdot)$ , that is,  $Y=f(X)$ , where  $f(\cdot)$  is continuous and differentiable over the range of  $X$ .

**Theorem 2:** Let  $X$  and  $Y$  be two continuous random variables such that  $Y = f(X)$  almost surely, and let  $C(u, v)$  be the copula value for the pair  $(x, y)$ . The function  $f(\cdot)$  has a global maximum at  $(x_1, y_{max})$  with a copula value  $C(u_1, v_1)$  or a global minimum at  $(x_2, y_{min})$  with a copula value  $C(u_2, v_2)$  if and only if

$$a) C(u_1, v_1) = M(u_1, v_1) = W(u_1, v_1) = \Pi(u_1, v_1) = u_1; \quad (1)$$

$$b) C(u_2, v_2) = M(u_2, v_2) = W(u_2, v_2) = \Pi(u_2, v_2) = 0. \quad (2)$$

The proof of Theorem 2 is given in the appendix. For a general definition of the copula, the reader is referred to Nelsen [13].

**Corollary 1:** Let  $X$  and  $Y$  be two continuous random variables such that  $Y = f(X)$ , almost surely. If  $f(\cdot)$  is a periodic function, then (1) and (2) holds true at all the global maxima and global minima, respectively.

The proof of Corollary 1 directly follows from Theorem 2. This corollary is demonstrated in Fig. 1, which displays the graph of the projections on the  $(u, C(u, v))$  plane of the

empirical copula  $C(u, v)$  associated with a pair  $(X, Y)$ , where  $X$  is uniformly distributed over  $[-1, 1]$ , and  $Y = \sin(2\pi X)$ . It is observed that at each one of the four optima of the sine function,  $C(u, v) = M(u, v) = W(u, v) = \Pi(u, v)$ .

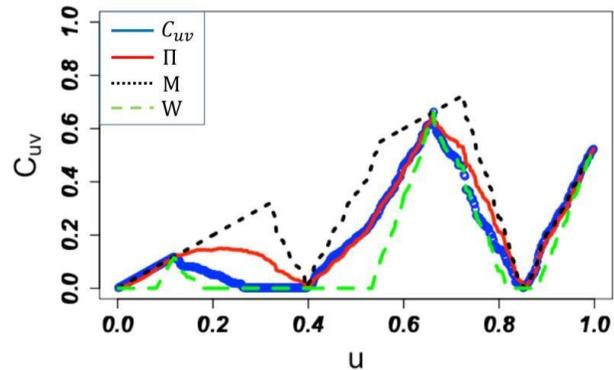


Fig. 1. Graph (in blue dots) of the projections on the  $(u, C(u, v))$  plane of the empirical copula  $C(u, v)$  associated with a pair of random variables  $(X, Y)$ , where  $X \sim U(-1, 1)$  and  $Y = \sin(2\pi X)$ . The  $u$  coordinates of the data points are equally spaced over the unity interval. Similar graphs are shown for the  $M(u, v)$ ,  $W(u, v)$  and  $\pi(u, v)$  copulas.

**Theorem 3:** Let  $X$  and  $Y$  be two continuous random variables such that  $Y = f(X)$ , almost surely where  $f(\cdot)$  has a single optimum and let  $C(u, v)$  be the copula value for the pair  $(x, y)$ . Then,  $C(u, v) = M(u, v)$  if and only if  $df(x)/dx \geq 0$  and  $C(u, v) = W(u, v)$  otherwise.

The proof of Theorem 3 is provided in the appendix. Theorem 3 is illustrated in Fig. 2.

### III. THE RELATIVE DISTANCE FUNCTION

A metric of proximity of the copula to the upper or the lower bounds with respect to the  $\Pi$  copula is defined and its properties are investigated.

**Definition 4:** The relative distance function,  $\lambda(C(u, v))$ :  $[0, 1] \rightarrow [0, 1]$ , is defined as

- a)  $\lambda(C(u, v)) = (C(u, v) - uv) / (\text{Min}(u, v) - uv)$  if  $C(u, v) \geq uv$ ;
- b)  $\lambda(C(u, v)) = (C(u, v) - uv) / (\text{Max}(u+v-1, 0) - uv)$  if  $C(u, v) < uv$ .

A graphical illustration for the relative distance is shown in Fig. 3.

**Theorem 4:**  $\lambda(C(u, v))$  satisfies the following properties:

- a)  $0 \leq \lambda(C(u, v)) \leq 1$  for all  $(u, v) \in \mathbf{I}^2$ ;
- b)  $\lambda(C(u, v)) = 0$  for all  $(u, v) \in \mathbf{I}^2$  if and only if  $C(u, v) = uv$ ;
- c) If  $Y = f(X)$  almost surely, where  $f(\cdot)$  is monotonic, then  $\lambda(C(u, v)) = 1$  for all  $(u, v) \in \mathbf{I}^2$ ;
- d) If  $Y = f(X)$  almost surely, then  $\lambda(C(u, v)) = 1$  at the global optimal points of  $f(\cdot)$ .

**Proof:** Property a) follows from Definition 5 and (3) while properties b), c) and d) follow from Definition 5 and Theorem 1 and 2. ■

**Corollary 2:** If  $Y = f(X)$  almost surely, where  $f(\cdot)$  has a single optimum, then  $\lambda(C(u, v)) = 1$  for all  $(u, v) \in \mathbf{I}^2$ .

#### IV. THE COPULA STATISTIC

The empirical copula is first defined; then, the copula statistic is introduced, and finally, an algorithm that implements it is provided. One possible definition for the CoS is the mean of  $\lambda C(u,v)$  over  $\mathbf{I}^2$ , that is,  $CoS(X,Y) = E[\lambda(C(u,v))]$ . However, according to Theorems 5 and 6,  $CoS \leq 1$  for functional dependence with multiple optima, which is not a desirable property. This prompts a better definition of the CoS based on the empirical copula as explained next.

##### A. The Empirical Copula

Let  $\{(x_i, y_i), i=1, \dots, n, n \geq 2\}$  be a 2-dimensional data set of size  $n$  drawn from a continuous bivariate joint distribution function,  $H(x, y)$ . Let  $R_{xi}$  and  $R_{yi}$  be the rank of  $x_i$  and of  $y_i$ , respectively. Deheuvels [22] defines the associated empirical copula as

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(u_i = \frac{R_{xi}}{n} \leq u, v_i = \frac{R_{yi}}{n} \leq v). \quad (3)$$

The empirical relative distance,  $\lambda(C_n(u,v))$ , satisfies Definition 4 by replacing  $C(u,v)$  with the empirical copula given by (3).

##### B. Defining the CoS Statistic for Bivariate Dependence

Let  $X$  and  $Y$  be two continuous random variables with a copula  $C(u,v)$ . Consider the ordered sequence,  $x_{(1)} \leq \dots \leq x_{(n)}$ , of  $n$  realizations of  $X$ . This sequence yields  $u_{(1)} \leq \dots \leq u_{(n)}$  since  $u_i = R_{xi}/n$  as given by (3). Let  $\mathcal{D}$  be the set of  $m$  contiguous domains  $\{\mathcal{D}_i, i = 1, \dots, m\}$ , where each  $\mathcal{D}_i$  is a  $u$ -interval associated with a non-decreasing or non-increasing sequence of  $C_n(u_i, v_j)$ ,  $i = 1, \dots, n$ . Let  $C_i^{min}$  and  $C_i^{max}$  respectively denote the smallest and the largest value of  $C_n(u,v)$  on the domain  $\mathcal{D}_i$ . Let  $\gamma_i$  be defined as:

$$\gamma_i = \begin{cases} 1 & \text{at a local optimum of } Y = f(X) \text{ on } \mathcal{D}_i, \\ \frac{\lambda(C_i^{min}) + \lambda(C_i^{max})}{2}, & \text{otherwise.} \end{cases} \quad (4)$$

**Definition 5:** Let  $n_i$  denote the number of data points in the  $i$ -th domain  $\mathcal{D}_i$ ,  $i = 1, \dots, m$ , while letting a boundary point belong to two contiguous domains,  $\mathcal{D}_i$  and  $\mathcal{D}_{i+1}$ . Then, the copula statistic is defined as:

$$CoS(X, Y) = \frac{1}{n+m-1} \sum_{i=1}^m n_i \gamma_i \quad (5)$$

**Corollary 3:** The CoS of two random variables,  $X$  and  $Y$ , has the following asymptotic properties:

- $0 \leq CoS(X, Y) \leq 1$ ;
- $CoS(X, Y) = 0$  if and only if  $X$  and  $Y$  are independent;
- If  $Y = f(X)$  almost surely, then  $CoS(X, Y) = 1$ .

##### C. Algorithmic implementation of the Copula Statistic

Given a two-dimensional data sample of size  $n$ ,  $\{(x_j, y_j), j=1, \dots, n, n \geq 2\}$ , the algorithm that calculates the CoS consists of the following steps:

1) Calculate  $u_j, v_j$  and  $C_n(u,v)$  as follows:

- $u_j = \frac{1}{n} \sum_{k=1}^n \mathbf{1}\{k \neq j: x_k \leq x_j\}$ ;

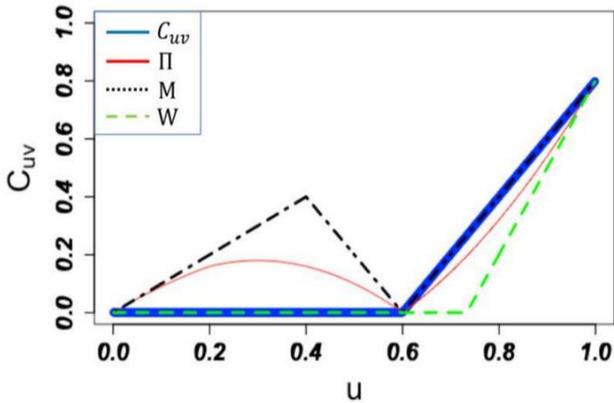


Fig. 2. Graph (blue circles) of the projections on the  $(u, C(u,v))$  plane of  $C(u,v)$  associated with  $X \sim U(-5,5)$  and  $Y = f(X) = (X-1)^2$ . The  $u$  coordinates of the data points are equally spaced. The minimum of the function  $f(\cdot)$  is associated with  $u = 0.6$  and  $C(u,v) = 0$ . Similar graphs are shown for  $M(u,v)$  (dotted black),  $W(u,v)$  (dashed green), and  $\Pi(u,v)$  (solid red). Here,  $C(u,v) = W(u,v)$  for  $0 \leq u \leq 0.6$ , which corresponds to  $f'(x) \leq 0$ , and  $C(u,v) = M(u,v)$  for  $0.6 \leq u \leq 1$ , which corresponds to  $f'(x) \geq 0$ .

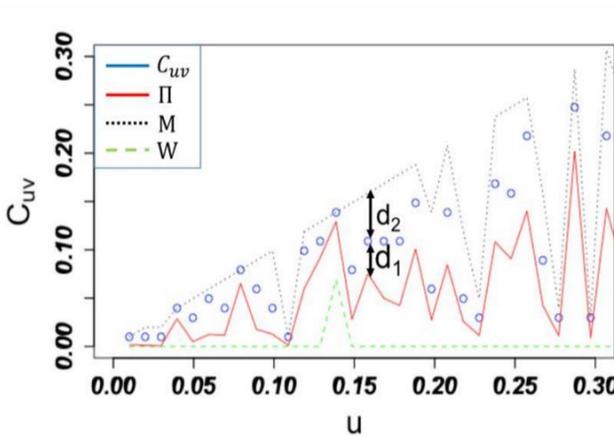


Fig. 3. Graph (blue circles) of the projections on the  $(u, C(u,v))$  plane drawn from the Gaussian copula  $C(u,v)$  with  $\rho_p = 0.5$ . Similar graphs are shown for  $M(u,v)$  (dotted black),  $W(u,v)$  (dashed green), and  $\pi(u,v)$  (solid red). The empirical relative distance function is given by  $\lambda(C(u,v)) = d_1/d_2$ , where  $d_1$  is the distance from  $C(u,v)$  to  $\Pi(u,v)$  and  $d_2$  is the distance from  $M(u,v)$  to  $\Pi(u,v)$ .

Now, the question that arises is the following: Is  $\lambda(C(u,v)) = 1$  for all  $(u,v) \in \mathbf{I}^2$  when there is a functional dependence with multiple optima, be they global or local? The answer is given by the following two theorems:

**Theorem 5:** If  $Y = f(X)$  almost surely where  $f(\cdot)$  has at least two global maxima or two global minima and no local optima on the domain  $\mathcal{D} = \text{Range}(X) \times \text{Range}(Y)$ , then there exists a non-empty interval of  $X$  for which  $\lambda(C(u,v)) < 1$ .

The proof of Theorem 5 is provided in the appendix.

**Theorem 6:** If  $Y = f(X)$  almost surely, where  $f(\cdot)$  has a local optimum, then  $\lambda(C(u,v)) \leq 1$  at that point.

The proof of Theorem 6 is provided in the appendix.

- b.  $v_j = \frac{1}{n} \sum_{j=1}^n \mathbf{1}\{k \neq j: y_k \leq y_j\}$ ;
- c.  $C_n(u, v) = \frac{1}{n} \sum_{j=1}^n \mathbf{1}\{u_j \leq u, v_j \leq v\}$ ;

2) Order the  $x_j$ 's to get  $x_{(1)} \leq \dots \leq x_{(n)}$ , which results in  $u_{(1)} \leq \dots \leq u_{(n)}$  since  $u_j = R_{x_j}/n$ , where  $R_{x_j}$  is the rank of  $x_j$ ;

3) Determine the domains  $\mathfrak{D}_i, i = 1, \dots, m$ , where each  $\mathfrak{D}_i$  is a  $u$ -interval associated with a non-decreasing or non-increasing sequence of  $C_n(u_{(j)}, v_p), j = 1, \dots, n$ .

4) Determine the smallest and the largest value of  $C_n(u, v)$ , denoted by  $C_i^{min}$  and  $C_i^{max}$ , and find the associated  $u_i^{min}$  and  $u_i^{max}$  for each domain  $\mathfrak{D}_i, i = 1, \dots, m$ .

5) Calculate  $\lambda(C_i^{min})$  and  $\lambda(C_i^{max})$ ;

6) If  $\lambda(C_i^{min})$  and  $\lambda(C_i^{max})$  are equal to one, go to Step 8;

7) Calculate the absolute difference between the three consecutive values of  $C_n(u_{(i)}, v_j)$  centered at  $u_i^{min}$  (respectively at  $u_i^{max}$ ) and decide that the central point is a local optimum if (i) both absolute differences are smaller than or equal to  $1/n$ ; and (ii) there are more than four points within the two adjacent domains,  $\mathfrak{D}_i$  and  $\mathfrak{D}_{i+1}$ ;

8) Calculate  $\gamma_i$  given by (16);

9) Repeat Steps 2 through 7 for all the  $m$  domains,  $\mathfrak{D}_i, i = 1, \dots, m$ ;

10) Calculate the  $CoS$  given by (17).

### V. STATISTICAL PROPERTIES OF THE $CoS$

The finite-sample bias of the  $CoS$  is analyzed for the independence case, then a statistical test of bivariate independence is developed.

#### 1) Finite-Sample Bias of the $CoS$

Table 1 displays the sample means and the sample standard deviations of the  $CoS$  for independent random samples generated from three monotonic copulas. As observed, the  $CoS$  has a bias for small to medium sample sizes. Fig. 4(a) shows a bias curve given by  $CoS = 8.05 n^{-0.74}$ , fitted to 19 mean bias values for Gauss(0) using the least-squares method. It is observed that the  $CoS$  bias becomes negligible for a sample size larger than 500. Fig. 4(b) shows values taken by the sample standard deviation  $\sigma_n$  of  $CoS$  for increasing sample size,  $n$ , and for Gauss(0). A fitted curve obtained is also displayed; it is expressed as  $\sigma_n = 2.99 n^{-0.81}$ .

#### 2) Independence Test

One common practical problem is to test the independence of random variables. To this end, hypothesis testing can be applied to the  $CoS$  based on Corollary 3b). The goal is to test the null hypothesis,  $\mathcal{H}_0$ : the random variables are independent, against its alternative,  $\mathcal{H}_1$ . The  $CoS$  is standardized under  $\mathcal{H}_0$  to get

$$Z_n = \frac{CoS - \mu_{n0}}{\sigma_{n0}} \quad (6)$$

Where,  $\mu_{n0}$  and  $\sigma_{n0}$  are the sample mean and the sample standard deviation of the  $CoS$ , respectively. Note that as observed in Fig. 4(a) and (b), for a number of samples  $n$  larger than 500,  $\mu_{n0}$  becomes negligible and  $\sigma_{n0}$  is approximately equal to 0.01.

Hypothesis testing consists of choosing a threshold  $c$  at a significance level  $\alpha$  under  $\mathcal{H}_0$  and then applying the following

decision rule: if  $|z_n| \leq c$ , accept  $\mathcal{H}_0$ ; otherwise, accept  $\mathcal{H}_1$ . Table 2 displays Type-II errors of the statistical test applied to the  $CoS$  for Gauss(0) for sample sizes ranging from 100 to 3000. It is observed that Type II-errors decrease as increases for a given  $n$  and sharply decrease with increasing  $n$ .

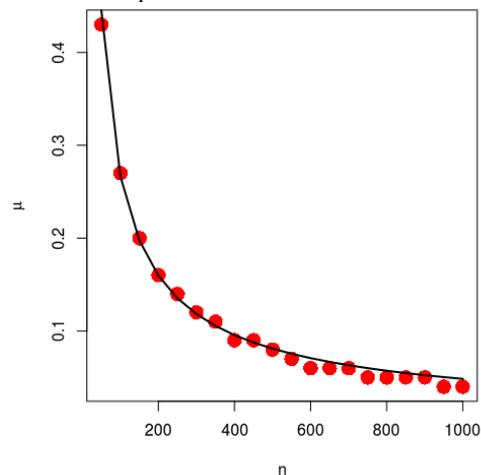
TABLE I. SAMPLE MEANS AND SAMPLE STANDARD DEVIATIONS OF THE  $CoS$  FOR THE GAUSSIAN, GUMBEL, AND CLAYTON COPULA IN THE INDEPENDENCE CASE

$n$	Gauss(0) $\rho_P = 0$		Gumbel(1) $\rho_S = 0$		Clayton(0) $\rho_S = 0$	
	$\mu_n$	$\sigma_n$	$\mu_n$	$\sigma_n$	$\mu_n$	$\sigma_n$
100	0.28	0.08	0.28	0.08	0.28	0.08
500	0.08	0.02	0.08	0.03	0.08	0.02
1000	0.04	0.01	0.04	0.01	0.05	0.01
2000	0.02	0.01	0.02	0.01	0.02	0.01
3000	0.02	0.01	0.02	0.01	0.02	0.01

For monotonic dependence, simulation results show that  $CoS = 1$  for all  $n \geq 2$ . For non-monotonic dependence, there is a bias that becomes negligible when the sample size is sufficiently large. As an illustrative example, Table 3 displays the sample mean,  $\mu_n$ , and the sample standard deviation,  $\sigma_n$ , of the  $CoS$  for increasing sample size,  $n$ , for the sinusoidal dependence,  $Y = \sin(aX)$ . It is observed that as the frequency of the sine function increases, the sample bias,  $1 - \mu_n$ , increases for constant  $n$ .

Table 4 displays  $\mu_n$  and  $\sigma_n$  of the  $CoS$  calculated for increasing  $n$  and for different degrees of dependencies of two dependent random variables following the Gaussian copula. It is interesting to note that for  $n \geq 1000$ , the  $CoS$  is nearly equal to the Pearson's  $\rho_P$  for the Gaussian copula and to the Spearman's  $\rho_S$  for other copulas.

#### 3) Bivariate Dependence



(a)

VI. COMPARATIVE STUDY

In this section, bivariate synthetic datasets and multivariate datasets of breast tumor cells are analyzed.

A. Synthetic Datasets

In this section, the performances of the *CoS*, *dCor*, *RDC*, *Ccor*, and of the *MICe* for various types of statistical dependencies is compared. Székely *et al.* [11] define the distance correlation, *dCor*, between two random vectors, *X* and *Y*, with finite first moments as

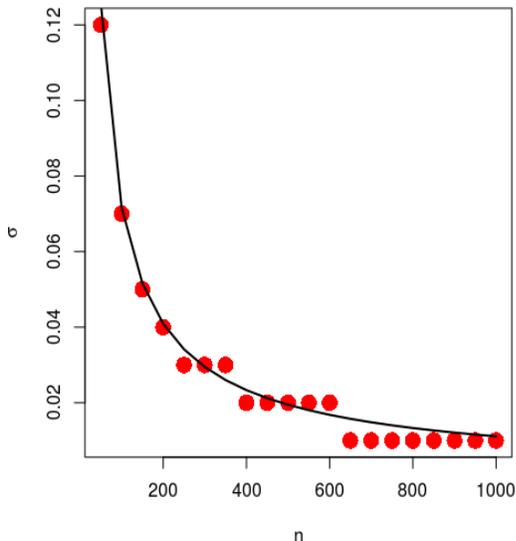
$$dCor(X, Y) = \begin{cases} \frac{v^2(X, Y)}{\sqrt{v^2(X)v^2(Y)}} & \text{for } v^2(X)v^2(Y) > 0, \\ 0 & \text{for } v^2(X)v^2(Y) = 0, \end{cases} \quad (7)$$

where  $v^2(X, Y)$  is the distance covariance. Lopes-Paz *et al.* [9] define the *RDC* as the largest canonical correlation between *k* randomly chosen nonlinear projections of the copula transformed data. Ding *et al.* [9], [10] define the copula correlation (*Ccor*) as half of the  $L_1$  distance between the copula density and the independence copula density. As for the *MIC*, it is defined by Reshef *et al.* [6] as the maximum taken over all *x*-by-*y* grids *G* up to a given grid resolution, typically  $x \cdot y < n^{0.6}$ , of the empirical standardized mutual information,  $I_G(x, y) / \log(\min\{x, y\})$ , based on the empirical probability distribution over the boxes of a grid *G*. Formally,

$$MIC(X, Y) = \max \left\{ \frac{I_G(x, y)}{\log(\min\{x, y\})} \right\}. \quad (8)$$

1) Bias Analysis for Non-Functional Dependence

A bias analysis is performed for the *MICe*, the *Ccor*, the *CoS*, the *RDC*, and the *dCor* and using three data samples drawn from a bivariate Gaussian copula with  $\rho_p(X, Y) = 0.2, 0.5$  and  $0.8$ , which models a weak, medium and strong dependence, respectively.



(b)

Fig. 4. (a) Bias mean values and (b) standard deviation values (red solid circles) for the *CoS* along with fitted curves (solid lines) using the least-squares method for the independence case.

TABLE II. TYPE-II ERRORS OF THE STATISTICAL TEST OF BIVARIATE INDEPENDENCE BASED ON *CoS* FOR GAUSS(0)

<i>N</i>	$\mu_{n0}$	$\sigma_{n0}$	Type-II error for $\rho_n = 0.1$	Type-II error for $\rho_n = 0.3$
100	0.28	0.08	97%	46%
500	0.08	0.02	27%	0%
1000	0.04	0.01	0%	0%
2000	0.02	0.01	0%	0%
3000	0.02	0.01	0%	0%

TABLE III. SAMPLE MEANS AND SAMPLE STANDARD DEVIATIONS OF THE *CoS* FOR THREE SINUSOIDAL FUNCTIONS OF INCREASING FREQUENCY

<i>N</i>	<i>Sin(x)</i>		<i>Sin(5x)</i>		<i>Sin(14x)</i>	
	$\mu_n$	$\sigma_n$	$\mu_n$	$\sigma_n$	$\mu_n$	$\sigma_n$
100	1.00	0.00	0.91	0.10	0.67	0.10
500	1.00	0.00	0.99	0.03	0.88	0.07
1000	1.00	0.00	1.00	0.01	0.96	0.04
3000	1.00	0.00	1.00	0.00	1.00	0.01
5000	1.00	0.00	1.00	0.00	1.00	0.00

TABLE IV. SAMPLE MEANS AND SAMPLE STANDARD DEVIATIONS OF THE *CoS* FOR THE NORMAL COPULA

<i>N</i>	Gauss(0.1) $\rho_p = 0.1$		Gauss(0.3) $\rho_p = 0.3$	
	$\mu_n$	$\sigma_n$	$\mu_n$	$\sigma_n$
100	0.33	0.09	0.49	0.09
500	0.14	0.05	0.36	0.05
1000	0.11	0.03	0.33	0.04
2000	0.09	0.02	0.32	0.03

The sample sizes range from 50 to 2000, in steps of 50. From Fig. 5, it is observed that unlike the *MICe* and *Ccor*, the *CoS*, *RDC*, and the *dCor* are almost equal to for large sample size.

2) Functional Dependence

Another series of simulations are conducted to compare the performance of the *MICe*, the *Ccor*, the *CoS*, the *RDC*, and the *dCor* when they are applied to four data sets drawn from an affine, polynomial, periodic, and circular bivariate relationship with an increasing level of white Gaussian noise. Described in [23], the procedure is executed with  $N = n = 1000$ , where  $n$  is the number of realizations of a uniform random variable  $X$  and  $N$  is the number of times the procedure is executed.

It is inferred from Table 5 that while the *CoS*, *dCor*, *Ccor* steadily decrease as the noise level  $p$  increases, the *MICe* sharply decreases as  $p$  grows from 0.5 to 2 and then reaches a plateau for  $p > 2$ . The *RDC* also decreases steadily with an increase in noise level for the functional dependencies considered, except for the quadratic dependence where it maintains a high power even under heavy noise level.

3) Ripley's Forms and Copula's Induced Dependence

Table 6 reports values of the *MICe*, the *Ccor*, the *CoS*, the *RDC*, and the *dCor* for Ripley's forms, and copula-induced dependencies for a sample size  $n = 1000$  averaged over 1000 Monte-Carlo simulations. The values of the Spearman's  $\rho_s$  for Gumbel(5), Clayton(-0.88), Galambos(2), and BB6(2, 2) copulas are calculated using the copula and the CDVine toolboxes of the software package R. As for the four Ripley's forms displayed in Fig. 6, a linear congruential generator using the Box-Muller transformation is used to generate several bivariate sequences with nonlinear dependencies.

Table 6 shows that the *CoS*, *MICe*, *RDC*, and *Ccor* correctly reveal some degree of nonlinear dependence for Ripley's form 2, with the *Ccor* detecting the highest level of dependence and the *dCor* the lowest level. It is observed that the *Ccor* is the only metric to correctly reveal some degree of nonlinear dependence for Ripley's form 3.

Furthermore, unlike the *MICe* values, the *dCor* and the *CoS* values are very close to the Pearson's  $\rho_p$  value for the Gaussian copula and to the Spearman's  $\rho_s$  values for the Gumbel, Clayton, Galambos and BB6 copulas.

B. Statistical Power Analysis

Finally, following Simon and Tibshirani [23], the power of the statistical tests based on the *CoS*, *dCor*, *RDC*, *TICe*, and the *Ccor* for bivariate independence subject to increasing additive Gaussian noise levels is tested. Six noisy functional dependencies at a noise level  $p$  ranging from 10% to 300% are considered. They include a linear, a quadratic, a cubic, a fourth-root, a sinusoidal, and a circular dependence. The results are shown in Fig. 7.

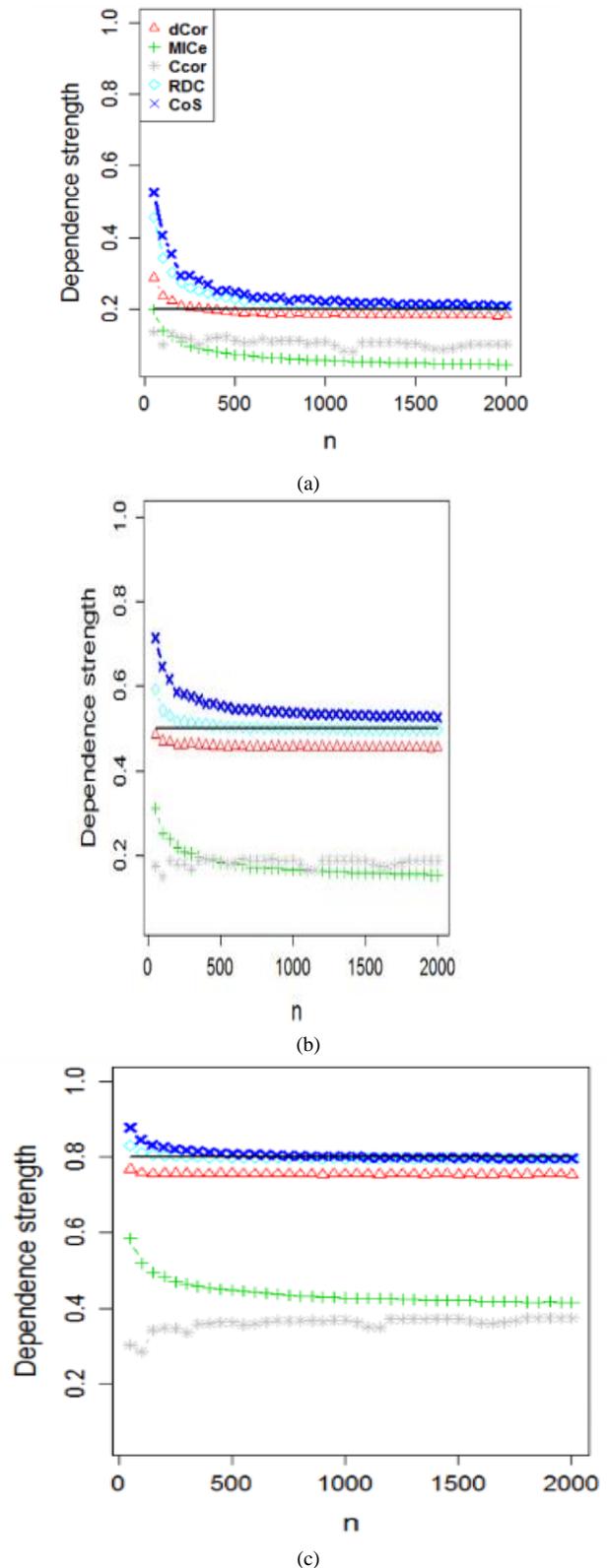


Fig. 5. Bias curves of the *CoS*, *MICe*, *dCor*, *RDC*, and *Ccor* for the bivariate Gaussian copula with  $\rho_P(X,Y) = 0.2, 0.5$  and  $0.8$ , which are displayed in a), b), and c), respectively, and for sample sizes that vary from 50 and 2000 with steps of 50.

TABLE V. SAMPLE MEANS OF THE *CoS*, *dCor* AND THE *MICe* FOR SEVERAL DEPENDENCE TYPES AND ADDITIVE NOISE LEVELS

Noise level $p$	Type of dependence	0.5	1	2	3
Affine: $Y = 2X+1$	<i>CoS</i>	0.86	0.72	0.41	0.29
	<i>dCor</i>	0.91	0.71	0.46	0.35
	<i>MICe</i>	0.88	0.46	0.26	0.22
	<i>RDC</i>	0.95	0.74	0.60	0.59
	<i>Ccor</i>	0.63	0.47	0.34	0.30
4 <sup>th</sup> -order Polynomial: $Y=(X^2-0.25)(X^2-1)$	<i>CoS</i>	0.64	0.41	0.29	0.26
	<i>dCor</i>	0.41	0.35	0.31	0.30
	<i>MICe</i>	0.79	0.54	0.49	0.48
	<i>RDC</i>	0.95	0.93	0.92	0.91
	<i>Ccor</i>	0.72	0.63	0.60	0.59
Periodic: $Y = \cos(X)$	<i>CoS</i>	0.53	0.46	0.28	0.23
	<i>dCor</i>	0.35	0.27	0.17	0.13
	<i>MICe</i>	0.78	0.40	0.22	0.19
	<i>RDC</i>	0.85	0.67	0.43	0.36
	<i>Ccor</i>	0.57	0.41	0.29	0.26

C. Feature Selection Applied to Breast Cancer Data

In order to reduce computation time, improve prediction performance and reducing irrelevant data in machine learning applications, the feature selection presents the all-important step required to choose the optimal subset of data.

The dependent variables provide useless information about the classes and thus serve as noise for the predictor. The rule of thumb here is that best feature selection must include independents features that have a strong dependence with the class or the label considered. The dimensionality reduction is a part of most known methods in machine learning such as filter, wrapper and embedded methods. Pearson correlation coefficient and mutual information are largely used in feature selection; nevertheless, the results are still unsuitable. A serious alternative here is using the *CoS* index to work out the feature selection problem.

A useful data for this purpose is the Wisconsin Diagnostic Breast Cancer (WDBC) data, available on UCI machine learning repository. The extraction of breast tumor tissue is performed using a fine needle aspiration (FNA). The procedure begins by obtaining a small drop of the fluid in hand by examining the characteristics of individual cells and important contextual features such as the radius of the nucleus, the compactness, the smoothness, among others.

A dataset of 569 cells (malignant and benign) and 30 input features is obtained [24]. Among the 30 features, 20 considered are computed from the others; hence, only 10 features are considered as initial subset. Table 6 reports the *CoS* measures for all pairwise feature dependence.

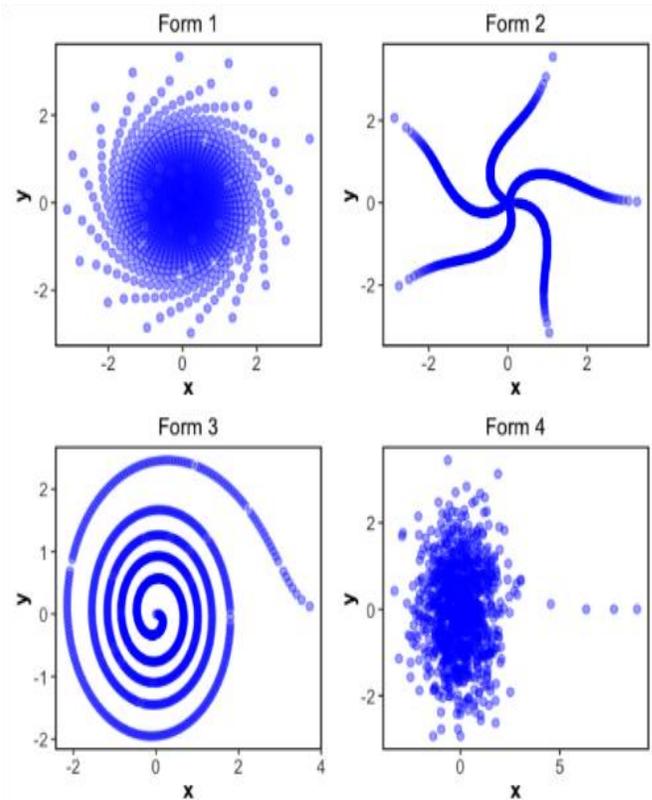


Fig. 6. Plots of four Ripley's forms generated using a linear congruential generator followed by the Box-Muller transformation. The parameters of the congruential generator,  $x_{i+1} = (a \cdot x_i + c)$  modulo  $M$ , are as follows: Form 1:  $a = 65, c = 1, M = 2048$ ; Form 2:  $a = 1229, c = 1, M = 2048$ ; Form 3:  $a = 5, c = 1, M = 2048$ ; Form 4:  $a = 129, c = 1, M = 2^{64}$ .

TABLE VI. DEPENDENCE INDICES FOR COPULA DEPENDENCIES AND RIPLEY'S FORMS

Type of Dependence	<i>CoS</i>	<i>dCor</i>	<i>MICe</i>	<i>RDC</i>	<i>Ccor</i>
Ripley's form 1	0.01	0.02	0.02	0.02	0.01
Ripley's form 2	0.52	0.19	0.42	0.42	0.84
Ripley's form 3	0.14	0.08	0.12	0.13	0.26
Ripley's form 4	0.03	0.04	0.03	0.08	0.09
Gaussian(0.1)	0.11	0.10	0.04	0.13	0.10
Gumbel(5)	0.92	0.93	0.72	0.96	0.62
Clayton(-0.88)	0.90	0.87	0.68	0.88	0.75
Galambos(2)	0.82	0.79	0.48	0.86	0.42
BB6(2,2)	0.84	0.83	0.57	0.92	0.48

Using 0.90 as a threshold to decide a total dependence, the subset is reduced to only 7 features. If the choice is spanned to a threshold of 0.85, the subset length is further reduced to five features [25]. Fig. 8 displays the scatters of the final subset empirical copulas while Fig. 9 displays the heat maps.

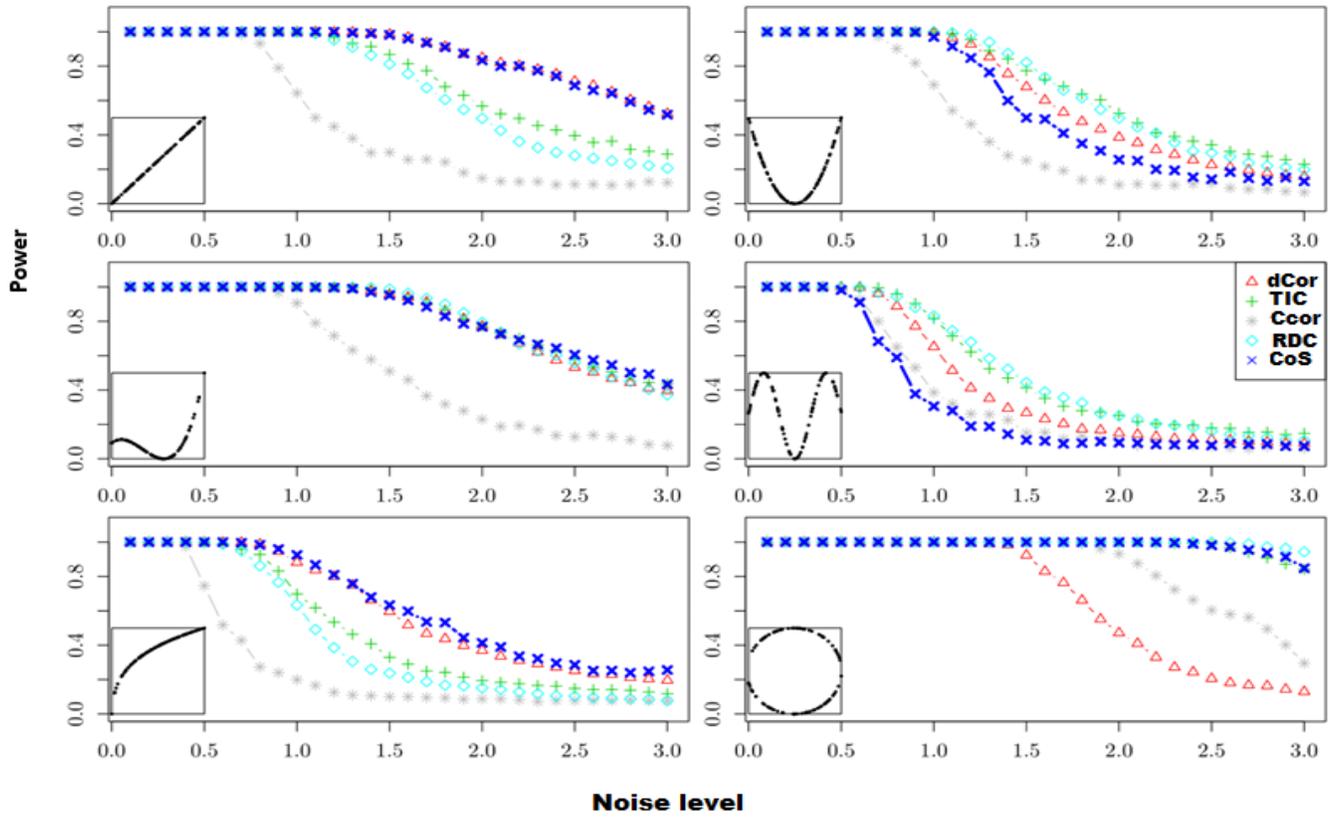


Fig. 7. Displays the power of the tests calculated using a collection of  $N = 500$  data samples, each of size  $n = 500$ , for a significance level  $\alpha = 5\%$  under the null hypothesis. As observed from that figure, the CoS is a powerful measure of dependence for the linear, cubic, circular and rational dependence.

TABLE VII. CoS VALUES FOR BIVARIATE DEPENDENCE BETWEEN FEATURES

	Radius	Texture	Perimeter	Area	Smooth	Compact	Concav	Nbrconcav	Sym	Fractal
Radius	1	0.23	<b>0.99</b>	<b>0.99</b>	0.37	0.33	0.60	0.72	0.18	0.41
Texture	-	1	0.24	0.25	0.18	0.27	0.26	0.21	0.17	0.18
Perimeter	-	-	1	<b>0.99</b>	0.40	0.40	0.60	0.78	0.17	0.33
Area	-	-	-	1	0.37	0.30	0.61	0.75	0.15	0.43
Smooth	-	-	-	-	1	0.80	0.75	0.74	0.72	<b>0.85</b>
Compact	-	-	-	-	-	1	<b>0.88</b>	0.84	0.76	0.84
Concav	-	-	-	-	-	-	1	<b>0.93</b>	0.67	0.64
Nbrconcav	-	-	-	-	-	-	-	1	0.56	0.48
Sym	-	-	-	-	-	-	-	-	1	0.71
Fractal	-	-	-	-	-	-	-	-	-	1

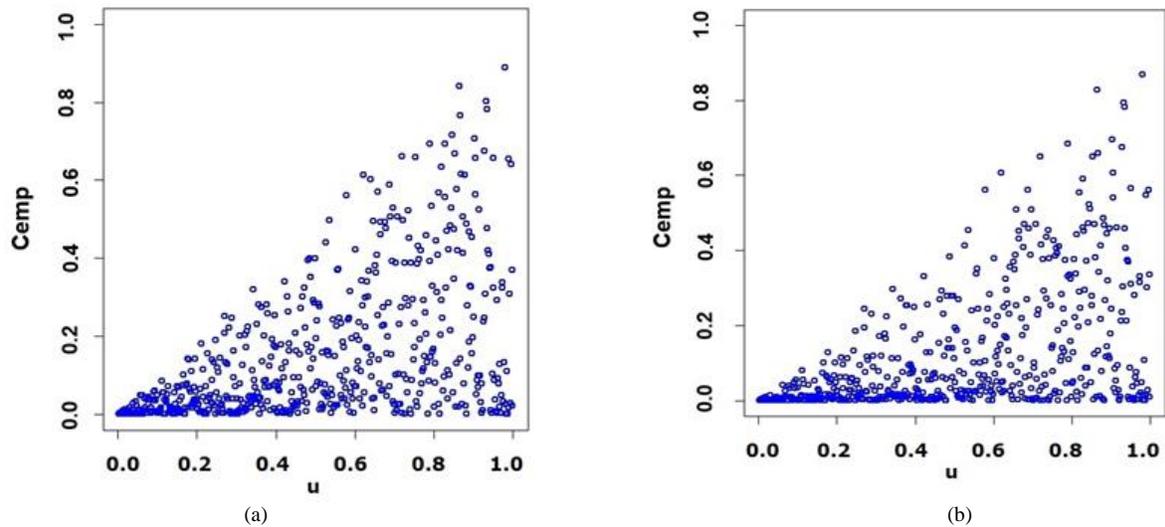


Fig. 8. Scatters of empirical copulas for a) the final feature subset except the perimeter and for b) the final subset, where the CoS values are respectively 0.27 and 0.26.

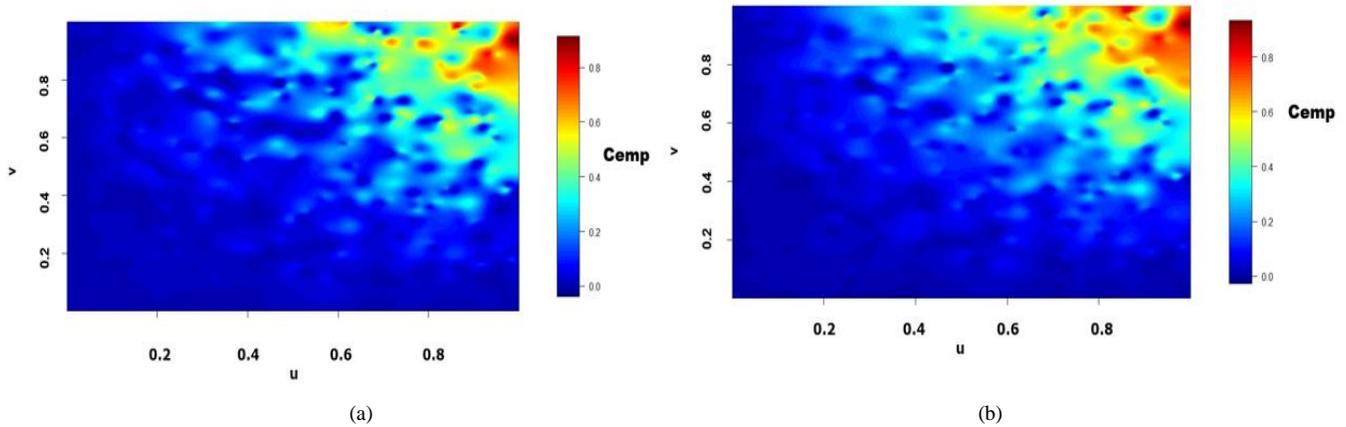


Fig. 9. Heat maps of empirical copulas for a) the final feature subset except the perimeter and for b) the final subset, CoS values are respectively 0.27 and 0.26.

## VII. CONCLUSIONS AND FUTURE WORK

A new reliable statistic for multivariate nonlinear dependence has been proposed and its statistical properties unveiled. In particular, it asymptotically approaches zero for statistical independence and one for functional dependence. Finite-sample bias and standard deviation curves of the *CoS* have been estimated and hypothesis testing rules have been developed to test bivariate independence. The power of the *CoS*-based test has been evaluated for noisy functional dependencies. Monte Carlo simulations show that the *CoS* performs reasonably well for both functional and non-functional dependence and exhibits a good power for testing independence against all alternatives. Good performance of the *CoS* was proved also with other application. Note that the code that implements the *CoS* is available on the GitHub repository.<sup>1</sup> As a future research work, the self-equitability of the *CoS* and other metrics will be assessed under various noise probability distributions and the robustness of the *CoS* to

outliers will be investigated. Furthermore, the *CoS* will be applied to common signal processing and more machine learning problems, including data mining, cluster analysis, and testing of independence. Another interesting property of the *CoS* that is not shared by the *MICe*, *RDC*, *Ccor*, and the *dCor* is its ability to measure multivariate dependence. This property will be investigated as a future work.

### REFERENCES

- [1] C. Spearman, The proof and measurement of association between two things, *The American Journal of Psychology* 15 (1) (1904) 72–101.
- [2] M.G. Kendall, A new measure of rank correlation, *Biometrika* 30 (1/2) (1938) 81–89.
- [3] T. Kowalczyk, Link between grade measures of dependence and of separability in pairs of conditional distributions, *Statistics and Probability Letters* 46 (2000) 371–379.
- [4] F. Vandenhende, P. Lambert, Improved rank-based dependence measures for categorical data, *Statistics and Probability Letters* 63 (2003) 157–163.
- [5] R.B. Nelsen, M. Úbeda-Flores, How close are pairwise and mutual independence? *Statistics and Probability Letters* 82 (2012) 1823–1828.

<sup>1</sup> <https://github.com/stochasticresearch/copulastatistic>

[6] D. N. Reshef, Y.A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Tumbaugh, E. S. Lander, M. Mitzenmacher, P. C. Sabeti, Detecting novel associations in large data sets, *Science* 334 (2011) 1518-1524.

[7] Y. A. Reshef, D. N. Reshef, P. C. Sabeti, M. M. Mitzenmacher, Equitability, interval estimation, and statistical power, *ArXiv pre-print, ArXiv: 1505.02212* (2015) 1-25.

[8] D. Lopez-Paz, P. Henning, B. Scholkopf, The randomized dependence coefficient, *Advances in neural information processing systems* 26, Curran Associates, Inc. (2013).

[9] A. Ding, Y. Li, Copula correlation: An equitable dependence measure and extension of Pearson's correlation, *arXiv:1312.7214* (2015).

[10] Y. Chang, Y. Li, A. Ding, J.A. Dy, Robust-equitable copula dependence measure for feature selection, *Proceedings of the 19<sup>th</sup> International Conference on Artificial Intelligence and Statistics*, 41 (2016) 84-92.

[11] G.J. Székely, M.L. Rizzo, N.K. Bakirov, Measuring and testing dependence by correlation of distances, *The Annals of Statistics* 35 (6) (2007) 2769-2794.

[12] G. Marti, S. Andler, F. Nielsen, P. Donnat, Optimal transport vs. Fisher-Rao distance between copulas for clustering multivariate time series, *arXiv:1604.08634v1* (2016).

[13] R. B. Nelsen, An introduction to copula, Springer Verlag, 2<sup>nd</sup> ed., New York (2006).

[14] A. Sundaresan, P. K. Varshney, Estimation of a random signal source based on correlated sensor observations, *IEEE Transactions on Signal Processing* (2011) 787-799.

[15] X. Zeng, J. Ren, Z. Wang, S. Marshall, T. Durrani, Copulas for statistical signal processing (Part I): Extensions and generalization, *Signal Processing* 94 (2014) 691-702.

[16] X. Zeng, J. Ren, Z. Wang, S. Marshall, T. Durrani, Copulas for statistical signal processing (Part I): Simulation, optimal selection and practical applications, *Signal Processing* 94 (2014) 681-690.

[17] S. G. Iyengar, P. K. Varshney, T. Damarla, A parametric copula-based framework for hypothesis testing using heterogeneous data, *IEEE Transactions on Signal Processing* 59 (5) (2011) 2308-2319.

[18] X. Zeng, T. S. Durrani, Estimation of mutual information using copula density function, *Electronics Letters* 47 (8) (2011) 493-494.

[19] A. L. Blum, P. Langley, Selection of relevant features and examples in machine learning, *Artif Intell* 1997; 97:245-70.

[20] C. Lazar, J. Taminau, S. Meganck, D. Steenhoff, A. Coletta, C. Molter *et al.*, A survey on filter techniques for feature selection in gene expression microarray analysis, *IEEE/ACM Trans Comput Biol Bioinform* 9 (2012).

[21] E. L. Lehmann, Some concepts of dependence, *The Annals of Mathematical Statistics* 37 (1966) 1137-1153.

[22] P. Deheuvels, La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance, *Bulletin de la Classe des Sciences, Academie Royale de Belgique* 65 (1979) 274-292.

[23] N. Simon, R. Tibshirani, Comment on "Detecting novel associations in large data sets" in [7], *Science* 334 (2011) 1518-1524.

[24] W. N. Street, W.H. Wolberg, and O.L. Mangasarian, Nuclear feature extraction for breast tumor diagnosis, *Proceedings of the IS&T/SPIE International Symposium on Electronic Imaging: Science and Technology* 1905 (1993) 861-870.

[25] G. Chandrashekar, S. Ferat, A survey on feature selection methods, *Computers and Electrical Engineering*, 40 (1) (2014) 16-28.

APPENDIX

In this appendix, Lemma 1 is stated and proved and then the proofs of Theorems 2, 3, 5, and 6 are provided.

**Lemma 1:** Let  $X$  and  $Y$  be two continuous random variables with copula  $C(F_1(x), F_2(y)) = H(x, y) = P(X \leq x, Y \leq y)$ . Then it follows:

a)  $P(X \leq x, Y > y) = F_1(x) - C(F_1(x), F_2(y));$  (9)

b)  $P(X > x, Y \leq y) = F_2(y) - C(F_1(x), F_2(y));$  (10)

c)  $P(X > x, Y > y) = 1 - F_1(x) - F_2(y) + C(F_1(x), F_2(y)).$  (11)

**Proof of Lemma 1:** Partition the domain  $\mathcal{D} = \text{Range}(X) \times \text{Range}(Y)$  of the joint probability distribution function,  $H(x, y)$ , into four subsets, namely

$\mathcal{D}_1 = \{X \leq x, Y \leq y\}$ ,  $\mathcal{D}_2 = \{X \leq x, Y > y\}$ ,  $\mathcal{D}_3 = \{X > x, Y \leq y\}$  and  $\mathcal{D}_4 = \{X > x, Y > y\}$ . Then it follows :  
 $P(X \leq x, Y \leq y) + P(X \leq x, Y > y) + P(X > x, Y \leq y) + P(X > x, Y > y) = 1.$   
 (12)

It also follows

$P(X \leq x, Y \leq y) + P(X \leq x, Y > y) = P(X \leq x),$  (13)

which yields (9), and

$P(X \leq x, Y \leq y) + P(X > x, Y \leq y) = P(Y \leq y),$  (14)

which yields (10). Substituting the expressions of  $P(X \leq x, Y > y)$  given by (9) and of  $P(X > x, Y \leq y)$  given by (10) into (12) produces (11). ■

**Proof of Theorem 2:** a) Under the assumption that  $Y = f(X)$ , suppose that  $(x_i, y_{max})$  is a global maximum of  $f(.)$ . Then, by definition  $C(F_1(x_i), F_2(y_{max})) = P(X \leq x_i, Y \leq y_{max}) = P(X \leq x_i)$ , implying that  $C(u_i, 1) = u_i$ . Additionally,  $\text{Min}(u_i, v_i) = \text{Min}(u_i, 1) = u_i$  and  $\text{Max}(u_i, v_i) = \text{Max}(u_i + 1 - 1, 0) = u_i$ , from which (1) follows. To prove the converse under the assumption that  $Y = f(X)$ , suppose that there exists a pair  $(u_i, v_i)$  such that  $C(u_i, v_i) = M(u_i, v_i) = W(u_i, v_i) = \Pi(u_i, v_i) = u_i$ . It follows that  $v_i = 1$ , which implies that  $C(u_i, 1) = u_i$  and  $C(F_1(x_i), F_2(y_{max})) = P(X \leq x_i, Y \leq y_{max})$ , that is,  $(x_i, y_{max})$  is a global maximum of  $f(.)$ .

b) Suppose that  $Y = f(X)$  and  $(x_2, y_{min})$  is a global minimum. Then, by definition  $C(F_1(x_2), F_2(y_{min})) = P(X \leq x_2, Y \leq y_{min}) = 0$ , implying that  $C(u_2, 0) = 0$ . Additionally,  $W(u_2, v_2) = \min(u_2, 0) = 0$ , and  $M(u_2, v_2) = \max(u_2 + 0 - 1, 0) = 0$ , from which (2) follows. To prove the converse under the assumption that  $Y = f(X)$ , suppose that there exists a pair  $(u_2, v_2)$  such that  $C(u_2, v_2) = M(u_2, v_2) = W(u_2, v_2) = \Pi(u_2, v_2) = u_2 v_2 = 0$ . It follows that either  $u_2 = 0$ , or  $v_2 = 0$ , or  $u_2 = v_2 = 0$ . Consider the first case where  $u_2 = 0$ . It follows that  $C(0, v_2) = 0$ , implying that  $C(F_1(x_{2min}), F_2(y_2)) = P(X \leq x_{2min}, Y \leq y_2) = 0$ . This means that  $(x_{2min}, y_2)$  is a global minimum of  $f(.)$ . Consider the second case where  $v_2 = 0$ . It follows that  $C(u_2, 0) = 0$ , implying that  $C(F_1(x_{2min}), F_2(y_2)) = P(X \leq x_2, Y \leq y_{2min}) = 0$ . This means that  $(x_2, y_{2min})$  is a global minimum of  $f(.)$ . Consider the third case where  $u_2 = v_2 = 0$ . It follows that  $C(0, 0) = 0$ , implying that  $C(F_1(x_{2min}), F_2(y_{2min})) = P(X \leq x_{2min}, Y \leq y_{2min}) = 0$ . This means that  $(x_{2min}, y_{2min})$  is a global minimum of  $f(.)$ . ■

**Proof of Theorem 3:** Suppose that  $Y = f(X)$  almost surely, where  $f(.)$  has a single optimum, which is necessarily a global one. Denote by  $S_1$  and  $S_2$  the non-increasing and the non-decreasing line segments of  $f(.)$ , respectively. Note that  $f(.)$  may have inflection points but may not have a line segment of constant value because otherwise  $Y$  will be a mixed random variable, violating the continuity assumption. Let  $A$  denote a point with coordinate  $(x, y)$  of the function  $f(.)$ . Consider the four subsets  $\mathcal{D}_1 = \{X \leq x, Y \leq y\}$ ,  $\mathcal{D}_2 = \{X \leq x, Y > y\}$ ,  $\mathcal{D}_3 = \{X > x, Y \leq y\}$  and  $\mathcal{D}_4 = \{X > x, Y > y\}$ . Suppose that  $A$  is a point of  $S_1$ . As shown in Fig. 10(a), either  $\mathcal{D}_1 \cap S_1 = \{A\}$  or  $\mathcal{D}_4 \cap S_1 = \emptyset$  depending upon whether  $f(.)$  has a global minimum or a global maximum point, respectively. In the former case,  $P(X \leq x, Y \leq y) = 0$ , implying that  $C(u, v) = 0$ , while in the latter case,  $P(X > x, Y > y) = 0$ , implying from (9) that  $C(u, v) = u + v - 1 \geq 0$ . Combining both cases, it follows that for all  $(x, y) \in S_1$ ,  $C(u, v) = \text{Max}(u + v - 1, 0)$ .

Now, suppose that  $A$  is a point of  $S_2$ . As shown in Fig. 10(b), either  $\mathcal{D}_2 \cap S_2 = \{A\}$  or  $\mathcal{D}_3 \cap S_2 = \emptyset$  depending upon whether  $f(.)$  has a global maximum or a global minimum point, respectively. In the former case,  $P(X \leq x, Y > y) = 0$ , implying from (7) that  $C(u, v) = u$  while in the latter case,  $P(X > x, Y \leq y) = 0$ , implying from (8) that  $C(u, v) = v$ . Combining both cases, it follows from (3) that for all  $(x, y) \in S_2$ ,  $C(u, v) = \min(u, v)$ . ■

**Proof of Theorem 5:** Suppose that  $Y = f(X)$  almost surely, where  $f(.)$  has at least two global maxima and no local optima. As depicted in Fig. 11(a), let  $B$  and  $C$  be two global maximum points of  $f(.)$  with coordinates  $(x_B, y_{max})$  and  $(x_C, y_{max})$ , respectively. This means that there exists  $\Delta x > 0$  such that  $f(x_B \pm \Delta x) < y_{max}$  and  $f(x_C \pm \Delta x) < y_{max}$ . Consider a point  $A$  with coordinate  $(x_A, y_A)$  such that  $x_B < x_A < x_C + \Delta x$ ,  $f(x_B - \Delta x) < y_A < y_{max}$  and  $f(x_C - \Delta x) < y_A < y_{max}$ . Denote by  $S_B$  and  $S_C$  the line segments of  $f(.)$  defined over the intervals  $[f(x_B - \Delta x), y_{max}]$  and  $[f(x_C - \Delta x), y_{max}]$ , respectively, which are shown as solid lines in Fig. 11(a). Partition the domain  $\mathcal{D}$  into four subsets,  $\mathcal{D}_1 = \{X \leq x_A, Y \leq y_A\}$ ,  $\mathcal{D}_2 = \{X \leq x_A, Y > y_A\}$ ,  $\mathcal{D}_3 = \{X > x_A, Y \leq y_A\}$  and  $\mathcal{D}_4 = \{X > x_A, Y > y_A\}$ . As observed in Fig. 12(a),  $\mathcal{D}_1 \cap S_B \setminus \{A\} \neq \emptyset$ ,  $\mathcal{D}_2 \cap S_B \neq \emptyset$ ,  $\mathcal{D}_3 \cap S_C \neq \emptyset$ , and  $\mathcal{D}_4 \cap S_C \neq \emptyset$ , yielding  $\square \square C(u, v) < 1$ . A similar proof can be developed for the case where  $f(.)$  has at least two global minima and no local optima. ■

**Proof of Theorem 6:** Suppose that  $Y = f(X)$  almost surely, where  $f(\cdot)$  has a local minimum point, say point  $A$  of coordinates  $(x_A, y_A)$  as shown in Fig. 11(b). This means that there exists  $\Delta x > 0$  such that  $f(x_A \pm \Delta x) > y_A$ . As depicted in Fig. 11(b), let  $S_{A1}$  and  $S_{A2}$  denote the line segments of  $f(\cdot)$  defined over  $x_A - \Delta x$  and  $x_A + \Delta x$ , respectively. Consider the four domains,  $\mathcal{D}_1 = \{X \leq x_A, Y \leq y_A\}$ ,  $\mathcal{D}_2 = \{X \leq x_A, Y > y_A\}$ ,  $\mathcal{D}_3 = \{X > x_A, Y \leq y_A\}$  and  $\mathcal{D}_4 = \{X > x_A, Y > y_A\}$ . As observed in Fig. 11(b),  $\mathcal{D}_2 \cap S_{A1} \neq \emptyset$  and  $\mathcal{D}_4 \cap S_{A2} \neq \emptyset$ . Now, because  $A$  is by hypothesis a local minimum point, there exist line segments of  $f(\cdot)$  denoted by  $S$  such that  $f(y) < y_A$ . Consequently, one of the following three cases arises: either  $\mathcal{D}_1 \cap S \setminus \{A\} \neq \emptyset$  and  $\mathcal{D}_3 \cap S \neq \emptyset$  as depicted in Fig. 11(b), or  $\mathcal{D}_1 \cap S \setminus \{A\} \neq \emptyset$  and  $\mathcal{D}_3 \cap S = \emptyset$ , or  $\mathcal{D}_1 \cap S \setminus \{A\} = \emptyset$  and  $\mathcal{D}_3 \cap S \neq \emptyset$ . In the first case,  $\square\square C, (u, v) < 1$  while in the last two cases,  $\square\square C, (u, v) = 1$ . A similar proof can be developed for  $f(\cdot)$  with a local maximum point. ■

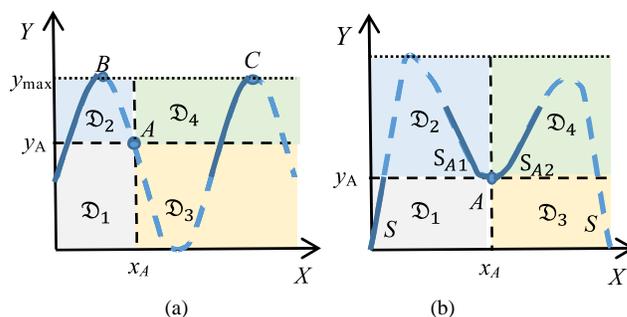


Fig. 11. (a) The graph of a function  $Y = f(X)$  having two global maximum points denoted by  $B$  and  $C$ , and one global minimum point, with two solid line segments denoted by  $S_B$  and  $S_C$ . (b) The graph of a function  $Y = f(X)$  having one local minimum point denoted by  $A$ , with line segments denoted by  $S_{A1}$ ,  $S_{A2}$ , and  $S$ . Four domains,  $\mathcal{D}_1, \dots, \mathcal{D}_4$ , are delineated by the vertical and horizontal lines at position  $X = x_A$  and  $Y = y_A$ , respectively.

AUTHOR PROFILE

**Mohsen Ben Hassine** received an engineering diploma and the M.S. degree in computer sciences from the École Nationale des Sciences de l'Informatique, Tunis, Tunisia, in 1993 and 1996, respectively. He is currently a graduate teaching assistant at the University of El Manar, Tunis, Tunisia. His research interests include statistical signal processing, mathematical simulations, and statistical bioinformatics.

**Lamine Mili** received an electrical engineering diploma from the Swiss Federal Institute of Technology, Lausanne, in 1976, and a Ph.D. degree from the University of Liege, Belgium, in 1987. He is presently a Professor of Electrical and Computer Engineering at Virginia Tech. His research interests include robust statistics, robust statistical signal processing, radar systems, and power system analysis and control. Dr. Mili is a Fellow of IEEE for contribution to robust state estimation for power systems.

**Kiran Karra** received a B.S. in Electrical and Computer Engineering and an M.S. degree in Electrical Engineering from North Carolina State University and Virginia Polytechnic Institute and State University, in 2007 and 2012, respectively. He is currently a research associate at the Virginia Tech and is studying statistical signal processing and machine learning for his PhD research.

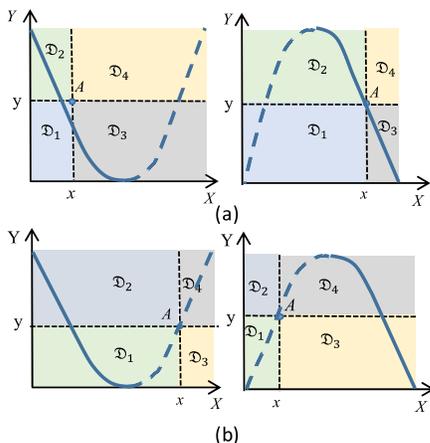


Fig. 10. Graphs of a function  $Y = f(X)$  having a single optimum. A point  $A$  with coordinate  $(x, y)$  is located either on the non-increasing part,  $S_1$ , shown as a solid line in (a) or on the non-decreasing part,  $S_2$ , shown as a dashed line in (b) of the function  $f(\cdot)$ . Four domains,  $\mathcal{D}_1, \dots, \mathcal{D}_4$ , are delineated by the vertical and horizontal lines at position  $X = x$  and  $Y = y$ , respectively.

# Security in OpenFlow Enabled Cloud Environment

Abdalla Alameen

Computer Science Department, College of Arts and  
Science, Prince Sattam bin Abdul Aziz University  
Wadi Adwasir, KSA

Sadia Rubab

Computer Science Department, COMSATS Institute of  
Information Technology,  
Attock Campus, Pakistan

Bhawna Dhupia

Computer Science Department, College of Arts and  
Science, Prince Sattam bin Abdul Aziz University  
Wadi Adwasir, KSA

Manjur Kolhar

Computer Science Department, College of Arts and  
Science, Prince Sattam bin Abdul Aziz University  
Wadi Adwasir, KSA

**Abstract**—Inception of flow tables as data plane abstraction, and forwarding rules that are managed by centralized controllers in emerging Software Defined Networks (SDN) has stemmed significant progress in OpenFlow based architectures. SDN is particularly fueled by data center networking and cloud computing. OpenFlow coupled with cloud solutions provide dynamic networking capabilities. With the benefits obtained from network services, security enforcement become more important and need powerful techniques for its implementation. Extensive researches in cloud security bring forward numerous methods of leveraging the SDN architecture with efficient security enforcement. The future of SDN and mobile networks is also enlightened if security models are satisfactory to cover dynamic and flexible requirements of evolving networks. This paper presents a survey of the state of the art research on security techniques in OpenFlow based cloud environments. Security is one of the main aspect of any network. A fair study and evaluation of these methods are carried out in the paper along with the security considerations in SDN and its enforcement. The security issues and recommendations for 5g network are covered briefly. This work provides an understanding of the problem, its current solution space, and anticipated future research directions.

**Keywords**—Software defined networks; OpenFlow; 5G network component; ONF; virtualization; SDN security framework; future security networks

## I. INTRODUCTION

Cloud computing cannot adequately handle the increased demands of its customers with the traditional network management techniques. Network devices are complex devices, which require individual configuration to change the network behavior. Therefore, to achieve high level of connectivity and communication, cloud computing deployment architectures are propelled by Software Defined Networks (SDN). SDN is introduced as a flexible way to control network in a more sophisticated and planned manner, with OpenFlow as the most commonly used SDN protocols [29]. SDN with centralized control on the network devices, make network open and programmable. Instead of configuring the network devices individually, an administrator can program the behavior of a network centrally. The organizations can develop and install applications for specific network behaviors. These applications

can be for security, traffic engineering, QoS, switching, routing, virtualization, load balancing and many others based on network evolution and new innovations. The core of SDN is to introduce flexibility that evolves with the speed of software. SDN is an enabling technology for 5G [53]. SDN with separated control and forwarding plan has also been applied in wireless networks [54]. The SDN architecture is presented in Fig. 1. The OpenFlow acts as an interface between the control plane and forwarding layer. Section I of the paper gives a brief introduction of SDN and related work. In Section II of the paper Dynamic Cloud Network Management is discussed, which includes various virtualization techniques and its implementation. It also explains the scope of cloud services in a virtualized environment. Section III of the paper elaborates the types of security considerations in SDN and their enforcement in various layers of SDN. Section IV deals with the detail regarding cloud security and OpenFlow. A brief discussion about the security in future network is also added in the last section of the paper. To start with, next section is regarding the OpenFlow based SDN.

### A. OpenFlow based SDN

The OpenFlow is a network control protocol maintained by Open Networking Foundation [1]. SDN concept decouples the control plane from the data plane, whereas OpenFlow define the rules for the communication between the controller and a switch [4]. While managing the network traffic, OpenFlow generates a flow table. The OpenFlow table contains match, priority, counters, instruction, time-outs and cookie fields. All these fields work together to identify and manage network traffic [30]. SDN implemented on OpenFlow offers many advantages for cloud environments. The OpenFlow in SDN can be programmed to streamline the network traffic with high security feature. SDN increase the network visibility of the devices offered, as the control is centralized and it can view both the real and the virtual devices. Full view of the resources facilitates the resource optimization; hence increase the elasticity of the cloud.

OpenFlow defines rules for packet forwarding in OpenFlow switches, describes rules for flow tables and handles the delivery of data packets from one location to another. The statistics about the traffic passing through the OpenFlow switch can be reviewed from flow tables. OFPacketIn,

OFPacketOut and OFFlownod are frequent OpenFlow events to which controller can listen and respond back. The OpenFlow protocol is accepted by many major switches and the routers manufacturing company to support and deliver products compatible with OpenFlow protocol.

### B. Redefined Cloud Network with OpenFlow

Open Networking Foundation (ONF) is the group who is working extensively in the field of dynamic networking architecture via SDN technology. Introduction of OpenFlow based SDN changed the way of networking in a cloud environment. SDN redefined the cloud network to meet the price, performance, scalability and dynamic demand of the clients for cloud services. The solution offered by the SDN is the mix of intelligence and flexibility of routing techniques with a higher capacity of integrated application. This concept supports the programmable control interface which decouples and abstracts the control plane from the data plane [29]. With

the introduction of SDN, control of the resources gets centralized which exponentially decrease the cost of the hardware infrastructure and cost of management of the network. OpenFlow switch permits selective forwarding of data based on inputs from the controller, which decreases the unwanted traffic in the network. The OpenFlow based SDN architecture allows devices to be programmed according to the need of the client [1], [30]. The cloud service providers can use SDN applications to supervise network conditions, provisioning of network resources and network flow traffic to increase the performance, security and quality of service in the cloud. In the OpenFlow network environment, the functionalities to control the devices are programmed in the control box and data packet forwarding behavior is handled by the central controller for all the devices. SDN control plane supports the traditional feature as well as latest functionalities introduced by the OpenFlow based SDN techniques [30].

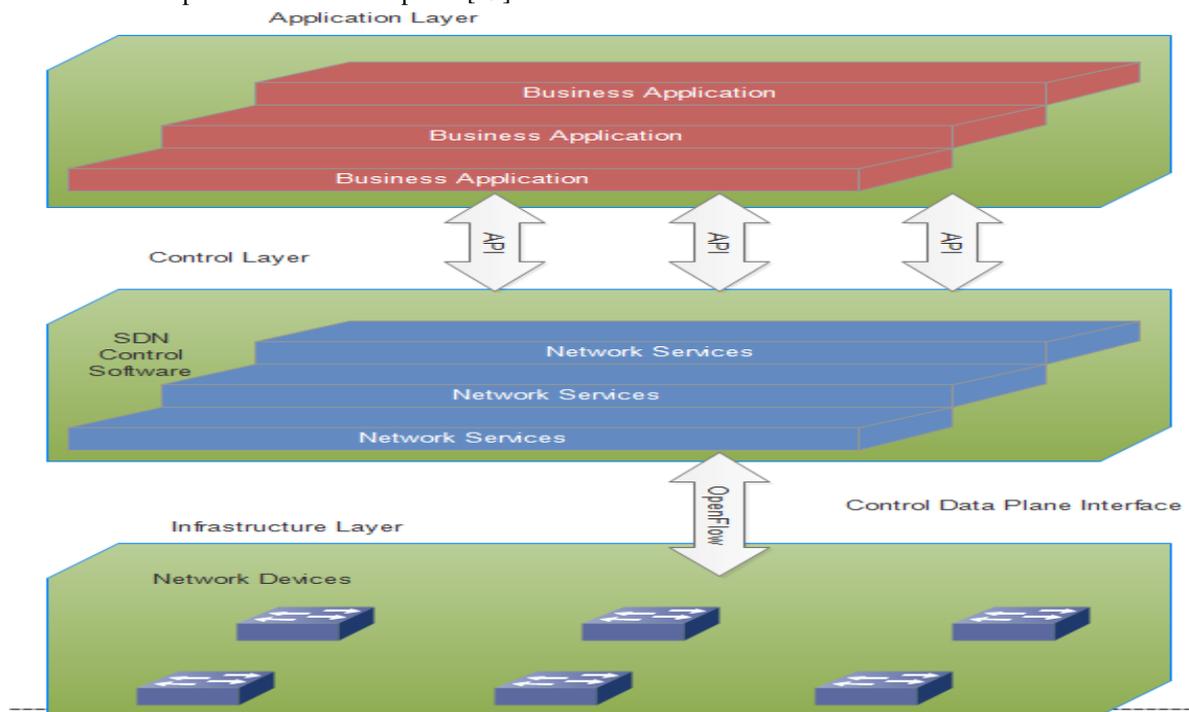


Fig. 1. Openflow based SDN architecture.

## II. DYNAMIC CLOUD NETWORK MANAGEMENT

Traditional cloud providers allow user to customize the facilities like, operating systems, hardware specification, storage and execution environment, however, they do not allow them to customize client's network environment. Virtual Private Cloud is an implementation of dynamic networking services. It permits users to develop network topologies to suit their requirements, with dynamic generation IP addresses, which facilitates dynamic IP addresses [5]. SDN supports virtual network services and simplifies data center infrastructure through dynamic connection between VMs and servers with resources.

Introduction of OpenFlow based SDN in cloud networks improved the security, scalability and portability of the resources in the cloud. It gives client exact location of the

physical resource shared in the cloud. In OpenFlow based networking environment, networking hardware itself act as a firewall. Network switches are programmed with OpenFlow rules to block the unwanted and malicious traffic in the network. To fulfill the increased demands of customers, scalability feature is improved with the help of dynamic networking [5], [6].

The main drivers for dynamic network virtualization are fast network, higher efficiency in utilizing the network resources and cost reduction. In dynamic network first requirement is the virtualization of the data center networks (DCN). Hence, the network virtual layer which is also called hypervisor is developed to virtualize the network resources to make a pool of network devices like, switches and routers

[7]. The hypervisor provides an actual mechanism to implement the functional network virtualization. The hypervisor's job is to fulfill the demands of clients by managing resources effectively. A multiple module virtual environment can have multiple hypervisors that can be used to provision and manage the resources available in the cloud [8].

#### A. Virtualization: Network of Virtual Machine Instances

Virtualization has a great impact on IaaS solutions of cloud service providers. Some examples of IaaS solutions are Nimbus, Open Nebula, Eucalyptus and Open stack [8]. Infrastructure virtualization separates the data plane from the control plane. A virtual infrastructure is defined by virtualizing the various physical resources and control programs on the data plane for provisioning. Each virtual infrastructure has its private set of control application programs within the control server. The users can choose the programmed control modules to prepare their customize environment, which includes, routing, network management, virtual machine migration control and so on [9].

OpenFlow, Open vSwitch, OpenNebula and One Cloud are the SDN based technologies used to implement the virtualization in the cloud. Open vSwitch is implemented through standard management interfaces such as, sFlow, NetFlow and can be customized programmatically. It is used to manage the network traffic between physical hosts and VMs [10]. OpenNebula frameworks are specially designed to manage the virtualized infrastructure which provides private, public and hybrid IaaS. The hypervisor supported by OpenNebula are KVM, VMware and Xen. This provides a centralized management interface for virtual and physical resources. It also supports high extensible plug-in framework management tools like, VM schedulers, virtual image managers to increase the capability of the resources offered [11]. One Cloud allows users to provision virtual machine instances using KVM hypervisor. It is an IaaS system. OpenFlow network and Open vSwitch are used as a virtual bridge to connect all the hypervisors physically.

#### B. Cloud Services Feasible in Virtualized Network

To improve the performance of the cloud we required to implement the virtualization of DCNs. This makes the cloud services feasible for almost all types of clients. The virtualization of DCNs optimizes the cloud service to its maximum extent. As we discussed above also that the maximum impact of virtualization of network is on the IaaS services. DCNs list all the VMs on DCNs servers so as all the available resources can be utilized at the maximum optimization. The DCNs are also responsible for the maintenance, repairs and addition and deletion of resources into the resource pool on DCN servers [12]. DCN provides connection of hundreds of data centers in such an efficient way, so that cloud computing services can expand easily. It helps to implement the reliable and secure communication between the various services.

### III. SECURITY CONSIDERATION IN SDN

Security is one of the major concerns in computer science as companies, government services and individuals rely on computer networks for their day to day activities. The amount

and sensitivity of data stored on network has considerably increased over the time. Viruses, Worms, Distributed Denial of Service (DDoS), Spyware and Trojan Horses the common heard jargons are generally considered as the common types of threats and attacks to network security. DDoS attacks overload the network with too much traffic by sending lots of information on the network. Thus it affects network bandwidth, memory, CPU and so on. If the network goes offline it can cause money and time loss to companies. Most importantly if DDoS takes down servers meant for network security, the whole network is open to the threats and attacks. Attackers try to steal data, they want to ruin the network, or try to use network for unlawful activities. Therefore every connection in network could be an opportunity for attackers. Strong security measures are required that not only protect network users and infrastructure from threats, but also evolve as networks evolve. Effective security not just creates a shield on network, it helps each entity associated with network to thrive and focus with freedom. In the light of given context, this topic elaborated how security has enhanced with SDN platform and cover the security challenges with SDN. SDN that was developed to obtain simplified and secure networking [13]-[18], attain security with dynamic access control, detection and mitigation of attacks and robust traffic monitoring.

#### A. Security Threats to SDN

SDN is an emerging network technology with flexible and agile environment for network traffic control proposed by many companies and researchers such as [19]-[21]. This novel approach comes across significant issues in availability, scalability and security [22]-[23]. Whenever there is new platform, service or infrastructure hackers try to go into that new option. Since SDN is based on programmable control interface, in SDN how the operators deal with the software has major impact on security. If the administrators and software developers not properly deal with security folks, it could have negative impact on security and result in security gap. If there is any compromise in controller and applications, whole network is affected. Fig. 2 represents the points in different layers that could be target of security attacks in SDN. The attacks on threat points cause forge traffic flow between switches, affect communication between controller and switch.

A physical damage, link failure, or attacks on a vulnerable controller [24]-[25] in worst cases results in paralysis of whole network. With the compromised controller, traffic could transfer to compromised nodes, and hacker could insert malware, monitor traffic and could even modify the packet contents. As in SDN route flows among security devices, the consequence of an affected controller could be very catastrophic. In case the communication between data plane and control plane is not appropriately secured entire network security is compromised. Man in the middle attack (with DNS spoofing, ARP Spoofing, session hijacking etc.) between controller and switch, DoS attacks, reply attacks, dodged network access policy by malicious users and unclear management of encrypted packets are threats to data forwarding plane layer [22], [26]. Whereas DDoS is a threat for control layer [22], [26] and illegal access affects application layer [22]. Effective diagnosis of faults and assurance of speedy recovery in SDN remains a security issue if trusted

resources are not incorporated in infrastructure [27]. A thorough analysis of SDN challenges and threats reveals that still lots of research, investigations and efforts required for

transforming this novel techniques into more reliable and secure infrastructure [28].

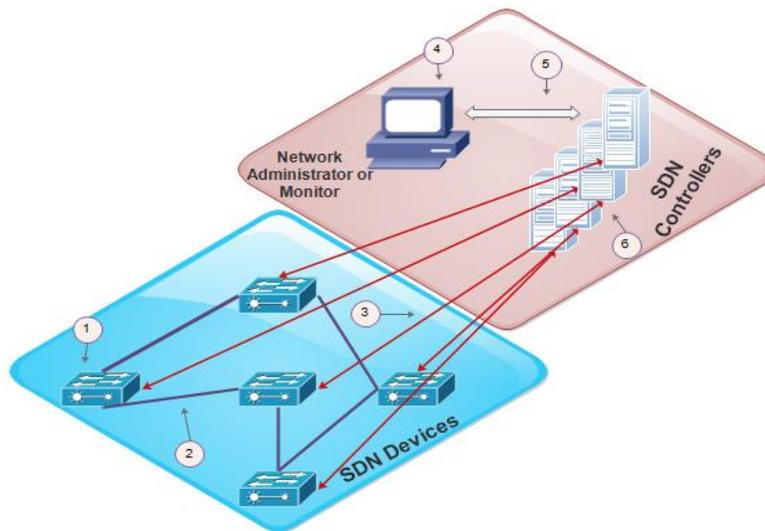


Fig. 2. Threat points in SDN infrastructure.

### B. Enforcing Network wide Security Policies

So far, brief introduction of SDN in context of cloud and the concerns regarding the security around SDN is given. The intense research and development in SDN has resulted in dramatically evolving technology with varied security solutions to threats and attacks. The different ways to define and enforce security policies are summarized in this section. The security policy in network is generally defined as the routing rules, or the actions that are allowed or forbidden in the network. Effective security policy improves quality of service (QoS), and filters the traffic. The security policy for information shared in network could be based on measurable documented standards, and procedure for policy enforcement and disaster recovery. Security policies in SDN spread security in the network with better visibility to network traffic, segregation of services in infrastructure, rigorous monitoring of sensitive data and through inspection of suspicious devices.

Careful design and implementation of SDN controllers and their access control policy is strongly required for network security. Dynamic access control framework implemented in Resonance [29], where controller implements policies with programmable switches. In this research controller provide central authentication and compromised host is segregated. A four state architecture that resembles framework elaborated in [29] was presented with high level policy language in [30]. Access control list and polices for NOX controller as core component of flow based policy enforcement discussed in [31]. The issues between access control and high level policies in switch and firewall are tackled in [32]. Security enforcement kernel FORT-NOX designed for role based authentication, conflict detection and resolution and rule based authorization covered in [33]. An interesting feature of [33] is OpenFlow switches are prevented from packet header modification to bypass firewall. Enforcement of security policy support QoS, network monitoring in real-time environment and problem detection [34]-[35]. SDN utilized in [36] for better

performance and to resolve traditional network security threats. Network services and security improved in this architecture with validation of sources addresses. Flexible, secure and scalable network management architecture for large scale networks [16] enforces security policy and dynamic load balancing in hybrid network with programmable devices. In FLOWGUARD (a centralized firewall) violations in firewall security policy are detected and resolved with policy enforcement on top of the controller [37]. SIMPLE, the work in [38] introduces resource management, flow correlation and policy manager for sending packets through various middle boxes. Load balancing ensured in SIMPLE while installing the rules (translated middleboxes specific policies) in SDN switches.

The scripting language to add security policy to SDN in conjunction with NOX controller incorporated in FRESKO security framework [39]. Scripting language used in development of security applications. Controller program for monitoring and configuration of policies [40], an advance language for network policies for control architecture Procera [41], and network security policy in human readable language for responding to alerts in OpenSec framework [42] represent few of the research work in policy language.

### C. Security Control in Multiple Layers

A review of techniques used to secure network by implementing security policies reveals that security studies are generally layer oriented. Each security solution pays more attention to threats and attacks on some specific layer. In SDN architecture, the three layer architecture is represented as application layer, control layer, and infrastructure layer as shown in Fig. 1. FlowNAC [15] support user authentication and grant rights to the users for network access. Frameworks in [15] and [42], OpenSec designed to deal with the security issues of switches. Switch level security in [42] is improved with spam detection, packet inspection and intrusion detection.

Research work in [16]-[17], [33], [41] provide countermeasures to security issues with control layer. They have more impact on security with automatic user authentication [16], virtual IP allocated to host [17], monitoring with varied granularity [18], avoidance of conflicts in rules [33], and with detection and tracking capabilities [37]. The security of both infrastructure layer and control layer are considered in [13], [29], [36] with DDoS detection, access control and unwanted traffic control respectively. Authors in [39] elaborate security mechanism for control and application layer with threat detection and mitigation.

#### D. How OpenFlow Enhance SDN Security

OpenFlow as one of the most popular specification of SDN significantly improved network security and reliability. Many of the switches in SDN are now adopted with OpenFlow interface. OpenFlow protocol provides secure communication between switches and controller with SSL and TLS encryption. SDN based on OpenFlow delivers better performance in terms of load balancing, routing, firewall configuration and traffic management [43]. In OpenFlow based SDN it's easy to alter packet flow rules, and now many researchers have introduced it in intrusion detection systems, mobile networks, and wireless sensor networks [44]. With the survey of security solutions for SDN, few research efforts that enhance security of OpenFlow based SDN architecture is being identified.

DDoS attacks detection with NOX controller and OpenFlow switches covered in [13]. OpenFlow Random host mutation [16] uses OpenFlow for IP mutation in SDN. In this approach [24] end host is protected from adversaries through unpredictable and random mutation of host IP. Resonance [29] for dynamic access control and distributed network monitoring uses OpenFlow switches and controllers. FORT-NOX [33] directly implemented on NOX as C++ extension is security enforcement kernel. Prioritization of security rules covered in [33] with southbound API. VAVE platform [36] was presented for OpenFlow architecture and engaged OpenFlow protocol for effective validation of source address. VAVE ensures information privacy and prevents spoofed and forged attack on data that pass through OpenFlow interface. FRESCO [39] a development framework for SDN based security applications incorporate security enforcement kernel that is integrated with OpenFlow controller. Research efforts in [42] provide solution to security challenges in switches with OpenFlow enabled framework.

#### E. Threats and Vulnerabilities with SDN utilization in Cloud

As discussed in Sections 1 and 2 that SDN and OpenFlow align well with cloud computing due to the scaling and dynamic nature of cloud. Since SDN is a step towards offering dynamic virtualization services to clouds, and fully virtualized data centers, therefore extensive research and development required to have a clear understanding of its security implications. SDN has some related threats and attacks that are either similar to challenges in traditional networks or specific for SDN. The different aspects of security and vulnerabilities with SDN based cloud briefly covered in this section. A detailed discussion on mechanisms elaborated in researches for development of secure OpenFlow based cloud environment covered in Section 4. The opportunities and vulnerabilities

related to cloud security with SDN are thoroughly discussed in [45].

A centralized and global view of SDN when employed in cloud environment with multiple tenants and shared resources require well defined boundaries for user privileges and limited functionalities passed on to users. SDN provide better control of VLAN and firewall implementation. In dynamic cloud environment quick response to attacks is very important. Though evolution of SDN evolves virtual machine migration, but appropriate security measures are required to avoid attacks on VM traffic. Reliability of cloud could be a major issue if numbers of SDN controllers are limited. As failure of controller has drastic impact on whole network. Problems of inaccuracy and unreliability in network management problem extend to control plane. SDN with central controller and network switches can enforce policy, but it is not as simple as it seems to be. Defining policies on high level such as in cloud require refined frameworks that ensure security with performance. Any negligence in configuration of security policies may cause data leakage, unauthorized access of controller, modification of flow rules, and data modification [55]. How SDN could help to decrease security risks, or worsens the security risk in cloud environment covered in [55].

## IV. CLOUD SECURITY AND OPENFLOW

Though, potentials of SDN for on demand services and applications for the user community cannot be denied. This advancement can induce unpredictable traffic across Cloud network which cannot be eluded by means of traditional approaches. Hence, SDN should equipped with centralized tools to monitor traffic flows so as the security modules of the SDN networks. In this section, the plethora of research work suggested for SDN will be elaborated.

### A. Security Frameworks of SDN

In [46], researchers proved that, adversaries passively and actively can fingerprint SDN networks. They have showed probability of fingerprinting of SDN network by means of RTT and packet pair dispersion. SDN never considered impairment confinement policy with respective damage recovery as mentioned by [47].

Control plane guides traffic rules whenever the data plane requires to control network. However, these kinds of scripting policies can harm SDN seriously when the data plane floods the request flow change to the control plane. Furthermore, Cloud computing techniques hide resources, such as physical servers, used processors, and OS from users. This technique enables network admin to divide a single processor into many independent servers. This flexibility also allows migrating server into other machine in case of failure. However, this migration should consider network topology of the failure machine. Because physical sever is connected to the network which won't be moved, moving virtual LAN from one network to other network creates networking issues. However, SDN is relatively new technology and creates new risks. Specifically, compromised network elements can affect the whole SDN architecture because of its centralized approach [47].

Static address assignment to the network always allows the remote users to scan and send the probe to the remote network.

However, because of limitation of DHCP and NAT protocols, yet IP address change is required at random pace and frequently so as to avoid scanning [27]. Nevertheless, this method is not effective for the DDoS and application layer attacks. However, to distinguish between normal traffic and abnormal traffic of huge size is very difficult in a distributed network environment. These challenges are addressed in [13]. They have implemented self-organizing Maps, which are traffic aware unsupervised artificial neural network. These Maps are used to distinguish normal and abnormal traffic flow.

One more daunting task is security monitoring of large scaled network. In [48], they have focused on how to route network traffic for network equipment's than analyzing the network traffic. Yet another, approach in monitoring traffic is Cloud Watchers[39], in this approach, network traffic bypasses to the network security devices by using programming scripts. In [Resonance], they have delegated traffic management to the network devices. Researchers [Resonance] have used programmable network elements to control network traffic. These lower level network switches are programmed to drop and redirect traffic whenever they sense real-time alerts in the traffic. In [49], middle boxes are used to induce network-wide policy enforcement on the outgoing to packets to provide useful information on host and source states.

Scripting network policies on vendor specific network devices is rigorous. Furthermore, these network policies induce configuration complexity apart from latency to adopt dynamic nature of the network traffic. Hence, network policy scripting is challenging process for the dynamic environment of SDN. In [42], an OpenFlow-based security framework is proposed and it allows a network security administrator to generate and apply security policies scripted in English like languages.

However, scripting network policies, such as enforcing RPKI-Based Routing Policy on the Data Plane at an Internet Exchange [50], will remain main concern for the network operators till SDN adopts new patches of security policies. To mention few, IDS is the choice for the network administrator to identify abnormal rate bounds at the control panel. Another set of solution in line with new security policies is to use autonomic trust management solution for SDN [51]. This policy is built on the adaptive trust model which enables requirement, assessment, instituting, and guaranteeing the trust of network elements according to observation measured at the runtime. Network operators can make use of cryptographic models of multiple certification environments between various network subdomains. In [46], they have proposed solution for the adversary attack (injecting probe) on the SDN network by passively or actively collecting traffic exchanged with the SDN network.

In [52], they have proposed API for packet generation for OpenFlow switches. The API for packet generation allows partition between controller's switch and functions, hence allowing the controller to bypass assigned tasks. However, these methods are not tested well for the security concern. In view of this discussion topics has been summarized few SDN security use cases that are essential to conquer security breaches. SDN vulnerable to attacks but these use cases if

implemented to some extent harden the security attacks.

- Traffic filtering, the major network elements such as SDN switches can act as firewall so the content that is not permitted is denied.
- DDoS Mitigation, communication between DDoS controller system and northbound API configures controller for a clean traffic pass to destination.
- Network Slicing, logical separation in network by addition of slicing layer between control and data plane with strong isolation.
- Network Access Control, unauthorized access prevented through security checks. Nodes that pass checks can only join network and send /receive data.
- Security Traffic Monitoring, packet monitoring tap to assess the data flow improve security.

### B. Security in Future Networks

The rapid progress in networks reveals that in the years 2020 to 2030 almost 100 billion things will be connected with 5G [56]. It means that new approaches are needed while defining the security for 5G to gain the user trust. The features of SDN such as network management and applications management are supportive in getting dynamic nature of 5G. Therefore dynamic and flexible security mechanisms with new trust and delivery models are required. Evolving 5G network architecture is based on SDN for communication between clouds, satellite systems, gateways and other devices. Security, resilience, robustness, privacy, trust and data integrity are main focus of these diverse and functional network environments [57].

The common attacks related to 5G are data manipulation, equipment cloning, rogue devices, unprotected endpoint entry, man-in-the-middle attack, spoofing, and premium content privacy. Most of the attacks can affect all 5G segments and a multilayer security model applicable from network to user guarantee basic security requirements. General 5G network architecture and its layered security vision are depicted in Fig. 3 and 4 respectively.

The ongoing advances and future of cloud computing make it a favorable enabling technique for flexible 5G network. In this exploitation of advanced techniques spectrum, infrastructure, high performance computing will be available as services (anything as a service). In ANYaaS traditional data center services move to mobile connectivity [58], as mobile devices can function as resource providers [59]. OpenFlow operate as standard in SDN enabled mobile networks. But the traditional OpenFlow mechanism is not suitable in mobile networks in terms of security. The impact of attacks on OpenFlow control on SDN based mobile networks are covered in [61]. Proscribed use of frequency, snooping by attackers and privacy issues are the security issues that need to be explored for improved security in 5G networks [53]. The reduced client latency effects users' control over the data and it leads to privacy, authentication and traceability issues.

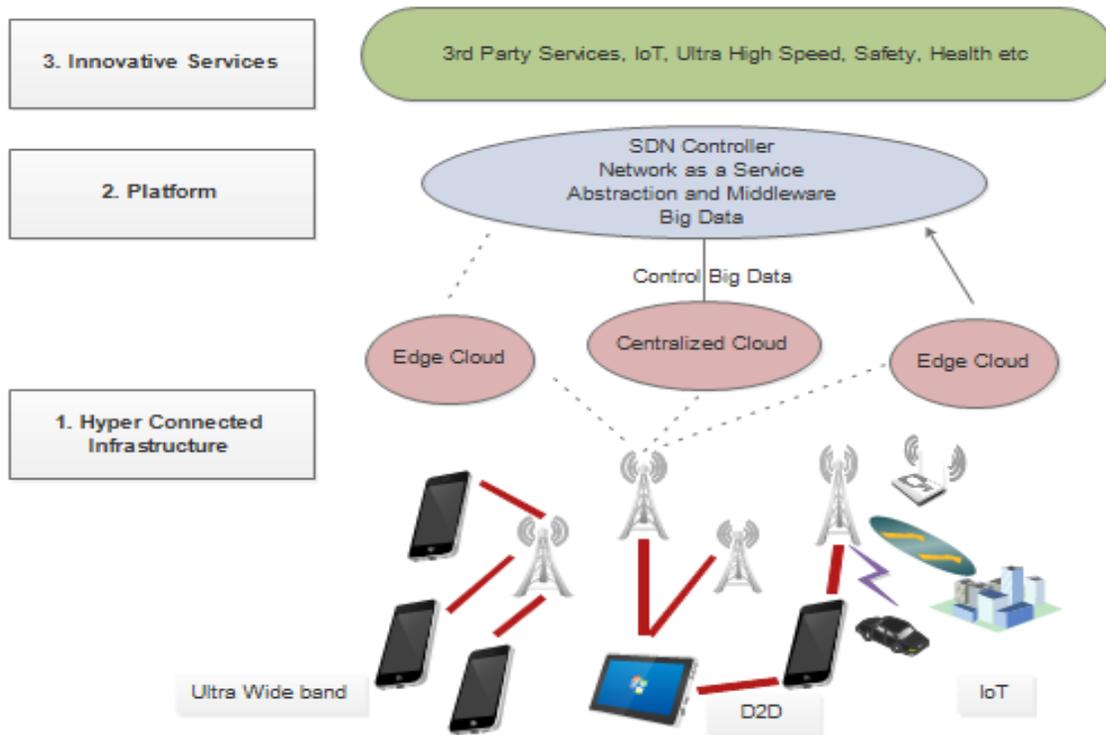


Fig. 3. 5G network architecture.

Intelligent security management, virtual security functions, authentication, authorization and residents isolation build up the concept of virtualized security (v-security) in 5G networks [60]. CHARISMA [60] with extension in OpenFlow protocol, intrusion detection, authentication, cluster encryption at physical layer, virtualization of network layer, and packet inspection defined a security protocol suitable in 5G network architecture. Comprehensive study in [61] on security of mobile networks revealed concerns and possible options to

handle these concerns. These both are really important to consider while adopting future network technologies that are evolution of mobile communication. Authors in [62] uncover security breaches that could occur due to DoS/DDoS attacks on control plane, system error or malicious software. The proposed security model in [62] is designed for telecommunication network where security of control and data plane, and the control- data interface is the major focus.

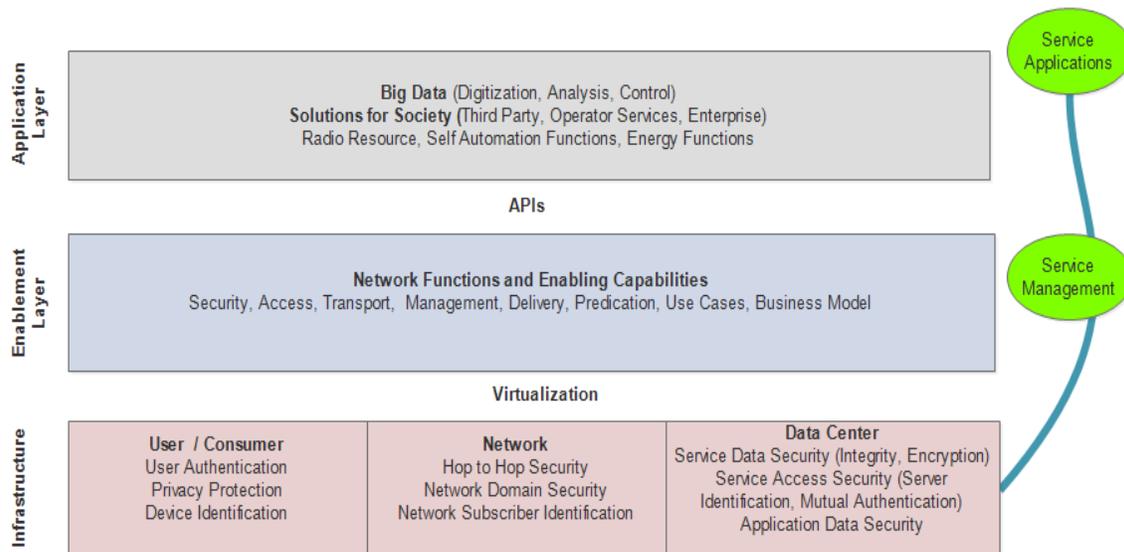


Fig. 4. 5G layered security vision.

## V. CONCLUSION

OpenFlow is considered in full scale deployments. Big giant such as Google, Cisco and many more cloud computing service providers are delivering services to their customers with SDN. Hence, security concern is considered to be important aspect of the technology. Very recently, plethora of research on this technology is suggested. Furthermore, number of researchers proved that, vulnerabilities do exist in the OpenFlow technology, namely, spoofing, tampering, repudiation, information disclosure, denial of service and elevation of privileges. These vulnerabilities can be had with adversary having access to single machine, group of machines and eventually taking control over complete network. Most of the suggested security modules with respect to OpenFlow are based on simulations of small scaled network. Undeniable, some of the methods considered here in the survey have considered lab based simulations involving very few processors. Hence, threats related to distributed attack such as DDoS is not full evaluated. Furthermore, latency sensitive applications running under SDN network with security enabled network elements is also not studied. Hence, before deployment of OpenFlow in a large scale requires through study of its vulnerabilities under the deliberation of Cloud benefits such as elasticity, on demand applications and smooth running of latency sensitive applications. The network is now moving from 3G, 4G to 5G. Cloud will help users to attain access to the capabilities of 5G. The evolution in network suggests modification in security models to gain user trust on network and services.

## REFERENCES

- [1] Alsmadi, Izzat, and Dianxiang Xu. Security of software defined networks: A survey, *Computers & Security* 53 (2015): 79-108.
- [2] Armbrust, Michael, and Armando Fox. Rean riffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia: Above the Clouds: A Berkeley View of Cloud Computing. Technical Report No. UCB/EECS-2009-28, 2009.
- [3] Bailey J, Pemberton D, Linton A, Pelsser C, Bush R. Enforcing rpk-based routing policy on the data plane at an internet exchange, In Proceedings of the third workshop on Hot topics in software defined networking 2014 Aug 22 (pp. 211-212). ACM.
- [4] Ballard, Jeffrey R., Ian Rae, and Aditya Akella. Extensible and Scalable Network Monitoring Using OpenSAFE. INM/WREN. 2008.
- [5] Benson, Theophilus, Aditya Akella, Anees Shaikh, and Sambit Sahu. CloudNaaS: a cloud networking platform for enterprise applications. In Proceedings of the 2nd ACM Symposium on Cloud Computing, p. 8. ACM, 2011.
- [6] Bernardos, Carlos J., Antonio De La Oliva, Pablo Serrano, Albert Banchs, Luis M. Contreras, Hao Jin, and Juan Carlos Zúñiga. An architecture for software defined wireless networking. *IEEE Wireless Communications* 21, no. 3 (2014): 52-61.
- [7] Blenk, Andreas, Arsany Basta, and Wolfgang Kellerer. HyperFlex: An SDN virtualization architecture with flexible hypervisor function allocation. In Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on, pp. 397-405. IEEE, 2015.
- [8] Braga, Rodrigo, Edjard Mota, and Alexandre Passito. Lightweight DDoS flooding attack detection using NOX/OpenFlow." In Local Computer Networks (LCN), 2010 IEEE 35th Conference on, pp. 408-415. IEEE, 2010.
- [9] Casado, Martin, Michael J. Freedman, Justin Pettit, Jianying Luo, Nick McKeown, and Scott Shenker. Ethane: taking control of the enterprise. In ACM SIGCOMM Computer Communication Review, vol. 37, no. 4, pp. 1- 12. ACM, 2007.
- [10] Cho, Hsin-Hung, Chin-Feng Lai, Timothy K. Shih, and Han-Chieh Chao. Integration of SDR and SDN for 5G. *IEEE Access* 2 (2014): 1196-1204.
- [11] Dabbagh, Mehiar, Bechir Hamdaoui, Mohsen Guizani, and Ammar Rayes. Software-defined networking security: pros and cons. *Communications Magazine, IEEE* 53, no. 6 (2015): 73-79.
- [12] Dejan Milojević and, Ignacio M. Llorente, and Ruben S. Montero. Opennebula: A cloud management tool. *Internet Computing, IEEE*, 15(2):11 –14, March-April 2011.
- [13] Fayazbakhsh, Seyed Kaveh, et al. FlowTags: enforcing network-wide policies in the presence of dynamic middlebox actions, Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking. ACM, 2013.
- [14] Feamster, Nick, Ankur Nayak, Hyojoon Kim, Russell Clark, Yogesh Mundada, Anirudh Ramachandran, and Mukarram Bin Tariq. Decoupling policy from configuration in campus and enterprise networks. In Local and Metropolitan Area Networks (LANMAN), 2010 17th IEEE Workshop on, pp. 1-6. IEEE, 2010.
- [15] Giotis, Kostas, Christos Argyropoulos, Georgios Androulidakis, Dimitrios Kalogeras, and Vasilis Maglaris. Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments. *Computer Networks* 62 (2014): 122-136.
- [16] H. Cui, G. O. Karame, F. Klaedtke and R. Bifulco, On the Fingerprinting of Software-Defined Networks, In *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 10, pp. 2160-2173, Oct. 2016.
- [17] He, Daojing, Sammy Chan, and Mohsen Guizani. Securing software defined wireless networks. *Communications Magazine, IEEE* 54, no. 1 (2016): 20-25.
- [18] Hinrichs, Timothy, Natasha Gude, Martin Casado, John Mitchell, and Scott Shenker. Expressing and enforcing flow-based network security policies. University of Chicago, Tech. Rep (2008).
- [19] Hu, Hongxin, Wonkyu Han, Gail-Joon Ahn, and Ziming Zhao. FLOWGUARD: building robust firewalls for software-defined networks. In Proceedings of the third workshop on Hot topics in software defined networking, pp. 97-102. ACM, 2014.
- [20] Jafarian, Jafar Haadi, Ehab Al-Shaer, and Qi Duan. Openflow random host mutation: transparent moving target defense using software defined networking. In Proceedings of the first workshop on Hot topics in software defined networks, pp. 127-132. ACM, 2012.
- [21] Kim, Hyojoon, and Nick Feamster. Improving network management with software defined networking. *Communications Magazine, IEEE* 51, no. 2 (2013): 114-119.
- [22] Kreutz, Diego, Fernando Ramos, and Paulo Verissimo. Towards secure and dependable software-defined networks. In Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking, pp. 55-60. ACM, 2013.
- [23] Lara, Adrian, and Byrav Ramamurthy. Opensec: A framework for implementing security policies using openflow. In Global Communications Conference (GLOBECOM), 2014 IEEE, pp. 781-786. IEEE, 2014.
- [24] Li, Wenjuan, Weizhi Meng, and Lam For Kwok. A survey on OpenFlow-based Software Defined Networks: Security challenges and countermeasures. *Journal of Network and Computer Applications* 68 (2016): 126-139.
- [25] Liyanage, Madhusanka, Ahmed Abro, Mika Ylianttila, and Andrei Gurtov. Opportunities and Challenges of Software-Defined Mobile Networks in Network Security Perspective. *IEEE Security and Privacy Magazine* (2015).
- [26] Liyanage, Madhusanka, et al. Security for future software defined mobile networks. Next Generation Mobile Applications, Services and Technologies, 2015 9th International Conference on. IEEE, 2015.
- [27] Ma, Duohe, Zhen Xu, and Dongdai Lin. Defending Blind DDoS Attack on SDN Based on Moving Target Defense. In International Conference on Security and Privacy in Communication Networks, pp. 463-480. Springer International Publishing, 2014.
- [28] Marinelli, Eugene E. Hyrax: cloud computing on mobile devices using MapReduce. No. CMU-CS-09-164. Carnegie-mellon univ Pittsburgh PA school of computer science, 2009.

- [29] Masoudi, Rahim, and Ali Ghaffari. Software defined networks: A survey, *Journal of Network and Computer Applications* 67 (2016): 1-25.
- [30] Matias, Jon, Eduardo Jacob, David Sanchez, and Yuri Demchenko. An OpenFlow based network virtualization framework for the cloud. In *Cloud computing technology and science (CloudCom)*, 2011 IEEE Third International Conference on, pp. 672-678. IEEE, 2011.
- [31] McKeown, Nick, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. OpenFlow: enabling innovation in campus networks. *ACM SIGCOMM Computer Communication Review* 38, no. 2 (2008): 69-74.
- [32] Nayak, Ankur Kumar, Alex Reimers, Nick Feamster, and Russ Clark. Resonance: dynamic access control for enterprise networks. In *Proceedings of the 1st ACM workshop on Research on enterprise networking*, pp. 11- 18. ACM, 2009.
- [33] Paim de Jesus, Wanderson, Daniel Alves da Silva, Rafael T. de Sousa, and Francisco Vitor Lopes Da Frota. Analysis of SDN Contributions for Cloud Computing Security. In *Utility and Cloud Computing (UCC)*, 2014 IEEE/ACM 7th International Conference on, pp. 922-927. IEEE, 2014.
- [34] Parker, M. C., G. Koczian, F. Adeyemi-Ejeye, T. Quinlan, S. D. Walker, A. Legarrea, M. S. Siddiqui et al. CHARISMA: Converged Heterogeneous Advanced 5G Cloud-RAN Architecture for Intelligent and Secure Media Access.
- [35] Porras, Philip, Seungwon Shin, Vinod Yegneswaran, Martin Fong, Mabry Tyson, and Guofei Gu. A security enforcement kernel for OpenFlow defined networks. In *Proceedings of the first workshop on Hot topics in software defined networks*, pp. 121-126. ACM, 2012.
- [36] Qazi, Zafar Ayyub, Cheng-Chun Tu, Luis Chiang, Rui Miao, Vyas Sekar, and Minlan Yu. SIMPLE-fying middlebox policy enforcement using SDN. In *ACM SIGCOMM computer communication review*, vol. 43, no. 4, pp. 27-38. ACM, 2013.
- [37] Raghavendra, Ramya, Jorge Lobo, and Kang-Won Lee. Dynamic graph query primitives for sdn-based cloud network management. In *Proceedings of the first workshop on Hot topics in software defined networks*, pp. 97-102. ACM, 2012.
- [38] Rekha, P. M., and M. Dakshayini. Dynamic network configuration and Virtual management protocol for open switch in cloud environment. In *Advance Computing Conference (IACC)*, 2015 IEEE International, pp. 143- 148. IEEE, 2015.
- [39] Roberto Bifulco, Julien Boite, Mathieu Bouet, and Fabian Schneider. 2016. Improving SDN with InSPired Switches. In *Proceedings of the Symposium on SDN Research (SOSR '16)*. ACM, New York, NY, USA,
- [40] Sasaki T, Perrig A, Asoni DE. Control-plane isolation and recovery for a secure SDN architecture. In *NetSoft Conference and Workshops (NetSoft)*, 2016 IEEE 2016 Jun (pp. 459-464). IEEE.
- [41] Schehlmann, Lisa, Sebastian Abt, and Harald Baier. Blessing or curse? Revisiting security aspects of Software- Defined Networking. In *Network and Service Management (CNSM)*, 2014 10th International Conference on, pp. 382-387. IEEE, 2014.
- [42] Shalinie, S. Mercy, and R. Nagarathna. An Intelligent Prototype to Lay the Road to Secure Next Generation Networks. *Cover Story* (2012): 29.
- [43] Shimonishi, Hideyuki, and Shuji Ishii. Virtualized network infrastructure using OpenFlow. In *Network Operations and Management Symposium Workshops (NOMS Wksp)*, 2010 IEEE/IFIP, pp. 74-79. IEEE, 2010.
- [44] Shin, Seungwon, Phillip A. Porras, Vinod Yegneswaran, Martin W. Fong, Guofei Gu, and Mabry Tyson. FRESCO: Modular Compassable Security Services for Software-Defined Networks. In *NDSS*. 2013.
- [45] Shu, Zhaogang, Jiafu Wan, Di Li, Jiayang Lin, Athanasios V. Vasilakos, and Muhammad Imran. Security in Software-Defined Networking: Threats and Countermeasures. *Mobile Networks and Applications*: 1-13.
- [46] Stabler, Greg, Aaron Rosen, Sebastien Goasguen, and Kuang-Ching Wang. Elastic IP and security groups implementation using OpenFlow. In *Proceedings of the 6th international workshop on Virtualization Technologies in Distributed Computing Date*, pp. 53-60. ACM, 2012.
- [47] Tsugawa, Maurício, Andréa Matsunaga, and José AB Fortes. Cloud Computing Security: What Changes with Software-Defined Networking?. In *Secure Cloud Computing*, pp. 77-93. Springer New York, 2014.
- [48] Voellmy, Andreas, Hyojoon Kim, and Nick Feamster. Procera: a language for high-level reactive network control. In *Proceedings of the first workshop on Hot topics in software defined networks*, pp. 43-48. ACM, 2012.
- [49] Wang, Bin, Zhengwei Qi, Ruhui Ma, Haibing Guan, and Athanasios V. Vasilakos. A survey on data center networking for cloud computing. *Computer Networks* 91 (2015): 528-547.
- [50] Wang, Kai, Yaxuan Qi, Baohua Yang, Yibo Xue, and Jun Li. LiveSec: Towards effective security management in large-scale production networks. In *Distributed Computing Systems Workshops (ICDCSW)*, 2012 32nd International Conference on, pp. 451-460. IEEE, 2012.
- [51] Y. Pu, Y. Deng, and A. Nakao, Cloud rack: Enhanced virtual topology migration approach with open vswitch, in *Information Networking (ICOIN)*, 2011 International Conference on, jan. 2011, pp. 160 –164.
- [52] Yan Z, Prehofer C. Autonomic trust management for a component-based software system. *IEEE Transactions on Dependable and Secure Computing*. 2011 Nov; 8(6):810-23.
- [53] Yang, Fengyi, Haining Wang, Chengli Mei, Jianmin Zhang, and Min Wang. A flexible three clouds 5G mobile network architecture based on NFV & SDN. *China Communications* 12, no. Supplement (2015): 121-131.
- [54] Yao, Guang, Jun Bi, and Peiyao Xiao. Source address validation solution with OpenFlow/NOX architecture. In *Network Protocols (ICNP)*, 2011 19th IEEE International Conference on, pp. 7-12. IEEE, 2011.
- [55] Zaalouk, Adel, Rahamatullah Khondoker, Ronald Marx, and Kpatcha Bayarou. OrchSec: An orchestrator-based architecture for enhancing network-security using Network Monitoring and SDN Control functions. In *Network Operations and Management Symposium (NOMS)*, 2014 IEEE, pp. 1-9. IEEE, 2014.

# Analysis of Received Power Characteristics of Commercial Photodiodes in Indoor LoS Channel Visible Light Communication

Syifaul Fuada<sup>1</sup>, Angga Pratama Putra<sup>2</sup>, Trio Adiono<sup>3</sup>

<sup>1,2,3</sup> University Center of Excellence on Microelectronics, Institut Teknologi Bandung  
PAU Building 4<sup>th</sup> floor, Tamansari street No.126, Bandung city, West Java, ZIP 40132  
INDONESIA

**Abstract**—To date, the photodiode still the first choice component is used in optical communication, especially for visible light communication (VLC) system. It has advantages of speed, energy consumption, and sensitivity, compared to other devices (e.g. image sensor). There are many practical implementations of high-speed VLC which uses photodiode. Commercially available photodiode typically have specific characteristics, so that it needs some consideration to be used as optimal receiver devices in VLC system. In this paper, analysis of received power characteristics of the photodiode in indoor line-of-sight (LoS) channel of VLC system is discussed. MATLAB® simulation is used as approach model (student version). The experiments are done by changing several parameters such as the semi-angle half power of the transmitter, distance from the transmitter to receiver, room size, field-of-view (FOV), lens index and optical filter gain. From the results, it can be known that distance, room size, FOV and LED power factor to have linear characteristic against the received power of commercial photodiode. Also in LoS channel model, the gain of optical filter and lens index plays an important role in defining the characteristics of received power.

**Keywords**—Commercial photodiodes; LoS channel; power received; visible light communication

## I. INTRODUCTION

In recent decades, many researchers are interested in research of Visible Light Communication (VLC) for various application, such as vehicle to vehicle communication, Light Fidelity (Li-Fi), hospitals data communication, real-time audio & video transmission, underwater communication, space communication, localization based movable devices (mobile robot or autonomous robot), WLANs, visible light ID system and so on.

There are many research about comparative study in VLC system that is interesting to be discussed to gain deeper understanding to develop VLC system, such as different modulation technique in VLC system [1], comparison study of OFDM multiplexing schemes (DCO, ACO, ADO) [2], placement optimization of LED-array as emitter [3], available bandwidth through red, green and blue phosphor LED [4], effect of color filter in VLC physical layer system using mica paper [5]-[6], noise analysis using variety Op-Amp and photodiode for VLC system [7]-[8], single versus multicarrier performance analysis [9], performance comparison using

variety QAM modulation from 4 to 512 based RGB LED [10], analysis of different LED array spacing [11], VLC system performance within and without analog filters [12], decoder performance within and without Viterbi [13], and so on.

Photodiode is a common photodetector device that can be used for precision measurement or optical communication application, e.g. VLC, fiber optic and infrared communication. Compared with other devices, such as the light dependent resistor (LDR), photo-IC, solar cell and phototransistor, The photodiode has several advantages in stability, precision and response time. The commercial photodiode can be divided into several types, those are: 1) precision photodiode that can be properly used for light measurement; 2) high-speed photodiode that has general characteristic of high cut-off frequency, which is suitable for optical application; or 3) integrated photodiode such as S8475 and S9295, manufactured by HAMAMATSU®, which already integrated with pre-amp in a single chip. Each type of photodiode has its own advantages and disadvantages.

There are various studies about the comparison of photodiode types for application of optical communication. The photodiode selection is important in communication system because it can affect sensitivity, speed, range, reliability, cost, and another factor in the communication system. Research scheme of photodiode types already been done by A. Boudkhal, et al. [14] who compares noise performances of PIN and APD photodiodes through an optical high debit transmission chain. Then P. Sharma, et al [15] compares PIN and APD performances with different modulation and wavelength of LED transmitter. Also M.A.A Ali [16] analyze APD performances for underwater communication application through combination scheme of the Jerlov water variable types (I, IA, IB) and photodiode material types (Si, Ge, InGaAs). Then O. Kharraz and D. Forsyth [17] analyze optical excess noise and thermal noise which exist in PIN and APD. Then Y. Chen, et al [18] experiment about optimizing collimating lens of the photodetector for supporting long range VLC. The author himself already done the investigation on capacitor junction (Cj) effect on total noise, which is an RMS function of voltage noise, current noise and feedback resistor in discrete TIA circuit [19]. All of the above six experiments are done through analytical calculus approach and proven by MATLAB® and other specific simulator tools.

The VLC system can be divided into three main parts, transmitter, channel and receiver. Channel in VLC system is a free space which can be implemented as Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) link. In LoS, one of its weak points is on shadowing effect which is caused by object blocking, such as by household equipment or human activities. Another LoS weak point is its limited covering area capability, so it's incapable of supporting mobile user because the LoS configuration requires transmitter and receiver to be placed in a straight line. The solution of this problem is by using photodiode which has broad FOV characteristic. Besides that, the LoS advantage is on its characteristic that can support high-speed data transfer for a relatively long distance and its invulnerability of distortion from multipath signal induction and ambient light noise. Illustration of the LoS link is shown in Fig. 1(a), in which photodiode as receiver placed on a straight line from LED. This link angular distribution is shown in Fig. 1(b). The mathematical derivation of the LoS link is explained in Section III of this paper.

The light information which transmitted from the LED will be weakened (fading) while transmitted on the free space channel, it means the farther the distance of the receiver, the weaker the signal received, and the information may not be received at all [20]. For that problem, the solution is to increase the LED power or to add more LED as the transmitter. But this is not the best solutions because it is not power efficient. Besides, adding more LED will add another problem, roaming. The ideal solution is by selecting the proper photodiode and optimizing the photodiode filter.

The ideal photodiode characteristics can be found in the datasheet from each manufacturer, where its specification can be analyzed through finding the relations of LED power of the VLC system with the received power in the photodiode (see Section III). By using this method, we can accurately predict the performance of the photodiode which will be implemented in VLC system. Based on the observations from many works of literature until this far, the discussion about photodiode's received power characteristic of the different commercially available photodiode is still rarely found. Related research has

been discussed by *K. Lee, et al.* [21] which analyzes the effect of photodiode's received power with LoS and NLoS scenarios with different wall type. Since this paper isn't exploring the received power characteristic based on different photodiode manufacturer, on this paper we will discuss that characteristic of the photodiode based on different manufacturer. The motivation of writing this paper is to fill that area of study.

Besides that, we also done other experiments that observing the effect of changing several parameters of the LoS channel against the received power on the photodiode, such as 1) changing semi-angle of the transmitter; 2) varying power of a single LED; and 3) changing the distances of the channel. After that, the effect of FOV, room dimension and internal concentrator of the photodiode against its received power characteristic will also be investigated. To find the ideal value of this characteristics, these experiments using simulation based approach using MATLAB® have been performed.

This paper is divided into several parts. The first part is an overview of VLC system, research area, channel system, problems and purpose of the experiments. The second part explained the photodiode consideration and several types of the photodiode. The third part explained the detail of the LoS channel which has been introduced in the introduction part. The fourth part discusses the experiment set-up, results, and analysis. And the last part consisted of conclusion, acknowledgment, and references.

## II. PHOTODIODE CONSIDERATIONS IN VLC SYSTEMS

Several considerations on selecting commercial photodiode for VLC application is as follows: 1) surface area; 2) generated short current; 3) capabilities to detect wavelength; 4) frequency cut-off; 5) rise-time; and 6) dark current and internal capacitance/junction capacitor ( $C_j$ ).

With broad surface area, for example, 10 mm x 10 mm, the photodiodes can be used to support mobility in VLC system. This sensing area capability can be improved by arranging the photodiode in an array setup such as being done by *J.H. Li, et al.* [22].

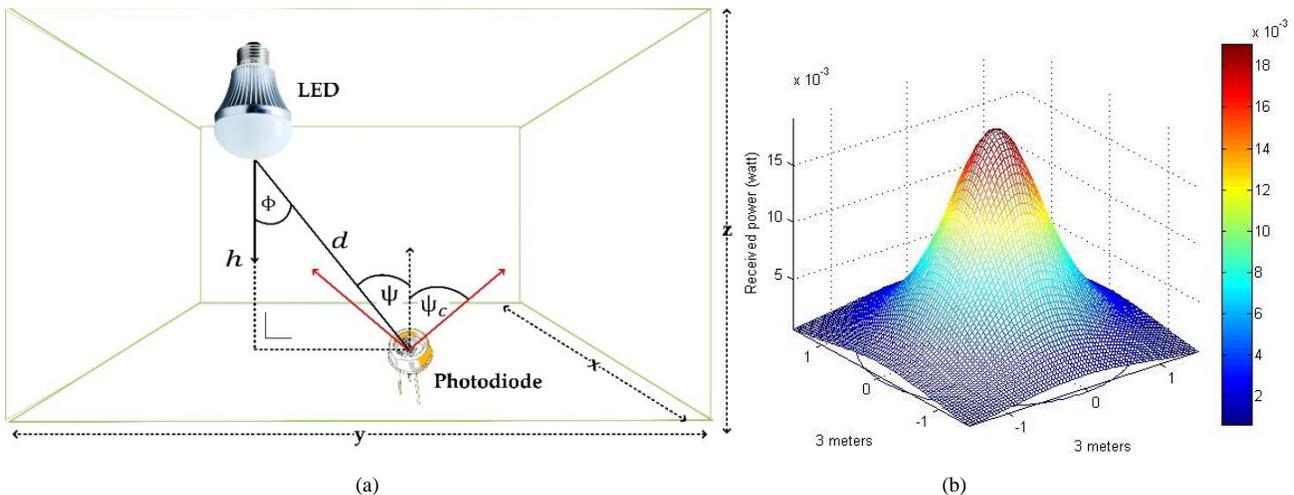


Fig. 1. (a) Geometry of LoS channel configuration in 3 m x 3 m x 3 m room; (b) MATLAB® simulation of light intensity distribution in room using LED with characteristics of 30 Watt,  $d = 1$  meter, FOV = 30°, number of transmitter = 1, number of grid = 25,  $A_r = 1 \text{ mm}^2$ , and  $\phi_{1/2} = 60^\circ$ .

Although, the larger the area, the cut-off frequency will become narrower and tend to be easily disturbed by the ambient light noise. For the high-speed application, this characteristic needs to be considered, to make the robust VLC system design that has high tolerances characteristic with noise from other light sources, such as sunlight.

After that, we also need to consider the short current. As we all know, the photodiode will generate current with linear characteristic against the light intensity, the brighter the light intensity, the higher the current that will be generated, at least around  $>100\mu\text{A}$ . The lower the short current, the higher the gain of the Op-Amp, and it will affect on the narrower bandwidth. The cut-off frequency ( $f_{-3\text{dB}}$ ) is inversely proportionated with the division of Op-Amp Gain Bandwidth Product (GBW) and voltage gain ( $A_v$ ). Another factor that needs to be considered is the wavelengths that can still be detected. On VLC application, the chosen the photodiode needs to has capabilities to senses wavelength in visible light spectrum range, i.e. 380 nm to 780 nm. Mistakes on the selection of the photodiode can affect the system to not be able to work optimally.

A photodiode with high cut-off frequency (~GHz scale) and fast rise time (~nanosecond scale) characteristics can be used for high-speed optical communication, although typically these characteristics have to trade off with the narrow sensing area (around 0.1 mm), so while it supports high-speed data transfer, it doesn't support the mobility of VLC system, and the receiver needs to have  $0^\circ$  elevation angle. Even though, in general, the VLC needs to able to provides mobility characteristic. This problem has several solutions such as using optic concentrator (collimator lens or polarizer) to focusing the light into the photodiode, adding more information light sources (for example, array LED based inter-cell system) or arranging the photodiode in an array based setup.

The dark current is generated current from a photodiode in a dark condition (no-light). The chosen PD needs to have low dark current characteristic and also low the  $C_j$ . In previous research, the lower the  $C_j$ , the higher the noise on the photodiode and the slower the response of photodiode amplifier. Both of these characteristics are important, and it will affect noise on the photodiode which have strong relations with  $e_n$  and  $i_n$  on the chosen Op- Amp [23].

All of those six factors above, can't be obtained simultaneously. Commercially available photodiode, typically only have one or two of those characteristics (no more than three), thus the selection process must be thorough. These limitations can be used as main consideration to create self-made photodiode in a research based method, as been done by *H. Chen, et al.* [24] and *W. Zou, et al.* [25].

In VLC practical demonstration, there are two types of the photodiode that are of primary interest, i.e. Positive-Intrinsic-Negative (PIN) and Avalanche photodiode (APD). The characteristics of both of these photodiodes has been discussed by *M. Azadeh* [26]. The comparison of PIN and APD is shown in Table 1. These data are gathered from many work of literature. To make it short, PIN provides higher sensitivity, higher bandwidth, lower operating voltage and also cheaper.

TABLE I. COMPARISON BETWEEN PIN AND APD IN VLC SYSTEM

Variable	PIN	APD
Materials	Si, Ge, InGaAs	
Bandwidth	To 40 GHz	
Life time	OK	
Spectral range	Tunable (Ultraviolet, Visible light, Near Infra-red)	
Form factor	Small	
Electromagnetic immunity	No	
Magnetic field sensitivity	No	
Large area	No	
Gain	1	$10^2$
Operating Voltage (V)	Low (0 - 5)	High (100 -100k)
Cost	Low	High
Efficiency (A/W)	Low	High
Response time	Fast	Slow
Sensing sensitivity	Low	High
Temperature sensitivity	Low	High
Bandwidth & bit rate	High	Medium
Damage by Stray light	No	Yes
Dark current	High	Low
Excess noise factor	Low	Medium
Mechanical Robustness	High	Medium

### III. THE LOS CHANNELS DESCRIPTION

As shown in Fig. 1(a), LED placed at height  $h$  relative at region 'x' and 'y' from the receiver. LED radiation angle of the transmitter to the receiver against transmitter normal, denoted with  $\phi$ . Whereas LED radiation angle to the receiver against receiver normal, denoted with  $\psi$ , where the receiver has a FOV. LED radiation can only be sensed while on the FOV range, where maximum angle range against the receiver normal denoted with  $\psi_c$ .

Fig. 1(b) shows the distribution of information light intensity inside a room that has an uneven distribution. This uneven distribution is shown with different color gradation in several areas. The maximum power that can be received by the photodiode on distances less than 250 cm is 1.4 to 1.2 dBm. Whereas, for distances, more than 500 cm, the power that can be received by the photodiode is around 1 dBm. The photodiode has three main part, those are: 1) the concentrator (*coating*) to focusing light; 2) the filter for passing signals only at a certain frequency range, or as band pass filter, so that noise of ambient light can be reduced; and 3) the photodetector to converts light into electrical currents. Illustration of the photodiode parts is shown in Fig. 2.

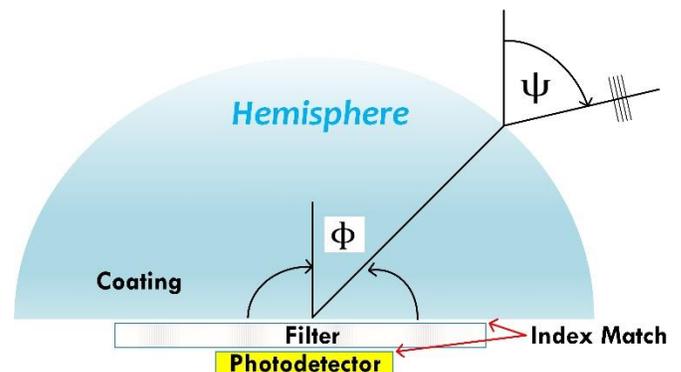


Fig. 2. The photodiode configuration.

Coating and filter have an effect on the received light on the photodiode, the coating will yield different refractive index of light which propagated from outside of the photodiode, so light propagation direction is changed. This phenomenon will be affecting the FOV of the receiver, as well as an effective region of the receiver ( $A_{eff}$ ) which denoted as  $g(\psi)$ , where  $n$  parameter is the refractive index of the concentrator and  $\psi_c$  is the maximum angle of FOV on the photodiode. The filter of photodiode is also affecting receiver effective region, this factor is denoted as  $T_s(\psi)$ .

$$g(\psi) = \begin{cases} \frac{n^2}{\sin^2 \psi_c}, & 0 \leq \psi \leq \psi_c \\ 0, & \psi > \psi_c \end{cases} \quad (1)$$

Then it will be shown how coating and filter will affecting  $A_{eff}$ . First is by ignoring the filter effect, concentrator effect, and losses in reflection. The receiver will receive the light radiation at an effective area which can be written mathematically as (2), where  $A_r$  is the surface area of the photodiode. Then by adding parameter of coating factor,  $g(\psi)$  and filter factor  $T_s(\psi)$ . The  $A_{eff}$  can be expressed as (3).

$$A_{eff}(\psi) = \begin{cases} A_r \cos \psi, & 0 \leq \psi \leq \pi/2 \\ 0, & \psi > \pi/2 \end{cases} \quad (2)$$

$$A_{eff}(\psi) = \begin{cases} A_r \cos \psi T_s(\psi)g(\psi), & 0 \leq \psi \leq \psi_c \\ 0, & \psi > \psi_c \end{cases} \quad (3)$$

For the relationship between optical power transmitted by LED and received by the LED. In this case, the frequency response of the transmitted and received visible light is flat enough and can be denoted as DC gain ( $H_d(d)$ ). Thus, the relations between optical received power ( $P_r$ ) in watt and optical transmitted power ( $P_t$ ) could be expressed as (4).

$$P_r = H_d(d)P_t \quad (4)$$

Where,

$$P_t = \frac{I(\phi)}{R(\phi)} \quad (5)$$

Where,  $R(\phi)$  is radiant Lambertian, then could be expressed as Eq. 6,  $\phi$  is the angle irradiance form of the LED,  $m$  is the order of the Lambertian emission which defined by LED's semi-angle at half power ( $\phi_{1/2}$ ), where  $m = \ln(\frac{1}{2}) / \ln(\cos(\phi_{1/2}))$ .

$$R(\phi) = \frac{m+1}{2\pi} \cos^m(\phi) \quad (6)$$

The intensity of light received by the photodiode has a dividing factor of the quadratic of the distance ( $d^{-2}$ ) of the intensity transmitted by LED. Whereas, received power is a product of received intensity against the effective area of the receiver. Therefore, optical received power  $I_t(\phi)$  can be expressed as (7).

$$P_r = I_r(\phi) A_{eff}(\psi) \quad (7)$$

Where,  $P_r$  LOS is also equal to (8).

$$P_r = \frac{I(\phi)}{d^2} A_{eff}(\psi) \quad (8)$$

Variable  $d$  denotes the distance between the LED and the photodiode. By substituting (1), (2) and (5) to the (8), DC channel gain function will be obtained (9). Where  $H_d$  is equal to  $H_{LOS}$ .

$$H_d(0) = \begin{cases} \frac{A_r(m+1)}{2\pi} \cos^m(\phi) \cos \psi T_s(\psi)g(\psi), & \text{untuk } 0 \leq \psi \leq \psi_c \\ 0, & \text{others } (\psi > \psi_c) \end{cases} \quad (9)$$

$$P_r(abm) = 10 \log_{10}(P_r(watt)) \quad (10)$$

According to calculation, it can be shown that (4) is a multiplication of (9) with  $T_s(\psi)$  and (1), then from (4), an 3D model can be simulated by MATLAB® using parameters that will be obtained in Section IV. To make the analysis become easier, the unit of watt will be transformed into dBm through (10). Another discussion about LoS channels of VLC can be shown in [27]-[29].

## IV. DISCUSSION

### A. Experiment Set-up

In this experiment, the photodiode manufactured by OSRAM optoelectronics is used. From its datasheet, it can be obtained information that has been explained in Section III. There are five PIN type of photodiode that is used: BPW21, BPW34B, BPX65, SFH213, SFH221. The detail is shown in Table 2, where there are nine variables, i.e., the features of the photodiode,  $A_r$ ,  $I_{sc}$ ,  $C_t$ ,  $\psi_{1/2}$ ,  $T_s(\psi)$ ,  $P_{tot}$ ,  $\lambda$ , and  $S$ . While parameters for simulation experiment is shown in Table 3, where there are three scenarios of simulation. Variable  $A_r$  in this experiment is used as static variable because this variable is an intrinsic variable of the photodiode and can't be changed.

Based on the recommendation from [30], for an indoor application (assumed dimension of 3 m x 3 m x 3 m) minimum lumen requirement is around 250 to 500 lm/m<sup>2</sup>. Based on that information, LED that is capable of works in that lumen range and has a maximum power of 5 Watt is chosen. The chosen LED is CREE XLamp® XT-E LE which has maximum ~629 lm and configured in parallel so its maximum power is 50 Watt.

The receiver devices is placed at the distances of 3 meters from the information source, and the LED placed on the coordinat (1.5, 1.5, 3), or exactly at the middle of the room. Because in this experiment is use the LoS channel, the reflectivity of the wall can be ignored. In the datasheet, the gain filter and the concentrator is not specified, so these also ignored. The effect of changing transmitter's placement coordinate is not addressed in this paper because in this calculation only use a single LED which has been explored before on [11] and [31]. Since the price of the photodiode was changed in every time, the cost factor is not use in product comparison.

On the scenario A, B, and C, (9) is used to be computed in MATLAB®, however in the datasheet of each photodiode, parameters  $T_s(\psi)$  and  $g(\psi)$  is not specified, so calculation references will be based on (9) and (10).

TABLE II. PHOTODIODE SPECIFICATION

Variable	Notation	BPW21	BPW 34B	BPX 65	SFH 213	SFH 221
Features	-	Si PD for the Visible Spectral Range	Si PD with Enhanced Blue Sensitivity	Si PIN PD	Si PIN PD	Si Dual PD
Physical area of photo-detector ( $mm^2$ )	$A_r$	7.45	7.45	1	1	1.54
Spectral response range ( $nm$ )	$\lambda$	350 – 820	350 – 1100	350 – 1100	400 – 1100	400 – 1100
Short circuit current at 100 lux ( $\mu A$ )	$I_{sc}$	10	7.4	10	125	24
Terminal Capacitance ( $pF$ )	$C_t$	580	72	11	11	25
Half angle ( $^\circ$ )	$\psi$	55	60	40	10	55
Gain of optical filter	$T_s(\psi)$	-	-	-	-	-
Spectral sensitivity ( $nA/lx$ )	S	10	75	10	-	24
Total power dissipation ( $mW$ )	$P_{tot}$	250	150	250	150	50

TABLE III. EXPERIMENTAL PARAMETERS

Variable	Notation	Scenario A	Scenario B	Scenario C
Room dimension	$w \times l \times h$	3m x 3m x 3m		
Transmitter coordinator	-	Center (1.5, 1.5, 3)		
Number of Transmitter	-	Single LED		
Reflectivity of wall	$\gamma$	ignored		
PD concentrator refractive index	$n$	ignored		
Gain of optical filter	$T_s(\psi)$	ignored		
Field of view (FOV) semi angle of the receivers	$\psi_{1/2}$	Sesuai kemampuan PD		
Transmitter's semi-angle at half power	$\phi_{1/2}$	15°, 30°, 45°, 60°, 75°, 90°	45°	
Distance between LED and PD	$d$	2 m	0.5m, 1 m, 1.5m, 2 m, 2.5m, 3m	2m
Maximum optical power of LED	$P_{LED}$	50 W		5W, 10W, 20W, 30W, 40W, 50W

$$\frac{A_r(m+1)}{2\pi} \cos^m(\phi) \cos \psi \quad (11)$$

B. Scenario I

In this scenario, transmitter's angle is a function of received power. Settings of this scenario are  $\psi_{1/2}$  = capability of the photodiode, where this parameter can be found in Table 2. The channel distance is fixed, i.e. 2 meters, with transmitter power 50 Watt and  $\phi_{1/2}$  is changed from minimal 15° and maximum 90° with range difference of 15°. The result of the simulation is shown in Fig. 3.

From that figure, it can be known that the larger  $\phi_{1/2}$ , the smaller received power in the photodiode. This is matched with the characteristic of the LoS channel, where the received power will be larger if the deviation angle of the receiver from the transmitter is closer to 0°.

BPX65 and SFH213 have similarities in physical area, the same with BPX65 and SFH213. Even though there are differences in FOV, the received power is relatively the same if semi-angle half power is changed. This is because  $T_s(\psi)$  and the refractive index of the lens ( $n$ ) are not included in the calculation.

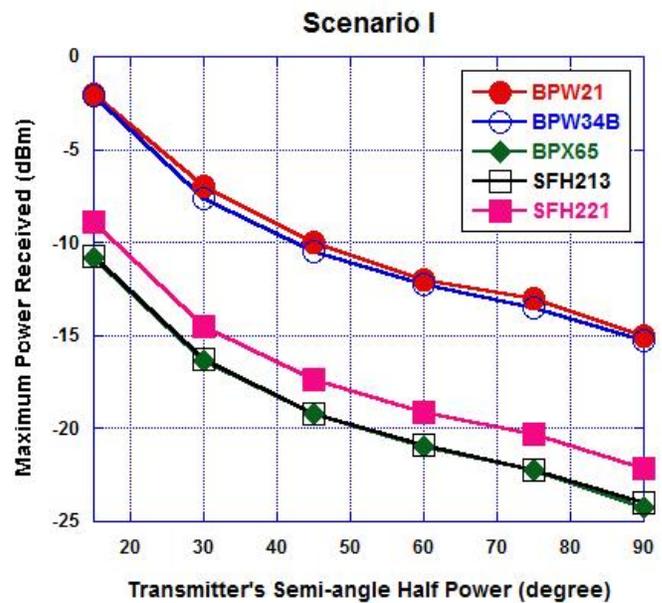


Fig. 3. Semi-angle half power of the transmitter vs photodiode's received power.

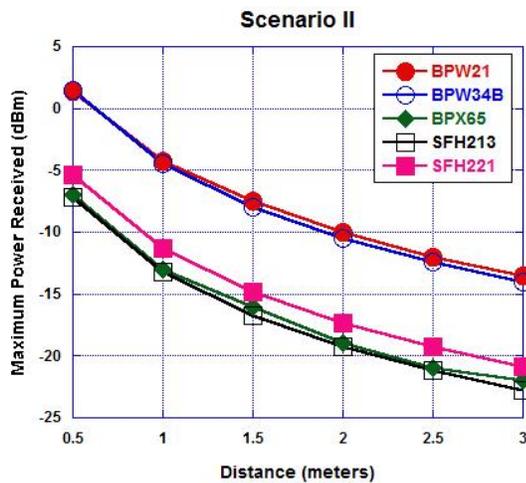


Fig. 4. Distance vs photodiode's received power.

The result of this simulation is shown in Fig. 4, where it can be known that the further the distance, the smaller the received power at the photodiode. It is because of the characteristic of the channel LoS, where the closer the transmitter to the photodiode, the higher the intensity of light that is received by the photodiode.

Short current is the current which generated by the photodiode that is linear with light intensity. But, the smaller the variable  $d$ . Therefore, to simulate the real condition, the minimum distance should be 1 meter.

### C. Scenario III

In this scenario, the effect of changing the LED power is observed. The parameters of this simulation is as follows: 1)  $d = 2$  meters, this is the fixed distance of LED to photodiode, and LED to the object (e.g. Table 2), it means the height of the object is assumed to be 1 meter, 2)  $\phi_{1/2}$  of  $45^\circ$  and transmitter power varied from 10 Watt, 20 Watt, 30 Watt, 40 Watt and 50 Watt. On the implementation, power setting can be done by configuring the forward voltage ( $V_f$ ) of the LED. The larger the  $V_f$ , the larger the power. The result of this simulation is shown in Fig. 5, where LED power is linear with the received power at the photodiode.

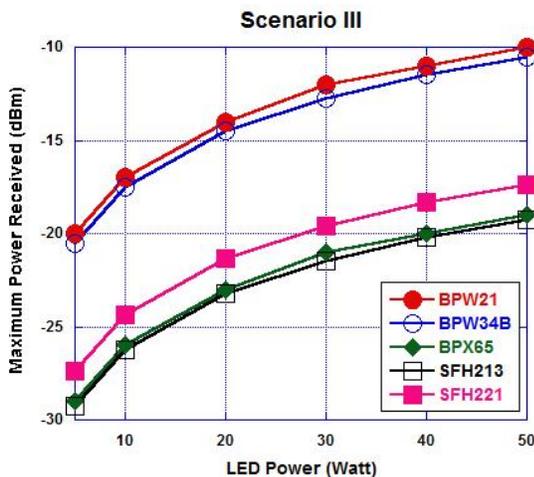


Fig. 5. LED power vs photodiode's received power.

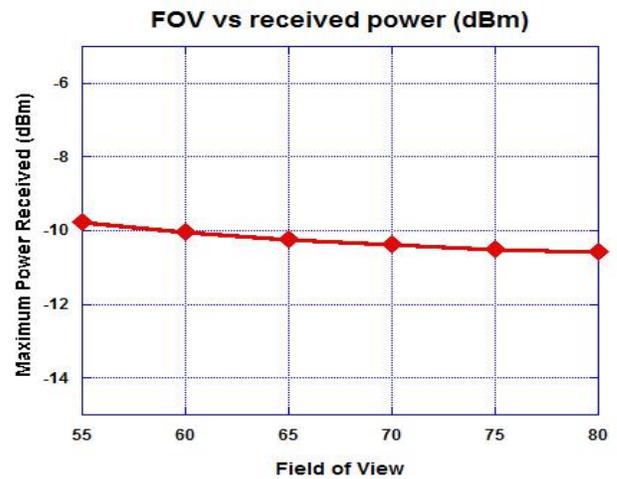


Fig. 6. FOV vs photodiode's received power.

### D. FOV of the Photodiode

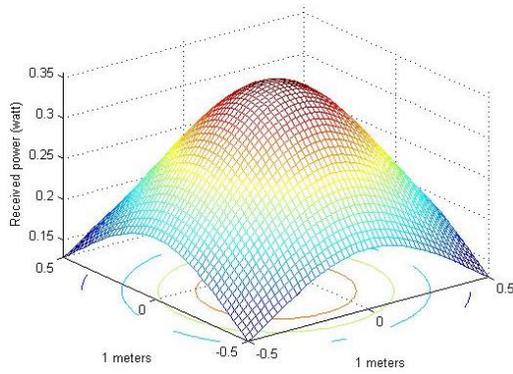
As has been addressed on the background, on this paper, the authors are also interested to find the effect of difference photodiode's Semi-angle Half power FOV against the received power. The parameter of this simulation is shown in Table 2, with  $\phi_{1/2} = 30^\circ$ , LED power = 50 Watt,  $d = 2$  meters. The photodiode is chosen with large semi-angle characteristic, that is BPW34B ( $A_r = 1 \text{ mm}^2$ ), then the value of  $\psi_{1/2}$  will be varied, from 55, 60, 65, 70, 75 to 80. Then the value of  $T_s(\psi) = 1$  and  $n = 1$ . The result of this simulation is shown in Fig. 6 and it shows the characteristic of FOV on the photodiode is affecting the received power, although insignificantly. Therefore, even though insignificant, the value of FOV can be used as consideration in choosing photodiode for VLC application.

### E. Effect of Changing Room Size

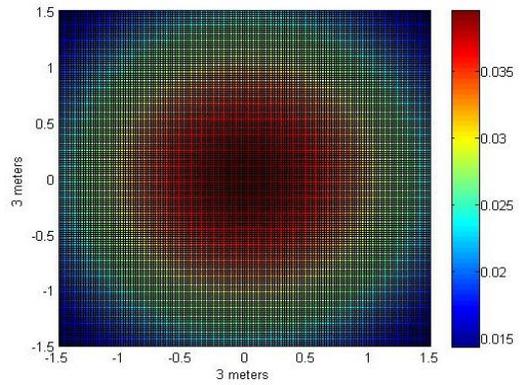
The purpose of this experiment is to prove that "the dimension of the room is linear with the received power". This experiment use photodiode BPW 34B with  $A_r = 7.45 \text{ mm}^2$ ,  $\phi_{1/2} = 45^\circ$ , and LED power = 50 Watt. Fig. 7 shows the result of the simulation from the side view with different room variation as follow: a) 1 m x 1 m x 1 m with  $d = 1$  meter, in 3D view; b) 2d view of (a); c) 3 m x 3 m x 3 m with  $d = 3$  meters in 3 dimension view; d) 2 dimension view of (c); e) 5 m x 5 m x 5 m with  $d = 5$  meters in 3D view; f) 2D view of Fig. 7(e). The results show that the hypothesis is correct, that the larger the dimension of the indoor room, the weaker the intensity of light and the distribution of light can't reach small sides in the room.

### F. Effect of Changing Filter and Concentrator

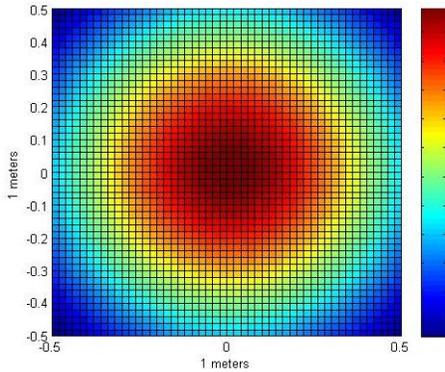
It has been addressed in Section III that concentrator is a part of the photodiode, even if on the datasheet (as shown in Table 2),  $T_s(\psi)$  and  $g(\psi)$  are not specified. Because of that, in this paper, will be investigated if the changes of both of those variable affecting the photodiode significantly. On the implementation, if on  $T_s(\psi)$  and  $g(\psi)$  of the photodiode is not available, the filter and the external concentrator can be added.



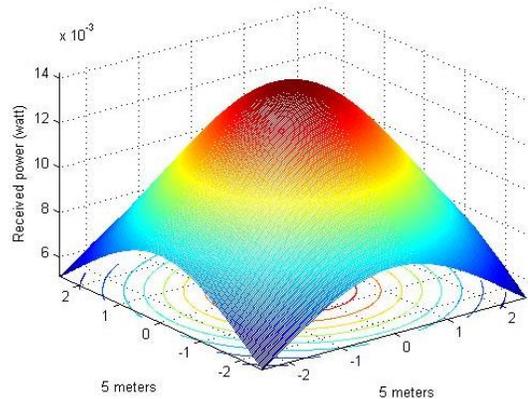
(a) The 3D view of Room size 1 m x 1 m x 1 m with  $d = 1$  meter (Number grid = 50)



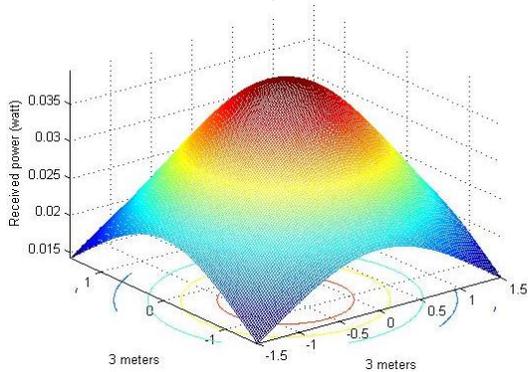
(d) The 2D view of Room size 3 m x 3 m x 3 m with  $d = 3$  meters (Number grid = 50)



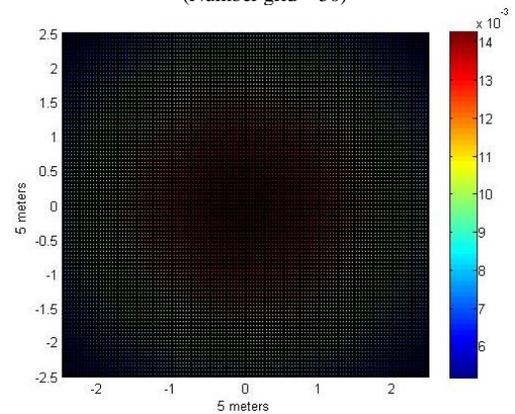
(b) The 2D view of Room size 1 m x 1 m x 1 m with  $d = 1$  meter (Number grid = 50)



(e) The 3D view of Room size 5 m x 5 m x 5 m with  $d = 5$  meters (Number grid = 50)



(c) The 3D view of Room size 3 m x 3 m x 3 m with  $d = 3$  meters (Number grid = 50)



(f) The 3D view of Room size 5 m x 5 m x 5 m with  $d = 5$  meters (Number grid = 50)

Fig. 7. (a) The 3D view of Room size 1 m x 1 m x 1 m with  $d = 1$  meter (Number grid = 50), (b) The 2D view of Room size 1 m x 1 m x 1 m with  $d = 1$  meter (Number grid = 50), (c) The 3D view of Room size 3 m x 3 m x 3 m with  $d = 3$  meters (Number grid = 50), (d) The 2D view of Room size 3 m x 3 m x 3 m with  $d = 3$  meters (Number grid = 50), (e) The 3D view of Room size 5 m x 5 m x 5 m with  $d = 5$  meters (Number grid = 50), (f) The 3D view of Room size 5 m x 5 m x 5 m with  $d = 5$  meters (Number grid = 50).

For that, the simulation parameters are, the distance of the photodiode to the LED of 2 meters,  $\phi_{1/2} = 1^\circ$ , which is perpendicular toward LED, transmitter power of 10 Watt (the minimum value) and FOV =  $45^\circ$ . This experiment is done on photodiode SFH 213 ( $A_r = 1 \text{ mm}^2$ ).

Fig. 8 shows the simulation result with parameters setting  $T_s(\psi) = 1$  and refractive index of lens ( $n$ ) = 0.5, 1, 1.5, 2 and 2.5, where the calculation of  $g(\psi)$  is based on (1). Then, Fig. 9 shows the simulation result with setting  $n = 1$  and  $T_s(\psi) = 1.5, 2, 2.5, 3, 3.5, \text{ and } 4$ .

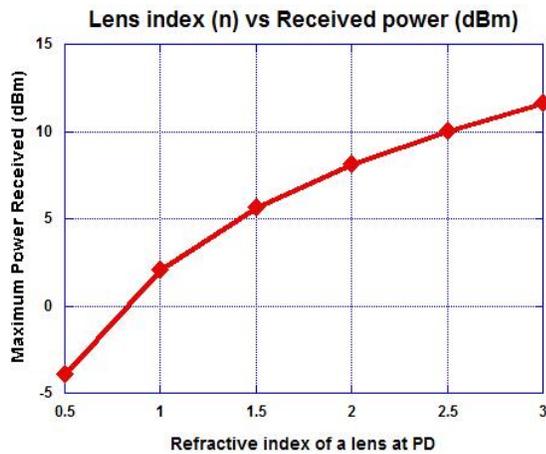


Fig. 8. Lens index vs photodiode's received power.

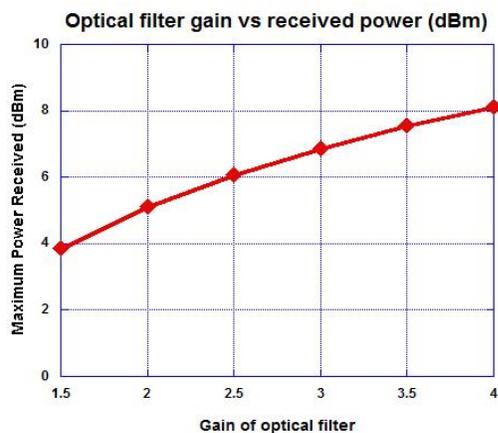


Fig. 9. Optical filter gain vs photodiode's received power.

## V. CONCLUSION

The investigation of received power characteristic on several commercially available photodiode for VLC system with LoS channel has been done clearly. The results of this research show that FOV is not affecting the received power characteristic if  $T_s(\psi)$  and  $n$  factor are ignored. This can be seen from the result of Scenario I which is done by changing the semi-angle of the transmitter. Then on Scenario II, which is done by varying the LED power and Scenario III, by varying the channel distances which denoted by  $d$  (meters). On the next scenario, it is known that  $T_s(\psi)$  and  $n$  factor plays an important role to improve the received power at the photodiode significantly.

Since this article is only investigated the LoS area, the investigation of other channels (i.e. NLoS) is an interesting topic for upcoming issues.

## ACKNOWLEDGEMENT

The authors wish to acknowledge the "PUT-PT Mikroelektronik" and also administrator at the IC design laboratory, Institut Teknologi Bandung for their financial support. We would also like to show our appreciation towards the Institut Teknologi Bandung for providing us with the MATLAB® student license to accomplish this work.

## REFERENCES

- [1] T. Adiono, A. Pradana, R.V.W. Putra, and Y. Aska, "Experimental Evaluation for PWM and OFDM Based Visible Light Communication," *Proc. of the 2<sup>nd</sup> Int. Conf. on Electrical Engineering and Computer Science (ICEECS)*, 2016.
- [2] S.D. Dissanayake and J. Armstrong, "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD Systems," *J. of Lightwave Technology*, Vol. 31(7), pp. 1063- 1072, April 2013.
- [3] I. Stefan and H. Haas, "Analysis of Optimal Placement of LED Arrays for Visible Light Communication," *Proc. of the 7<sup>th</sup> Vehicular Technology Conference (VTC Spring)*, January 2014.
- [4] A-E. Marcu, R-A. Dobre, and M. Vlădescu, "Investigation on Available Bandwidth in Visible-Light Communications," *Proc. of the 22<sup>nd</sup> Int. Symposium for Design and Technology in Electronic Packaging (SIITME)*, pp. 244-247, December 2016.
- [5] T. Adiono, S. Fuada, A.P. Putra, and Y. Aska, "Desain Awal Analog Front-End Optical Transceiver untuk aplikasi Visible Light Communication," *J. Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, Vol.5 (4), pp. 319-327, November 2016. DOI: [10.22146/jnteti.v5i4.280](https://doi.org/10.22146/jnteti.v5i4.280).
- [6] S. Fuada, A.P. Putra, Y. Aska, and T. Adiono, "Short-range Audio Transfer through 3 Watt White LED based on LOS channels," *Unpublished*.
- [7] S. Fuada, A.P. Putra, T. Adiono, and Y. Aska, "Noise Analysis of Trans-impedance Amplifier (TIA) in Variety Op Amp for use in Visible Light Communication System," *unpublished*.
- [8] S. Fuada, T. Adiono, Y. Aska, and A.P. Putra, "Noise Analysis in VLC Optical Link based Discrete OP-AMP Trans-Impedance Amplifier (TIA)," *Unpublished*.
- [9] M.A. Kashani and M. Kavehrad, "On the Performance of Single- and Multi-carrier Modulation Schemes for Indoor Visible Light Communication Systems," *Proc. of Optical Networks and Systems Symposium GLOBECOM*, pp. 2084-2089, 2014.
- [10] F. M. Wu, C. T. Lin, C. C. Wei, C. W. Chen, Z. Y. Chen, H. T. Huang, and S. Chi, "Performance Comparison of OFDM Signal and CAP Signal over High Capacity RGB-LED-Based WDM Visible Light Communication," *IEEE Photonics J.*, Vol. 5, (4), August 2013.
- [11] J. Ding, K. Wang, and Z. Xu, "Impact of Different LED-Spacing in Arrayed LED Transmitter on VLC Channel Modeling," *Proc. of the 6<sup>th</sup> Int. Conf. on Wireless Communications and Signal Processing (WCSP)*, December 2014.
- [12] T. Adiono, A. Pradana, R.V.W. Putra, and S. Fuada, "Analog Filters Design in VLC Analog Front-End Receiver for Reducing Indoor Ambient Light Noise," *Proc. of the IEEE Asia Pacific Conference on Circuit and Systems (APCCAS)*, pp. 581-584, October 2016. DOI: [10.1109/APCCAS.2016.7804058](https://doi.org/10.1109/APCCAS.2016.7804058).
- [13] T. Adiono, Y. Aska, A.A. Purwita, S. Fuada, and A.P. Putra, "Modeling OFDM system with Viterbi Decoder Based Visible Light Communication," *Proc. of Int. Conf. on Electronics, Information, and Communication*, 2017.
- [14] A. Boudkhal, A. Ouzzani, and B. Soudini, "Analysis of Fundamental Photodetection Noises and Evaluation of PIN and APD Photodiodes Performances using an Optical High Debit Transmission Chain Simulated by Optisystem," *Int. J. of Computer Applications*, Vol. 115(18), pp. 21-29, April 2015.
- [15] P. Sharma and H. Sarangal, "Performance Comparison of APD and PIN Photodiodes using Different Modulation and Different Wavelengths," *Int. J. of Signal Processing, Image Processing and Pattern Recognition*. Vol. 9(4), pp.257-264, 2016.
- [16] M.A.A. Ali, "Transmitter Inclination Angle Characteristics for Underwater Optical Wireless Communication in a Variety of APD Detectors," *World Scientific News (WSN)*, Vol. 45(2), 355-372, 2016.
- [17] O. Kharraz and D. Forsyth, "Performance comparisons between PIN and APD photodetectors for use in optical communication systems," *Optik Elsevier*, Vol. 124, pp. 1493 – 1498, 2013.
- [18] Y. Chen, S. Wen Y. Wu, Y. Ren, W. Guan, and Y. Zhou, "Long-range visible light communication system based on LED collimating lens," *Optics Communications*, 377, pp. 83–88, 2016.

[19] S. Fuada and T. Adiono, "Pengaruh Junction Capacitor (C<sub>j</sub>) Photodiode terhadap Noise pada TIA untuk Aplikasi Visible Light Communications (VLC)," *unpublished*.

[20] S. Fuada, A.P. Putra, Y. Aska and T. Adiono, "A First Approach to Design Mobility Function and Noise Filter in VLC System Utilizing Low-cost Analog Circuits," *Int. J. of Recent Contributions from Engineering, Science, and IT (IJES)*, Vol.5(2), pp. 14-30, 2017.

[21] K. Lee, H. Park, and J. R. Barry, "Indoor Channel Characteristics for Visible Light Communications," *IEEE Communications Letters*, pp. 1-3, 2015.

[22] J-H Li, *et al.*, "An integrated PIN-array receiver for visible light communication," *J. Opt.* 17,105805, pp. 1-5, 2015.

[23] T. Adiono, R.V.W. Putra, and S. Fuada, "Noise and Bandwidth Considerations in Designing Op-Amp Based Transimpedance Amplifier for VLC," *unpublished*.

[24] H. Chen, P. Verheyen, P. De Heyn, G. Lepage, J. De Coster, S. Balakrishnan, P. Absil, W. Yao, L. Shen, G. Roelkens, and J. Van Campenhout, "1 V bias 67 GHz bandwidth Si-contacted germanium waveguide P-I-N photodetector for optical links at 56 Gbps and beyond," *Optics Express*, Vol. 24(5), pp. 4622-463, 2016.

[25] W. Zou, Y. Xia, D. Chen, and Y. Zeng, "A High-Speed Lateral PIN Polysilicon Photodiode on Standard Bulk CMOS Process," *Elsevier: Solid-State Electronics*, December 2016. DOI: <http://dx.doi.org/10.1016/j.sse.2016.12.005>.

[26] M. Azadeh, *PIN and APD detectors*. Fiber Optics Engineering-Optical Networks, Springer Science+Business Media, 2009.

[27] T. Komine, S. Haruyama and M. Nakagawa, "A Study of Shadowing on Indoor Visible-Light Wireless Communication Utilizing Plural White LED Lightings," *Wireless Personal Communications* 34, pp. 211-225, 2005.

[28] M.V. Bhalerao, M. Sumathi, and S.S. Sonavane, "Line of sight model for visible light communication using Lambertian radiation pattern of LED," *Int. J. Commun. Syst.*, pp. 1-7, 2016.

[29] Y. Qiu, H-H. Chen, and W-Xi. Meng, "Channel modeling for visible light communications - a survey," *Wirel. Commun. Mob. Comput.*, pp. 1-19, 2016.

[30] T.Adiono and S. Fuada, "Desain dan Implementasi LED Driver Linier untuk Aplikasi Visible Light Communication," *Unpublished*.

[31] C-W. Chow, Y. Liu, C-H. Yeh, C-Y Chen, C-N. Lin, and D-Z. Hsu, "Secure communication zone for white-light LED visible light communication," *Optics Communications*, Vol. 344, pp. 81-85, 2015.

[32] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, "Optical Wireless Communications System and Channle Modelling with MATLAB®," Boca Raton: CRC Press Taylor and Francis Group, 2013.

APPENDIX

In this work, the MATLAB® computations are adopted from [32]. In order to obtain specific point, the authors modify from 3D model to 2D model and then analyze by viewing maximum point of curve characteristic (shown in Fig. 10). A sample of the MATLAB® codes to calculate the LoS channel gain of variety commercial photodiode based 2D view is shown in the program below:

```
theta=45; %semi-angle at half power
m=-log10(2)/log10(cosd(theta)); %lambertian order of emission

P_total=50;%transmitted optical power by individual LED
Adet=1.54e-3; %detector physical area of a photodiode in cm

FOV=55*pi/180;%FOV at receiver
lx=3; ly=3; lz=3; %room dimension in meter

h=2; %the distance between source and receiver plane

XT=0; YT=0; %position of LED
Nx=lx*25; Ny=ly*25;%number of grid receiver plane
```

```
x=-lx/2:lx/Nx:lx/2;y=-ly/2:ly/Ny:ly/2;[XR,YR]=meshgrid(x,y);
%receiver plane grid

D1=sqrt((XR-XT(1,1)).^2+(YR-YT(1,1)).^2+h^2);
%distance vector from source 1

cosphi_A1=h/D1;
%angle vector

H_A1=(m+1)*Adet.*cosphi_A1.^(m+1)./(pi*D1.^2);
%channel DC gain for source 1

P_rec=P_total.*H_A1;
%received power from source 1

P_rec_dBm=10*log10(P_rec);
plot(max(P_rec_dBm));
```

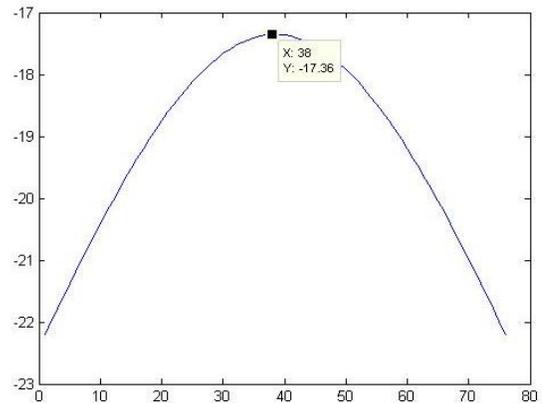


Fig. 10. The 2D view of maximum point of the photodiode's received power at -17.36 dBm.

AUTHORS PROFILES



**Syifaul Fuada** received B. Ed degree on Electrical Engineering Education major from State University of Malang (UM), Indonesia in 2014/2015 and M. Sc degree on Microelectronics engineering major, School of Electrical Engineering and Informatics Institut Teknologi Bandung, Indonesia in 2016/2017. Currently he is working in University

Center of Excellence on Microelectronics Institut Teknologi Bandung. His research interests include Analog circuits design, instrumentation system, circuit simulation, engineering education, multimedia learning, game concepor, and Visible Light Communication.



**Angga Pratama Putra** received his B. Sc. and M.Cs degree in Electrical Engineering from Institut Teknologi Bandung (ITB), Indonesia. His research interests include Embedded System, Software Engineering, VLSI, System-on-Chip, Internet of Things, Digital Signal Processing (DSP) and Visible Light Communication (VLC).



**Trio Adiono** obtained his Ph. D. degree in VLSI Design from Tokyo Institute of Technology, Japan, in 2002. Currently, he is a lecturer at the School of Electrical Engineering and Informatics, a Head of the Microelectronics Center and IC Design Laboratory, ITB. He currently serves as a chair of the IEEE SSCS Indonesia Chapter. His research

interests include VLSI, Signal and Image Processing, Visible Light Communication, Smart Card, Electronics Solution Design and Integration.

# An Efficient Spectral Amplitude Coding (SAC) Technique for Optical CDMA System using Wavelength Division Multiplexing (WDM) Concepts

Hassan Yousif Ahmed

Electrical Engineering Department  
College of Engineering at Wadi Aldawaser  
Prince Sattam bin Abdulaziz University,  
Wadi Aldwasser,  
Riyadh Region,  
Kingdom of Saudi Arabia

Medien Zeghid

<sup>1</sup>Electrical Engineering Department  
College of Engineering at Wadi Aldawaser  
Prince Sattam bin Abdulaziz University, Wadi Aldwasser,  
Riyadh Region, Kingdom of Saudi Arabia  
<sup>2</sup>Electronics and Micro-Electronics Laboratory (E.  $\mu$ . E. L),  
Faculty of Sciences, University of Monastir, Tunisia

**Abstract**—This article introduces an improved method for Optical Code Division Multiple Access system (OCDMA). In this scheme, a hybrid technique is used in which Wavelength Division Multiplexing (WDM) is merged with Spectral Amplitude Coding (SAC) to efficiently diminish Multiple Access Interference (MAI) and alleviate the impact of Phase Induced Intensity Noise (PIIN) appearing in photo-detecting process. The proposed technique SAC-OCDMA/WDM MP (SW-MP) is implemented by using Matrix Partitioning (MP) code family, which is constructed via merging mathematics sequence and algebraic approaches. The key notion is to create the code patterns in SAC domain, then diagonally replicate it in the wavelength domain as blocks which preserves the same code patterns of a given code weight. The SW-MP code family preserves convenient code length property which gives flexibility in transmitter-receiver design. It is reported that the proposed scheme has potential to remove MAI proficiently and improve the system performance significantly.

**Keywords**—Optical Code Division Multiple Access System (OCDMA); Multiple Access Interference (MAI); Spectral Amplitude Coding (SAC); Wavelength Division Multiplexing (WDM); SAC-OCDMA/WDM MP (SW-MP) code; Cross Correlation (CC)

## I. INTRODUCTION

Lately WDM system is measured as a promising technique to expand the optical network capacity without changing the backbone fiber optics. Researchers in both academia and industry sectors have proposed various designs that integrate WDM into access networks by [1]-[2]. OCDMA technique has been deemed as a promising technique for light communication networks [3]. Out of the entire OCDMA techniques, SAC system has gained a lot of consideration due to its ability to eliminate MAI completely via balance detection technique [4]-[6]. On the other hand, PIIN considers as an intrinsic noise that impairs the system performance and it occurs when various light fields are occurrence on a receiver, because of the square-law detection. Several systems have been proposed to be integrated with OCDMA scheme for the sake of MAI

elimination and provide full cardinality in optic network [7]. In this regard temporal/spatial OCDMA network is presented in [7] to improve cross correlation and autocorrelation properties. In [8]-[10], optical pulses are used to mark one chip in time-wavelength domain to improve the MAI cancellation property. Some schemes have been proposed utilizing differential detection to diminish the MAI [11]-[13]. Nevertheless, these schemes are suffering from different problems somehow or other to eliminate an MAI impact on the end system. Eventually tough interference took place which reduces the involvement of high number of users. In SAC system, fiber Bragg grating (FBG) could be utilized as the major part of both transmitter and receiver structures of each use at large number of users, FBG sizes will become unworkable. To overcome size problem, a two dimensional coding techniques might be used but at the cost of extra passive optical components [14]. In this paper, an SW-MP technique is built by merging WDM and SAC system which keeps MAI cancellation characteristic and PIIN mitigation in OCDMA network. The SW-MP code words are described by the code length  $L$ , the number of users  $N$ , the code weights  $W$ , and the cross correlation  $\lambda_c$ . The SW-MP scheme is built with  $\lambda_c \leq 1$  aiming to remove MAI impact. Despite an MAI effect can be removed via balanced detection scheme, PIIN attributed to spontaneous emission from optical source plays a significant role in system degradation too and should be addressed as well [15]. An effective way to reduce the PIIN is by reducing the interference at the optical layer itself, which means the value of  $\lambda_c$  should be kept to the minimum. The remaining parts of this paper are arranged as follows. Section 2 shows the mathematical steps of MP code construction. The mathematical models of MP code and SW-MP code systems are described in Section 3. The system design and description of SW-MP are demonstrated in Section 4. In Section 5, codes comparison and evaluation is discussed. The performance analysis of the SW-MP in the OCDMA network is explained in Section 6. Hypothetical analysis and mathematical findings are drawn in Section 7. Lastly, the summary of the paper is given in Section 8.

## II. MATHEMATICAL MODEL OF SW-MP CODE

### A. Explanations

Let AS = (W, W - 1, W - 2, W - 3, ... ..., 1) represents an arithmetic sequence. The sum of W elements of the arithmetic sequence (AS) can be calculated by “(1)”.

$$S_w = \frac{W}{2}(W + 1) \quad (1)$$

The value of  $S_w$  is the number of columns of matrix partition (MP). Table 1 shows the mapping procedure of MS to MP. Every component in MS will be linked to matching block in MP. The block length is computed as follows:

$$L_g = W - g + 1 \quad (2)$$

Where,  $g = 1, 2, 3, \dots, N$  is the number of groups.

TABLE I. MAPPING ELEMENTS IN AS TO BLOCKS IN MP MATRIX

Block <sub>w</sub>	Block <sub>1</sub>	Block <sub>2</sub>	Block <sub>3</sub>	...	Block <sub>(w-1)</sub>
← W - (W - 1) →	← W →	← W - 1 →	← W - 2 →	-----	← W - (W - 2) →

### B. MP Code Family Construction Steps

- Step1: Construct the AS as follows:

$$AS = (W, W - 1, W - 2, W - 3, W - 4, \dots, 1)$$

The elements of AS indicate the number of blocks (h) of MP matrix.

- Step 2: Compute b value as follows in “(3)”

$$B_1 \leq b \leq B_2 \quad (3)$$

Where,

$$B_1 = 2 + (h - 1)(W - \frac{h-2}{2}) \quad (4)$$

$$B_2 = 1 + h(W - \frac{h-1}{2}) \quad (5)$$

If the calculation of  $B_1$  surpasses the L value; b takes the value 1 (i.e., if  $B_1 > L$  then b,  $B_1$ ,  $B_2$  will be given the value 1).

- Step 3: Compute L's value using “(6)”

$$L = \frac{W \times N}{2} \quad (6)$$

- Step 4: Compute the position of “1s” c at the first row of each block as follows in “(7)”.

$$c = (h, b) \quad (7)$$

- Step 5: Compute the positions of CC “1s” d in every block as follows by “(8)”.

$$d = (h + b - B_1 + 1, b) \quad (8)$$

- Step 6: The  $\frac{W(N-2)}{W-1}$  #0s# values were filled every row.

### C. Code Examples

Apply above steps and as mentioned in [16] the following code patterns were built:

$$MP = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad (9)$$

## III. MATHEMATICAL MODEL OF SW-MP SYSTEM

An SW-MP is a scheme where the entire code created in SAC domain (MP) and replicated diagonally in WDM domain as blocks (SW-MP) [17]. Each block keeps the similar number of users for specified weight of SAC code as displayed in Fig. 1. In this SW-MP system, the code sequences are separated into g blocks, where  $g = 1, 2, 3, \dots$ . Each user is labeled as user #(z,t) and given a code sequence  $C_{z,t}$ ,  $z = 1, 2, 3, \dots, g$  and  $t = 1, 2, 3, \dots, N$ . The code length L is computed using “(10)”.

$$L = g \frac{W \times N}{2} \quad (10)$$

Equations (11) and (12) associated with c and d respectively and determine the positions of “1s” at every row of each block and the positions of CC “1s” in each block, respectively.

$$c = (h + (g - 1)N, b + (g - 1)L) \quad (11)$$

$$d = (h + m - B_1 + (g - 1)N + 1, b + (g - 1)L) \quad (12)$$

$$(3) \begin{bmatrix} (SW-MP)_1 & 0 & 0 & \dots & 0 \\ 0 & (SW-MP)_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & (SW-MP)_N \end{bmatrix}$$

Fig. 1. Matrix representation of SW-MP code.

To explain the SW-MP structure we dealt with code patterns when  $W=2, N=6$  and  $g=2$ .

In Sections 2.2 and 3, we applied Steps 1-2 and 5 and “(12-13)” to compute the place of “1s” in the first row of each block and the locations of CC “1s” in every block respectively (as in Table 1). Therefore, the points coordinates attained for c are (1,2), (1,3), (2,1) using “(7)” and (4,5), (4,6) by applying “(12)”. While for d are (2,2), (3,3), (3,1) utilizing “(8)” and (5,5), (6,6), (6,4) by applying “(12)”.

TABLE II. SW-MP CODE SEQUENCES FOR W=2, G=2 AND N=6

z	T	MP code words $C_{z,t}$					
1	1	0	1	1	0	0	0
1	2	1	1	0	0	0	0
1	3	1	0	1	0	0	0
2	4	0	0	0	0	1	1
2	5	0	0	0	1	1	0
2	6	0	0	0	1	0	1

## IV. SYSTEM DESIGN AND DESCRIPTION

For  $W = 4$  the transmitter section is built in Fig. 2 based on the SW-MP code sequence. Representing digital data in the

form of presence or absence is called on-off shift keying where it used in this work to modulate the data of the targeting user#1 from Table 2, 011000. The modulated data is then guided to an arranged fiber Brag grating. Each pulse of the desirable user is assigned specific wavelengths ( $\lambda_2, \lambda_3, \dots$ ). The center wavelengths of the FBGs depend on the positions of “1s” in the code sequences.

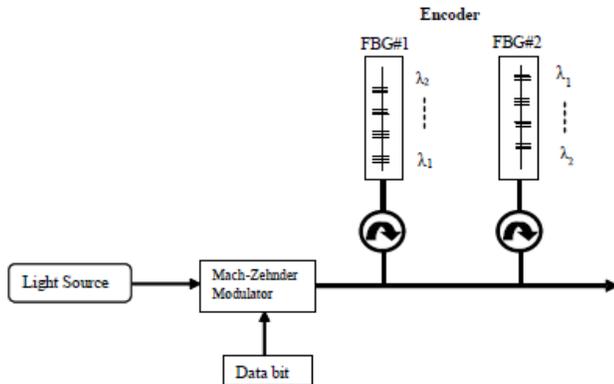


Fig. 2. Implementation of the SW-MP transmitter side.

The basic principle of the work of SW-MP detection procedure in which only pulses of desired users and pulses of overlapping users having the same frequencies in the same group are detected and removed. The configuration of the SW-MP receiver for hybrid SAC OCDMA is shown in Fig. 3. In this figure, the optical pulses are passed to an arranged fiber Brag grating. Each pulse of the desirable user is assigned specific wavelengths ( $\lambda_2, \lambda_3, \dots$ ). The position of the FBGs depends on the pulse value and only the pulse ‘1’ is represented.

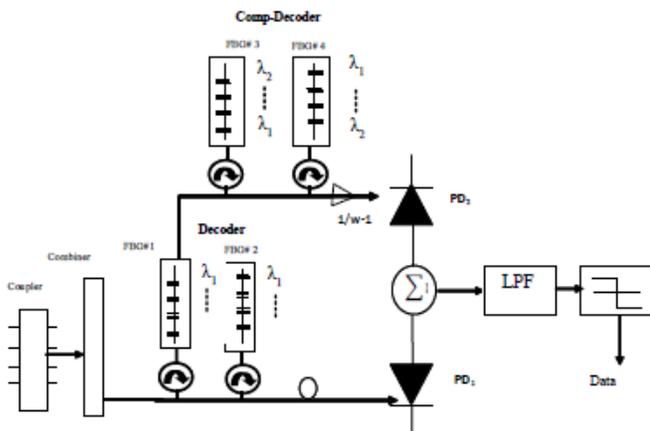


Fig. 3. Implementation of the SW-MP receiver.

In Fig. 3, the incoming pulse is deciphered by the decoder who has similar spectral response to the desired encoder for the data to be processed (Decoder). The intended signal spectrum and overlapping spectra from other interferers are detected as production from the decoder which is  $W$  power units for the desired user accompanied by  $\lambda$  power units for interferers.

The complementary decoder (Comp-Decoder) branch detects the complementary spectrum of the intended user (from Table 2, ( $\lambda_j$ )). Then the received signal is passing over FBG

pieces and the output passed to balanced photo-detectors. From the FBGs, different center wavelengths are placed along a piece of fiber and the wavelength elements of spectral codes are spread out in time. So, second fibers with FBGs in reverse positions are needed in each encoder and decoder in order to compensate the time spreading. To separate the unwanted signal from the wanted signal, a subtraction process is took place to deduct the interference signals from the required signal to yield the wanted signal. Code patterns reside in different groups pass via the decoder without being detected (user #4, user #5, and user #6 from Table 2). The advantage of the SAC/WDM (SW-MP) decoder design over conventional SAC techniques is that only code sequences in the same group are passed to balanced photo-detectors. In addition, the receiver complexity is reduced by using less filters.

## V. SYSTEM PERFORMANCE

To study the system based on Fig. 4 and 5 for SW-MP, let  $C_x(i)$  represents the  $i$ th element of the  $x$ th SW-MP code words; based on XOR subtraction scheme the code properties is formed as follows [18]-[19]:

$$\sum_{i=1}^L C_x(i)C_y(i) = \begin{cases} W, & x = y \\ 1, & x \neq y \end{cases} \begin{cases} \text{In the same group (g=1)} \\ \text{Not in the same group (g} \geq 2) \end{cases} \quad (13)$$

$$\sum_{i=1}^L (C_x(i) \oplus C_y(i))C_x(i) = \begin{cases} 0, & x = y \\ W - 1, & x \neq y \end{cases} \begin{cases} \text{In the same group (g=1)} \\ \text{Not in the same group (g} \geq 2) \end{cases} \quad (14)$$

The stipulation of  $x$  and  $y$  exist in the same group (i.e.,  $g = 1$ ) means that the two code sequences maybe located in SW-MP<sub>(1)</sub> or SW-MP<sub>(2)</sub> or SW-MP<sub>(m)</sub> as shown in Fig. 1. The stipulation of  $x$  and  $y$  not reside in the same group (i.e.,  $g \geq 2$ ) means that one user maybe located in SW-MP<sub>(1)</sub> and the other user is resided in SW-MP<sub>(2)</sub> or SW-MP<sub>(m)</sub>. Hence, the XOR process of  $(C_x(i) \oplus C_y(i))C_x(i)$  is valid for  $x \neq y$ . Still, the CC of  $(C_x(i) \oplus C_y(i))$  is only valid for  $x \neq y$  in “(14)” while from “(14)”, the CC of  $(C_x(i).C_y(i))$  is  $W$  when  $x = y$ . Consequently, an MAI impact is removed as the CC  $\sum_{i=1}^L (C_x(i) \oplus C_y(i))C_x(i)$  can be deducted from  $\sum_{i=1}^L C_x(i)C_y(i)$  when  $x \neq y$ . Therefore, the decoder that calculates “(15)” declines the MAI arriving from intrusive users and gets the desired data.

Hence,

$$\frac{\sum_{i=1}^L C_x(i)C_y(i) - \sum_{i=1}^L (C_x(i) \oplus C_y(i))C_x(i)}{W-1} = \begin{cases} W, & x = y \\ 0, & x \neq y \end{cases} \quad (15)$$

When  $x \neq y$  the lower branch in “(16)” equals zero which means an XOR subtraction technique is able to remove an MAI impact smoothly. To study the system performance we compute the coherence time of a thermal source ( $\tau_c$ ) as follows [20]:

$$\tau_c = \frac{\int_0^\infty G^2(v)dv}{\left[ \int_0^\infty G(v)dv \right]^2} \quad (16)$$

Where,  $G(v)$  is the single sideband power spectral density (PSD) of the optic source. The variance of photocurrent caused by the recognition of un-polarized thermal source, which produces by spontaneous emission and given as follows [4]-[5]:

$$\langle i^2 \rangle = \langle I_{shot}^2 \rangle + \langle I_{PIIN}^2 \rangle + \langle I_{thermal}^2 \rangle \quad (17)$$

Where,  $I_{shot}^2$  denotes shot noise,  $I_{PIIN}^2$  symbolizes intensity noise and  $I_{thermal}^2$  represents the thermal noise. Hence, "(18)" will be written as follows:

$$\langle i^2 \rangle = 2eI_B + I^2 B \tau_c + 4K_B T_n B R_L \quad (18)$$

Where,

- electronic charge symbolizes by  $e$
- average photocurrent symbolizes by  $I$
- equivalent electrical bandwidth of the receiver denotes  $B$ ;
- Boltzmann's constant denotes by  $K_B$
- Absolute receiver noise temperature denotes by  $T_n$
- load resistor of receiver denotes by  $R_L$ .

$$r(v) = \frac{P_{sr}}{\Delta v} \sum_{n=1}^N d_n \sum_{i=1}^L c_n(i) rec(i) \quad (19)$$

Where,  $P_{sr}$  represents the active power of a broad-band source at the receiver and  $d_n$  is the data bit of the  $n$ th user that is "1" or "0". The  $rec(i)$  function in "(19)" is written as follows as in [16]-[17]:

$$rec(i) = u \left[ v - v_o - \frac{\Delta v}{2L} (-L + 2i - 2) \right] - u \left[ v - v_o - \frac{\Delta v}{2L} (-L + 2i) \right] \quad (20)$$

Where,  $v_o$  is the central optical frequency and  $\Delta v$  is the optical source bandwidth in  $\text{Hz}$ . The unit step function  $u[v]$  is written as:

$$u[v] = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases} \quad (21)$$

The overall power occurrence at the Photodiode1 and Photodiode 2 as shown in Fig. 3 of the  $g$ th receiver during one bit period is formed as follows:

$$\begin{aligned} \int_0^{\infty} G_1(v) dv &= \int_0^{\infty} \frac{P_{sr}}{\Delta v} \sum_{f=1}^N d_f \sum_{i=1}^L c_f(i) C_g(i) \left( u \left[ v - v_o - \frac{\Delta v}{2L} (-L + 2i - 2) \right] - u \left[ v - v_o - \frac{\Delta v}{2L} (-L + 2i) \right] \right) dv \\ &= \frac{P_{sr}}{\Delta v} \frac{\Delta v}{L} \sum_{f=1}^N d_f \sum_{i=1}^L c_f(i) C_g(i) \\ &= \frac{P_{sr} W}{L} d_g + \frac{P_{sr}}{L} \sum_{f=1, f \neq g}^N d_f \\ \int_0^{\infty} G_2(v) dv &= \int_0^{\infty} \frac{P_{sr}}{\Delta v} \sum_{f=1}^N d_f \sum_{i=1}^L \frac{c_f(i) (C_f(i) \oplus C_g(i))}{W - 1} \left( u \left[ v - v_o - \frac{\Delta v}{2L} (-L + 2i - 2) \right] - u \left[ v - v_o - \frac{\Delta v}{2L} (-L + 2i) \right] \right) dv \\ &= \frac{P_{sr}}{\Delta v} \frac{\Delta v}{L} \sum_{f=1}^N d_f \sum_{i=1}^L \frac{c_f(i) (C_f(i) \oplus C_g(i))}{W - 1} \\ &= \frac{P_{sr}}{L} \sum_{f=1, f \neq g}^N d_f \end{aligned} \quad (22)$$

The current  $I$  of preferred user is computed by taking the difference of two photodiodes as follows:

$$I = I_1 - I_2 \quad (24)$$

The currents at Photodiode1 and Photodiode 2 are denoted by  $I_1$  and  $I_2$ , respectively.

$$\begin{aligned} I &= \Re \int_0^{\infty} G_1(v) dv - \Re \int_0^{\infty} G_2(v) dv \\ &= \Re \left( \frac{P_{sr} W}{L} d_g + \frac{P_{sr}}{L} \sum_{f=1, f \neq g}^N d_f - \frac{P_{sr}}{L} \sum_{f=1, f \neq g}^N d_f \right) \\ &= \Re \left( \frac{P_{sr} W}{L} d_g \right) \end{aligned} \quad (25)$$

The photo-detectors responsivity is  $\Re$  and represented by

$$\Re = \frac{\eta e}{h \nu_c} \quad (26)$$

Here,  $\eta$  is the quantum efficiency and  $h$  is the Planck's constant. The shot noise power can be written as:

$$\begin{aligned} \langle I_{shot}^2 \rangle &= 2eB \Re \left[ \int_0^{\infty} G_1(v) dv + \int_0^{\infty} G_2(v) dv \right] \\ &= 2eB \Re \left( \frac{P_{sr}}{L} \sum_{f=1, f \neq g}^N d_f + \frac{P_{sr}}{L} \sum_{f=1, f \neq g}^N d_f \right) \\ &= 2eB \Re \frac{P_{sr}}{L} \left( W d_g + 2 \sum_{f=1, f \neq g}^N d_f \right) \\ &= 2eB \Re \frac{P_{sr}}{L} (W + 2(N - 1)) \end{aligned} \quad (27)$$

$$\langle I_{shot}^2 \rangle = 2eB \Re \frac{P_{sr}}{L} [2(N - 1) + W]$$

Once all users conveying bit "1" and by approximating the summation from "(21)" via applying the average value as  $\sum_{f=1}^N C_f \cong \frac{NW}{L}$ . Based on the properties of SW-MP code, the PIIN noise power is given as [4]-[5]:

$$\begin{aligned} \langle I_{PIIN}^2 \rangle &= B I_1^2 \tau_{c1} + B I_2^2 \tau_{c2} \\ &= B \Re^2 \left[ \int_0^{\infty} G_1^2(v) dv + \int_0^{\infty} G_2^2(v) dv \right] \\ &= B \Re^2 \frac{P_{sr}^2}{\Delta v L} \sum_{i=1}^L \left\{ C_g(i) \left[ \sum_{f=1}^L d_f C_f(i) \right] \left[ \sum_{m=1}^N d_m C_m(i) \right] \right\} \\ &\quad + \frac{B \Re^2}{P^2} \frac{P_{sr}^2}{\Delta v L} \sum_{i=1}^L \left\{ (C_f(i) \oplus C_g(i)) \left[ \sum_{f=1}^L d_f C_f(i) \right] \left[ \sum_{m=1}^N d_m C_m(i) \right] \right\} \\ &\cong B \Re^2 \frac{P_{sr}^2}{\Delta v L} \sum_{i=1}^L \left\{ C_g(i) \frac{NW}{L} \left( \sum_{f=1}^L C_f(i) \right) \right\} \\ &\quad + \frac{B \Re^2}{P^2} \frac{P_{sr}^2}{\Delta v L} \sum_{i=1}^L \left\{ (C_f(i) \oplus C_g(i)) \frac{NW}{L} \left( \sum_{f=1}^L C_f(i) \right) \right\} \\ &\cong B \Re^2 \frac{P_{sr}^2}{\Delta v L} \frac{NW}{L} \sum_{f=1}^L \left( \sum_{i=1}^L C_f(i) \cdot C_g(i) \right) \\ &\quad + \frac{B \Re^2}{\Delta v L} \frac{P_{sr}^2}{L} \frac{NW}{L} \sum_{f=1}^L \left( \sum_{i=1}^L C_f(i) \cdot (C_f(i) \oplus C_g(i)) \right) \end{aligned}$$

$$\begin{aligned} \langle I_{min}^2 \rangle &\cong B \mathfrak{R}^2 \frac{P_{sr}^2}{\Delta v L} \frac{NW}{L} (W + N - 1) + \frac{B \mathfrak{R}^2}{\Delta v L} \frac{P_{sr}^2}{L} \frac{NW}{L} (N - 1) \\ &= \frac{B \mathfrak{R}^2 P_{sr}^2 NW}{\Delta v L^2} (W + 1 + 2(N - 1)) \end{aligned} \quad (28)$$

It should be point out that the probability of transmitting bit ‘1’ in any time of each user is 0.5, then “(27)” and “(28)” become respectively [4]:

$$\langle I_{shot}^2 \rangle = eB \mathfrak{R} \frac{P_{sr}}{L} [2(N - 1) + W] \quad (29)$$

and

$$\langle I_{min}^2 \rangle = \frac{B \mathfrak{R}^2 P_{sr}^2 NW}{2\Delta v L^2} \left( W + 1 + \frac{2(N - 1)}{g} \right) \quad (30)$$

To determine the overlapping from the other users hitting on the desired user we have studied two cases depending on the values of g. If g = 1 means the two users are reside in the similar group while for the condition g ≥ 2 means the two users are reside in another groups. Thus, “(28)” is simplified furthermore as:

The thermal noise is given as [4]-[5]:

$$\langle I_{thermal}^2 \rangle = \frac{4K_b T_n B}{R_L} \quad (31)$$

The SNR of the SW-MP scheme is calculated as:

$$SNR = \frac{I^2}{\langle i^2 \rangle} = \frac{(I_2 - I_1)^2}{\langle I_{shot}^2 \rangle + \langle I_{min}^2 \rangle + \langle I_{thermal}^2 \rangle} \quad (32)$$

Thus “(32)” based on “(25)”, “(29)”, “(30)” and “(31)” can be written as:

$$SNR = \frac{\mathfrak{R}^2 P_{sr}^2 (W - 1)^2 / L^2}{(P_{sr} e B \mathfrak{R} / L) [2N + W - 2] + \left( \frac{B \mathfrak{R}^2 P_{sr}^2 NW}{2\Delta v L^2} \right) \left( W + 1 + \frac{2(N - 1)}{g} \right) + 4K_b T_n B / R_L} \quad (33)$$

Using Gaussian approximation as in [4]-[5], the bit error rate (BER) is computed as follows:

$$BER = \frac{1}{2} \operatorname{erfc} \left( \sqrt{\frac{SNR}{8}} \right) \quad (34)$$

## VI. COMPARISON AND EVALUATION

For evaluation objective, the characteristic of the SW-MP and DEU codes are tabulated in Table 3 [18]-[19]. Table 3 shows that SW-MP and DEU codes exist for positive integer W, free cardinality, and ideal CC. In terms of code length and to support 8 users, the code lengths required by SW-MP (W=3) and DEU (W=3) are 12 and 17, respectively. To conclude SW-MP has short code length compared to DEU for the same parameters; long code length is not practical to be implemented in terms of hardware as the code is susceptible to either very extensive band source or narrow filter bandwidths are necessary.

TABLE. III. PROPERTIES OF SW-MP AND DEU CODES [18]-[19]

Property	DEU-technique	SW-MP-technique	Remarks
$\lambda$	$\leq 1$	$\leq 1$	<ul style="list-style-type: none"> <li>•The maximum CC equals one between any two adjacent DEU codes</li> <li>•The maximum cross correlation is zero when <math>g \geq 2</math> for SW-MP codes</li> </ul>
Existence	Any integer number	Any integer number	•More flexibility in code weight selection for SW-MP and DEU codes
Size	Free	Free	•Free cardinality for SW-MP and DEU codes
Code length	$L = N(W - 1) + 1$	$L = \frac{(W \times N)}{2}$	•SW-MP code has short code length compared to DEU code
Number of variables and parameters	W	W, g	•SW-MP has two parameters in construction while DEU has one parameter
Matrix Form	Yes	Yes	•Same steps of construction

## VII. THEORETIC AND SIMULATION FINDINGS

In this section the process of SAC domain combined with WDM domain is investigated by considering different types of noises such as PIIN, thermal and shot noises using the key parameters listed in Table 4.

TABLE. IV. ELEMENTS OF SNR AND BER ASSOCIATED WITH CORRESPONDING VALUES

Symbol	Translation of symbol	Symbol's value and representation
$\eta$	Quantum efficiency of photodiode	0.6
$V_c$	Line-width of the thermal source	3.75THz
$\lambda_0$	Transmission Window	1550 nm
$B$	Electrical bandwidth	80 MHz
$R_b$	Data bit rate	155 Mb/s
$T_n$	Absolute receiver noise temperature	300 K
$R_L$	Receiver load resistor	1030 $\Omega$
$e$	The electron charge	$1.6 \times 10^{-19}$ Coulomb's
$K_B$	Boltzmann's constant is	$1.38054 \times 10^{-23}$
$P_{sr}$	Efficient power of a broad-band source	-10 dBm
$N$	Number of simultaneous users	Vary
$W$	Code weight	Any integer number
$g$	Number of groups	Any integer number
$L$	Code's length	$L = \frac{(W \times N)}{2}$

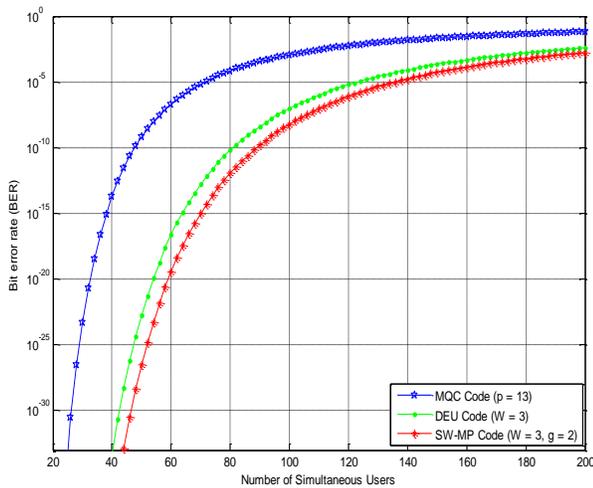


Fig. 4. Bit error rate against simultaneous users at 155Mbps/s.

In Fig. 4, SW-MP is compared as a function of the number of concurrent users versus BER with conventional SAC coding technique MQC and DEU. The comparison has been carried out in a free space setting with different values of parameters when  $P_{sr} = -10\text{dBm}$  at 155Mbps/s for MQC ( $p=16$ ), DEU ( $W=3$ ), and SW-MP ( $W=3, g=2$ ). As expected, SW-MP code is having good result in terms of performance with regard to DEU even for the same code's weight, which is 3 and this is due to the elimination of the interference from different users when the value of  $g \geq 2$ . As seen from the result, for lower weight ( $W=3$ ) the quality of the received signal is satisfactory and the  $\text{BER} = 10^{-9}$  is attained for almost  $\approx 80$  and  $\approx 90$  users for DEU and SW-MP, respectively.

SW-MP outperforms MQC even for higher code's weight, which is 13. This outperformance due to higher SNR in SW-MP as compared with MQC. However, MQC code is utilized for an ideal cross-correlation ( $\lambda=1$ ) where each system employing MQC still suffers from MAI effect, thus preventive the system performance for more improvement. Hence, this boosts the signal's impairment eventually system performance degradation.

In Fig. 5, performance of the system with regards to the effective power  $P_{sr}$  for 40 users at data rate of 155Mb/s for considering PIIN, shot and thermal noises for MFH ( $q=16$ ), MQC ( $p=13$ ), DEU ( $W=3$ ) and SW-MP ( $W=3, g=2$ ) codes is evaluated. The figure demonstrates that the  $P_{sr}$  of the BER of  $10^{-9}$  is achieved with  $P_{sr} \approx -28\text{ dBm}$  for the SW-MP code while the same BER is obtained as  $P_{sr} \approx -18\text{ dBm}$ ,  $P_{sr} \approx -17\text{ dBm}$  and  $P_{sr} \approx -23\text{dBm}$  for the MFH, MQC and DEU codes, respectively. Due to its good property the SW-MP code shows better performance where an MAI impact is minimized when  $g \geq 2$  whereas for MFH and MQC codes are "1", respectively as the number of concurrent users increases. In particular, for DEU code the maximum CC is "1" between any two neighboring users that minimizes the impact of MAI and this sees in better performance compared to MQC and MFH codes.

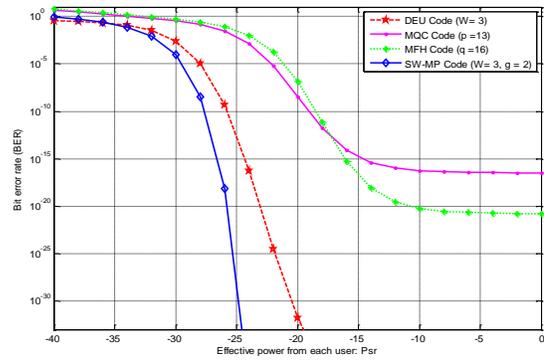


Fig. 5. BER plotted against  $P_{sr}$  for 40 users at the data rate 155Mbps/s.

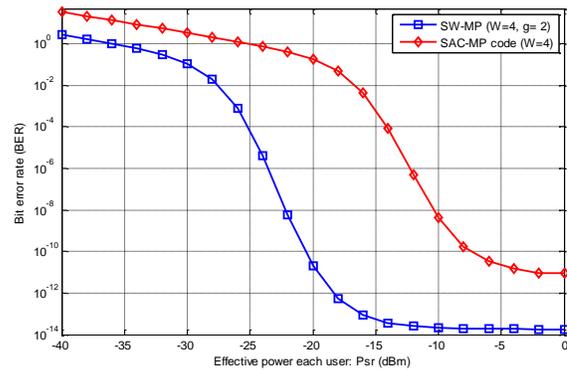


Fig. 6. BER drawn against  $P_{sr}$  for 15 users at the data rate 155Mbps/s.

Fig. 6 displays the BER drawn versus  $P_{sr}$  for 40 users at a data rate of 155Mbps/s considering the PIIN, thermal and shot noises for SW-MP (hybrid system) and SAC-MP (conventional system) codes. SW-MP is created with the parameters  $W=4$  and  $g=2$  (two groups); SAC-MP code is selected with the weight  $W=4$ . The acceptable error for reliable transmission ( $\text{BER} < 10^{-9}$ ) is achieved for the SW-MP code with  $\approx -22\text{ dBm}$  whereas the same BER is achieved with  $\approx P_{sr} > -8\text{ dBm}$  for SAC-MP code for 40 users. This is because the effect of MAI is diminished when  $g \geq 2$  for SW-MP while for SAC-MP is fixed as the number of concurrent users increases. The figure proves that SW-MP codes outperform SAC-MP codes by the magnitude of almost three times.

## VIII. CONCLUSION

The study in this paper has provided a promising strategy to constructing a code family for optical communication. The main aspects of this code family have been studied extensively. The theory and calculation findings of the system were compared with the codes mentioned in the literature review. The SW-MP technique is a SAC code reiterated diagonally in wavelength domain. The in-phase CC has maximum value of "1" in the similar group and zero with codes in different groups. SW-MP code has several features like the freedom of picking the number of users (free cardinality) than other codes, workable code length and easy to carry out through FBGs. It's

reported that the new code family was able to mitigate PIIN noise efficiently and enhance the system performance noticeably. It is mentioned that when the system is highly populated with concurrent users, the SW-MP surpass almost three times the traditional ones and the BER is reduced when the value of  $g$  increments.

Furthermore, top of its excellence performing the SW-MP codes need less complexity in terms of hardware for the transmitter-receiver structure.

#### ACKNOWLEDGMENT

This project was supported by the Deanship of Scientific Research at Salman bin Abdulaziz University under the research project # 2014/1/877.

#### REFERENCES

- [1] N.Frigo, P.Iannone, and K.Reichmann, "Spectral slicing in WDM passive optical networks for local access," Proc. Europ. Conf. Optical Commun. Tech. Dig.1, pp.119–120, 1998.
- [2] N.Froberg, S.Henion, H.Rao, B.Hazzard, S.Parikh, B.Romkey, M.Kuznetsov, "The NGI ONRAMP test bed: Reconfigurable WDM technology for next generation regional access networks," Journal of Lightwave Technology., vol. 18(12), pp.1697–1708, 2000.
- [3] J.A.Salehi, C. A.; Brackett, "Code division multiple access techniques in optical fiber network—Part II: System performance analysis," IEEE Transaction on Comm., vol. 37, pp.834–842, 1989.
- [4] Z.We, H. M. Shalaby, H.Ghafouri-Shiraz, "Modified Quadratic Congruence codes for Fiber Bragg-Grating-Based SAC-OCDMA," Journal of Lightwave Technology., vol.50, pp.1209-1212, 2002.
- [5] Z.H.We, H. Ghafouri-Shiraz, "Code for spectral amplitude-coding optical CDMA systems," Journal of Lightwave Technology, vol.20, pp.1284-1291, 2002.
- [6] T.H.Abd, S.A.Aljunid, H. A, Fadhil, R.B.Ahmad, M.N.Junita, "Enhancement of performance of a hybrid SAC-OCDMA system using dynamic cyclic shift code," Ukrainian Journal of Physical Optics., vol.13, pp.12-27, 2012.
- [7] E.Park, A.J.Mendez, E.M.Garnire, "Temporal/spatial optical CDMA networks-design, demonstration, and comparison with temporal networks," IEEE Photonics Technol. Lett.,vol. 4, pp.1160-1162, 1992.
- [8] R.M.H.Yim.; L.R.Chen, J. Bajcsy, "Design and performance of 2-D codes for wavelength-time optical CDMA," IEEE Photonics Technol. Lett., vol.14, pp.714-716, 2002.
- [9] A. J.Mendez; R. M.Gagliardi, V. J.Hernandez,. C. V.Bennett, W. J.Lennon, "Design and performance analysis of wavelength/time (W/T) matrix codes for optical CDMA," IEEE J. Lightwave Technol.,vol. 21, pp.2524-2533, 2003.
- [10] L.Song-Ming, Jen-Fa, H. Chao-Chin Yang, "Optical CDMA network codecs with merged-M-coded wavelength-hopping and prime-coded time-spreading," Opt. Fiber Technol.,vol. 13, pp.117-128, 2007.
- [11] S. P.Wan; Y.Hu , "Two-dimensional optical CDMA differential system with prime/OOC codes," IEEE Photonics Technol. Lett., vol.13, pp.1373-1375, 2001.
- [12] R.M.H.Yim, J. Bajcsy, L.R. Chen, " A new family of 2-D wavelength-time codes for optical CDMA with differential detection," IEEE Photonics Technol. Lett.,vol. 15,pp.165-167, 2003.
- [13] H.Heo; S.Min, Y.H.Won; Y.Yeon; B. K Kim, "A new family of 2-D wavelength-time spreading code for optical code-division multiple-access system with balanced detection," IEEE Photonics Technol. Lett., vol.16, pp.2189-2191, 2004.
- [14] C.C.Yang, J.F.Huang, "Two-Dimensional M-matrices coding in spatial/frequency optical CDMA networks," IEEE Photon. Technol. Lett., vol.15, pp.168–170, 2003.
- [15] E. D. Smith, R. J.Blaikie, D.P.Taylor, "Performance enhancement of spectral-amplitude-coding optical CDMA using pulse-position modulation," IEEE Trans. Commun., vol.46, pp.1176–1185, 1998.
- [16] H. Y.Ahmed, Z. M Gharsseldien, K.S.Nisar, and S. A. Aljunid "An Inclusive Comparison in LAN Environment between Conventional and Hybrid Methods for Spectral Amplitude Coding Optical CDMA Systems,IJACSA" vol.7, pp.335–344, 2016.
- [17] C. Yang, "Hybrid wavelength-division multiplexing/spectral-amplitude-coding optical CDMA system," IEEE Photon. Technol. Lett., vol.17, pp.1343–1345, 2005.
- [18] H. Y.Ahmed, K.S.Nisar, "Diagonal Eigenvalue Unity (DEU) code for spectral amplitude coding-optical code division multiple access," Optical Fiber Technology., vol.19, pp.335–347, 2013.
- [19] H.Y.Ahmed, K.S.Nisar, "Reduction of the fiber dispersion effects on MAI for long span high-speed OCDMA networks using Diagonal Eigenvalue Unity (DEU) code," Optik, vol.124, pp.5765– 5773, 2013.
- [20] J.W. Goodman, statistical optics, Wiley, New York, 1985.

# Efficient Key Agreement and Nodes Authentication Scheme for Body Sensor Networks

Jawaid Iqbal

Department of Information Technology  
Hazara University  
Mansehra, Pakistan

Noor ul Amin

Department of Information Technology  
Hazara University  
Mansehra, Pakistan

Arif Iqbal Umar

Department of Information Technology  
Hazara University  
Mansehra, Pakistan

Nizamud Din

Department of Computer Science  
IQRA National University  
Peshawar, Pakistan

**Abstract**—Technological evolvement of Wireless Sensor Networks (WSNs) gave birth to an attractive research area for health monitoring called Body Sensor Network (BSN). In BSN tiny sensor nodes sense physiological data of patients under medical health care and transmit this data to Base Station (BS) and then forward to Medical Server (MS). BSN is exposed to security threats due to vulnerable wireless channel. Protection of human physiological data against adversaries is a major addressable issue while keeping constrained resources of BSN under consideration. Our proposed scheme consists of three stages. In first stage deployment of initial secret key by the ward Medical Officer (MO), in second stage secure key exchange and node authentication, in third stage secure data communication are performed. We have compared our proposed scheme with three existing schemes. Our scheme is efficient in computation cost, communication overhead and storage as compared to existing schemes while providing enough security against the adversaries.

**Keywords**—Body sensor network; hash function; node authentication; key agreement; session key

## I. INTRODUCTION

The WSN applications in various fields like natural disasters, habitat monitoring, battle field, and other emergency services got the attention of the researchers [1], and WSN evolved to BSN for medical applications. In 1996 T.G. Zimmerman proposed the idea of Wireless Body Sensor Networks (WBSNs) for the first time. These networks were initially called Wireless Personal Area Network (WPANs). A typical sensor node's hardware consist on processor and memory, wireless communication stack, analog to digital converter and sensing [2]. BSN network comprised of low power, low processing, small size, light weight body sensors deployed on patient body which constantly monitor Electroencephalogram (EEG), respiratory rate, heart rate, Blood Pressure (BP) through by sensing then forward the real time sensed patient data to BS outside the body for onward transmitting to MS. After receiving patient data by MS, the ward physician gives feedback for the patient health care [3]. The bandwidth for transmission in BSN is 10Kbps to

10Mbps [2]. BSN being challenging area of research have a number of research directions such as its security, energy, memory and data management. We accept the challenge of secure and authentic transmission of patient physiological data to the legal user (MO) of the network while keeping adversary attacks and overall efficiency of BSN. For secure communication between communicating parties it is essential to confidentially share secret keys. Our proposed scheme addresses the efficient key management and authentication to encounter the possible attacks on the system and reduce the human life risk. The three stages scheme is proposed where in first stage the ward MO deploy initial master secret keys  $M_{sk}$  in BSN equipments and stores the IDs of all sensor nodes in BS for further establishing secure link, in second stage legal nodes are authenticated and secure keys are established for the transmission of the next stage data, in third stage secure transmission of physiological data is performed. We compare the efficiency of our scheme with existing schemes and obtained results show that our proposed scheme is efficient in communication overhead, computation cost and storage requirement while provide protection against the attackers.

The rest of the paper is organized in sections. In Section II "Related Work" the background and related security schemes are critically discussed. In Section III "Network Model", Section IV "Radio Model", Section V "Attack Model", Section VI "Proposed Protocol", Section VII "Security Analysis", Section VIII "Performance Analysis" and Section IX "Conclusion" are elaborated.

## II. RELATED WORK

In scheme [4], [5] the proposed protocols sensor association and key management have considered while in [4] public key based authentication is used for the secure association of sensor nodes with the controller, however sensor with patient authentication is not considered that leads to security lapse and an illegal node may join the network and pick the patient data. Association between sensor nodes and controller is tedious and high in computation cost. In [5] group keys establishment and authentication is performed by group device pairing where for obtaining group keys each sensor has

to perform  $n + 3$  times Modular Exponentiations (M-Exp) operations and  $n$  represents total of sensors in BSN which over burden the short resources sensor nodes and the communication overhead and computation cost of this scheme is quite high. In scheme [6], node to node authentication and key agreement is performed and Diffie-Hellman (DH) protocol is initially applied for the establishment of the secret key which is vulnerable to man in the middle attack. In this scheme, membership broad casting to all sensor nodes leads to high communication and memory overhead. The problem with [7] is that each received packet on each node is decrypted and hash function is applied which clearly increase computational cost and can't suitable for BSN environment. Scheme [1] is suitable for WSN but bulky for BSN as RSA public cryptosystem is used. In scheme [8], a hybrid approach is used where RSA is used for key agreement and symmetric cipher for session data transmission. As RSA is a public cryptosystem using 1024 bit key which is infeasible for the resource constrained sensor nodes and has high computation cost similarly it lacks node authentication. In [9] ECC is used for Key agreement instead of RSA which somehow reduced computation cost and memory requirement as ECC use 160 bits key but still costly due to hybrid (asymmetric and symmetric) approach and no mechanism for node authentication. Scheme [10] uses biometric technique for key agreement using electrocardiogram. The generated session key is used for secure transmission of patient data between sensors and base station. The keys generated through electrocardiogram are long and random. Identical ECG signals generate non linkable keys. Although this scheme provides security but obtaining two signals from accurate similar random biological signals is hard [11]. In scheme [12] asymmetric cryptosystem is used for key establishment and rekeying by utilizing DHECC and RSA. Specific routing algorithm is used for efficiency purposes. However, this scheme has unavoidable problems of high computation cost and storage requirement due to PKC, RSA and DHECC and inconvenient for tiny sensor nodes of BSN. Scheme [13] uses smart card and password based user authentication for patients' health care with two stages of registration and login and authentication. Before to access BSN each user has to be registered with the gateway first and then the gateway issue smart card to the system user which is used for accessing patient data. Smart card contains login information to network. After authentication a session for information communication is generated between communicating parties. This technique suffers from security flaws. In scheme [14] a preloaded secret key is shared amongst all node of the network. Then another secret session key for a specific session is generated by the cryptographic protocol. This scheme can be used for a large dynamic nature network. In proposed scheme [15] asymmetric mechanism is applied for sharing secret key amongst nodes. Then that key is used for the session data transmission securely using symmetric cryptosystem. This seems to be hybrid scheme where public key infrastructure is used for key establishment and symmetric for secure communication. This scheme is expensive in computational cost while using asymmetric technique for key establishment. Scheme [16] introduced a WBANs security suite; IAMKeys technique for WBAN key management and KEMESIS for inter-sensors

transmission, random keys generation and ensuring security by eliminating exchange of keys between body sensors. Inter sensor communication over burden network overhead. To avoid inter sensors communication over all network overhead can be reduced. In scheme [17] AES based encryption which is supported by CC2420 where all nodes involved in communication receive share secret key through by a specific server. MAC, CCM and CBC are used for encryption and authentication. This is a platform dependency scheme. In scheme [18] generation of 128 bits key is performed using IPI and time difference is calculated by the peaks of the ECG/PPG. Hamming distance error correction scheme is used. The limitation of the scheme is that by a minute difference in calculating IPI at sensor error correction code should be applied for balancing keys. Calculating of IPI values require enough time which slows down the BSN. In scheme [19] SCK and ECC is used for authentication using pair of keys. Sensor nodes are loaded with confidential data through KDC for this identity of each sensor. Various parameters of EEC are used for association of every patient in BSN. Association patients and sensors is very difficult so the scheme is impractical for large hospitals with hundreds of patients. In scheme [20] pair of keys is established using ECC amongst sensors and BS. Patients of BSN are authenticated using biometric device attached to every sensor node. Attachment of biometric device leads to more energy consumption and memory requirement of the sensor. In scheme [21], a three tier architecture is presented for health care application i.e. patients authentication through by biometric, ECC for key agreement and symmetric encryption for confidential session data transmission with integrity. Each sensor is connected with a small scanner for finger print for ID of patients. This is a secure scheme but expensive with respect to computation cost and energy overhead.

### III. NETWORK MODEL

The network model comprises of low power sensors, base station and medical server. Low power biosensor nodes are deployed on patient body for sensing vital signs data. This data is forwarded to a device called BS or Access Point (AP) which acts as a controller. All BSN sensors access the base station directly to avoid inter sensor communication and reduce the BSN traffic. Base station is resourceful equipment with no limits of storage, processing and energy. BS forward health status received from sensing sensors to medical server. MS stores health status record which is received by the ward physician for speedy treatment. For interoperability Zigbee/802.15.6 standards are preferred to be used and all nodes are accessible at maximum up to two hops. Fig. 1 represents the architecture of BSNs.



Fig. 1. BSNs architecture for patients in a ward of medical centre.

#### IV. RADIO MODEL

We would prefer to use first order radio model for the estimation of energy consumption by transmitting patient data wirelessly in BSN. The basic parameters of the model are  $E_t$  for energy transmission,  $l$  packet length and  $d$  transmission distance. Equation (1) for data transmission [22]:

$$E_t(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs} d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{mp} d^4, & d \geq d_0 \end{cases} \quad (1)$$

Where  $E_t(l, d)$  is the ratio of consumed power by a sensor node in data transmission, power consumed is directly proportional to the packet length  $l$  and  $d^2$  distance. Power consumption depends upon the communication distance, long distance more energy consumption and short distance less energy consumption.

$$E_r(l) = lE_{elec} \quad (2)$$

Equation (2) is used to measure the consumed energy on patient data receiving where  $E_r(l)$  Energy required for receiving data by a sensor node,  $l$  is packet length and  $E_{elec}$  Energy consumption per bit as:

$$E_{elec} = 50nJ/bit \\ d_0 = 100m$$

The distance in our scheme  $d < d_0$  so we use free space model  $\epsilon = \epsilon_{fs} = 10 \text{ pJ/bit/m}^2$ .  $\epsilon_{fs}$  is the free space model amplifier energy factor.

#### V. ATTACK MODEL

It is assumed that the BSN equipment are in reach of the attacker and may launch attacks like replay, eavesdropping, masquerading etc. BSN communicate patient physiological data which are the top personal secrets of the patient and should be protected from illegal use to safe the human life risk. For this purpose a cost effective and secure technique should be developed to tackle these issues. Preloading of initial secrets keys by the ward physician has to be done securely. Legal and illegal nodes should be differentiated through nodes authentication to protect the network from unauthorized access of patents personal diseases information and avoid masquerading attack. Secure exchange of secret keys for the session data communication is the requirement of our proposed scheme. As asymmetric cryptosystem is costly so we would prefer to use symmetric cryptosystem for the confidential communication of the patient data and avoid eavesdropping, chosen cipher and plain text attacks.

#### VI. PROPOSED SCHEME

Our proposed scheme comprised on three stages, deployment stage, node authentication stage and secure data communication stage. The notations used throughout this paper are listed in Table 1.

##### A. Deployment Stage

Deployment stage is the first stage in which initially required information are loaded to BSN devices. The corresponding ward MO generates a master secret key  $M_{sk}$  and deploys that key on MS, BS and sensor nodes.

TABLE I. NOTATION GUIDE

Symbols	Description
$P_i$	Patient in ward
$s_i$	Sensor node
PRNG	Pseudo Random Number Generator
R	random number
$S_k$	Session key
nonce	Number used once
ID	Identification number of sensor
DES	Data encryption standard
$M_k$	Master secret key
$h / hi'$	One-way hash function
$C_i$	Cipher text
$E_k/D_k$	Encryption / Decryption with key $k$

Unique IDs of the sensor nodes having  $M_{sk}$  are stored in BS and all relevant sensor nodes are deployed on patient body for monitoring health status of patients.

##### B. Node Authentication Stage

Node authentication is important in a situation where two or more biosensor nodes want to authenticate each other's identity or BS want to authenticate the identity of a legal node in a data communication networks. In this stage, biosensor nodes  $S_i$  send encrypted data to BS for authentication. BS decrypts the received data and authenticates biosensor nodes  $S_i$ . If authentication granted node will start secure communication using session key otherwise node is black listed and isolated from the network.

##### ALGORITHM.1. Key Agreement and Authentication

1. Preload Patient Master secret key  $M_{sk}$  on BS and Biosensor
2. Biosensor
  - a. Generate random number (R) called pre session key  $S_k$
  - b. Generate nonce
  - c. Computes  $C = E_{M_{sk}}(Nonce \parallel s_k \parallel ID)$
  - d. Sends C to BS
3. Base Station
  - a. Computes  $(Nonce \parallel s_k \parallel ID) = D_{M_{sk}}(C)$
  - b. If  $ID = \text{pre store ID so}$  Grant Authentication
  - Else
  - c. Blacklist the biosensor
  - d. If Grant Authentication
  - e. Computes  $C = E_{M_{sk}}(S_k)$
  - f. Sends C to MS and MO
4. Medical Server and Medical Officer
  - a. Computes  $S_k = D_{M_{sk}}(C)$
5. Base Station
  - a.  $C = E_{S_k}(Nonce + 1)$
  - b. Sends C to Biosensor
6. Biosensor
  - a.  $(Nonce + 1) = D_{S_k}(C)$

End

Authentication is required to ensure that only authorized nodes can join the network. Each sensor node has a default Pseudo Random Number Generator (PRNG) which generates a random number R called session key  $S_K$  and then generate *nonce*. Sensor node concatenate *nonce*, session key  $S_K$ , its own unique ID and encrypt on pre loaded master secret key  $M_{sk}$  then transmit to BS. At other end BS decrypt the received information by master secret key  $M_{sk}$  compare the received sensor node ID with its pre stored ID if matched,

node is legal and authentication is granted and otherwise the node is from intruder and black listed. After a node is authenticated BS increment the received *nonce* by 1 then encrypt it using session key  $S_K$  and sends to corresponding sensor node which decrypt the received message by its own  $S_K$ . Moreover, BS encrypts the session keys of authenticated nodes using  $M_{sk}$  and forward to MS and MO for onward secure communication. The overall scenario of proposed scheme is presented in Fig. 2.

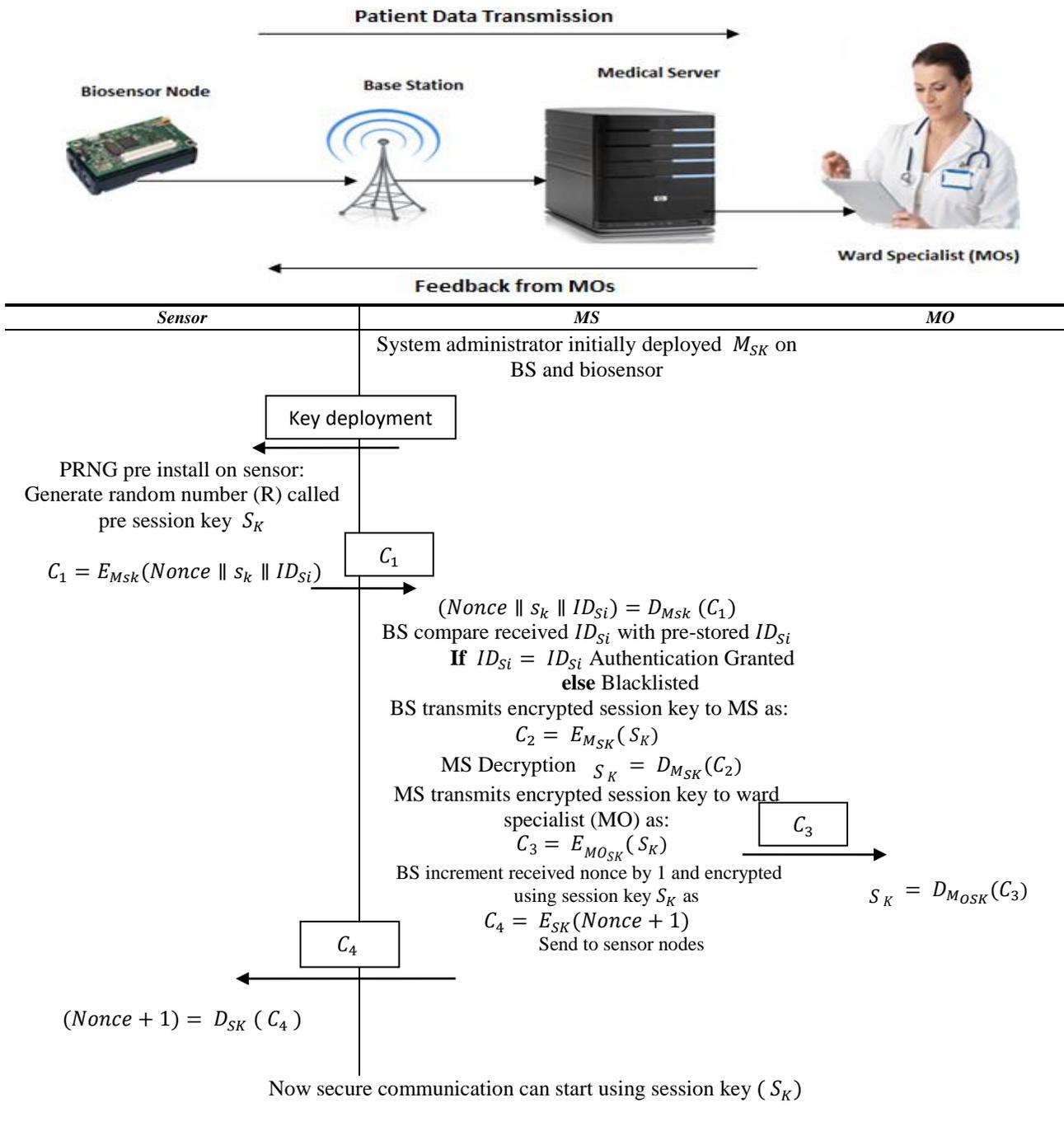


Fig. 2. Over all scenarios of proposed scheme.

### C. Secure Data Communication Stage

Secure communication of the sensed physiological data of vital signs is performed inside the ward of a hospital so the range of BSN in our proposed scheme is limited to a ward. Sensor nodes deployed on patients are directly connected with BS and the sensed patient data is forwarded to MS for quick response of the physician. Each node has its own Session key  $S_K$  and all these keys are also stored on BS and MS as in stage 2 which are further used for secure communication as when a message patient vital signs data packet is required to be sent to MO by a sensor node. This data packet will be encrypted by the session key  $S_K$  of that node and will be transmitted to medical server through BS. Similarly the integrity of patient data is gained by hash collision resistive Message Digest (MD5) as hash of patient data  $h(\text{patient data})$  is taken and hash value  $h_i$  is obtained then to obtain secure  $C_i$ , patient data and  $h_i$  is encrypted by session key  $S_K$  and  $C_i$  is transmitted to MS through BS. Now MS decrypt the received  $C_i$  by  $S_K$ ,  $h_i$  and patient data is obtained if  $h_i'$  (hash taken of received patient data by MS) is compared with  $h_i$  if found same then the received message is original and not changed otherwise changed by the attacker.

---

#### ALGORITHM.2. Secure Data Communication

---

##### A. Sensor Node

- I. For each sensor node  $s_i \in P_i$ 
  - {
  - a.  $h_i = h(\text{Patient data})$
  - b. Computes  $c_i = E_{s_k}(\{\text{Patient data}\}, h_i)$
  - c. Sends  $c_i$  to MS through BS
  - }
- II. End for

##### D. Medical Server

- I. For each biosensor node  $s_i \in P_i$ 
  - {
  - d.  $(\{\text{Patient data}\}, h_i) = D_{s_k}(c_i)$
  - e.  $h_i = h(\text{Patient data})$
  - f. Computes  $h_i' = h(\text{Patient data})$
  - g. Accept if  $h_i' = h_i$  hold
  - h. otherwise reject
  - }

End for

---

Security is depending upon two major parts. One as data security and the second is data privacy. In data security we study how data can be securely transmitted and stored and the second part only authorized users can access the patient personal information. In below Fig. 3 is represented the flow chart of secure data communication.

### VII. SECURITY ANALYSIS

The analysis to validate security features of our proposed scheme is represented here. Our proposed scheme provides the essentials security requirements of authentication, confidentiality and integrity.

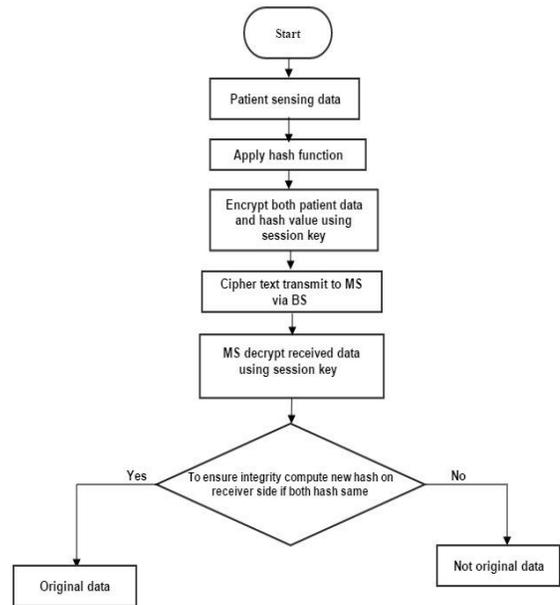


Fig. 3. Flow chart for secure data transmission.

##### A. Node Authentication

Upon receiving the request from a node for becoming the part of the network BS compare the ID of that sensor node with its pre installed IDs, if both of the IDs i.e. the received ID and the pre-stored ID are matched then that sensor node is authenticated otherwise rejected and thrown out from network by black listing that ID.

##### B. Data Confidentiality

In our proposed scheme, Master secret key is used for confidentially and sharing the session key ( $s_k$ ). Session key is used to make sure vital signs data transmission of patients between sensor nodes with base station and medical server. Confidentiality of patient is maintained through DES cipher which encrypt the sensed session data before to be communicated to the BS and MS so that to protect this personal data from the illegal reading .MS forward it to MO for quick health care.

##### C. Patient Data Integrity

Data integrity is that feature of our scheme which obstructs the alteration of the patient precious personal data from illegal use for any bad intention. Integrity in our scheme is achieved using hash collision resistive Message Digest (MD5) in such a way that the received hash  $h_i$  is compared with computed hash ( $h_i'$ ) is similar then data is safe and not changed otherwise incorrect data is received.

##### D. Scalability

Scalability is the property of our proposed scheme as whenever a sensor node is required to be added to the network or a sensor is to be removed from the network or a sensor is needed to be changed due to low battery power or any other fault by any of these activities the normal functionality of the network is not affected.

VIII. PERFORMANCE ANALYSIS

Performance analysis of our proposed scheme and two existing schemes with respect to computational cost, communication overhead, storage and energy consumption in term of efficiency is given below.

A. Computational cost

No expensive and major operations like ECPM and M-Exp are involved in our proposed scheme. In designed scheme [12], four ECPM and two M-Exp operations and in scheme [8] two M-Exp are used. Graph in Fig. 4 shows that our scheme is 90.99 % efficient in computation cost as compare to [12], 89.67 % as compare to [8] and 69.98 % as compare to [23]. In our scheme we implement the experiment done in [24] on MICA2 sensor that is operational with low power ATmega128 8-bit micro-controller at 7.3728 MHz, 128 KB nonvolatile memory ( ROM) and 4 KB volatile memory ( RAM). One major operation ECPM uses 0.81s using 160 bits elliptic curve [25] and RSA 1024 bits M-Exp takes 22 seconds [26]. DES encryption and decryption execution time [27] is same which 4.543859 seconds. We calculate the computation cost of our scheme in comparison with the [8], [12] on the basis of the results of [23], [24], [26]-[28].

According to scheme [28] the 3rd generation MICA2 needs 2.66s for pairing computation. The computational time of our proposed scheme is negligible as compared with others existing schemes [8], [12] because we used symmetric algorithm for encryption and decryption as well as our scheme is more suitable for resource constraint environment of BSN. One ECPM operation consumes 19.1Mj and one pairing computation operation consumes 62.73mJ energy [24], [28]. Our scheme have no major operation so energy consumption as compared to others existing schemes is negligible.

TABLE II. ASSOCIATED PARAMETER AND DATA SIZE INVOLVED IN DATA COMMUNICATION

Associated Parameters	Data Size
RSA Key	1024 bits
ECC Key	160 bits
DES Session Key	64 bits
Master Secret Key	128 bits
Sensor ID	48 bits
nonce	16 bits

Comparison of Computation Cost of existing and proposed

1) Computation cost of our proposed scheme is compared with scheme [8] as:

Computation cost efficiency:

$$\frac{(2 * 22)s - (4.543859) s}{(2 * 22)s} \% = 89.67 \%$$

2) Computation cost of our proposed scheme is compared with scheme [12] as:

Computation cost efficiency:

$$\frac{(2 * 22 + 4 * 1.61)s - (4.543859) s}{(2 * 22 + 4 * 1.61)s} \% = 90.99 \%$$

3) Computation cost of our proposed scheme is compared with scheme [23] as:

Computation cost efficiency:

$$\frac{(2*30.67+1(14.62))s-(4.543859) s}{(2*30.67+1(14.62))s} \% = 69.98 \%$$

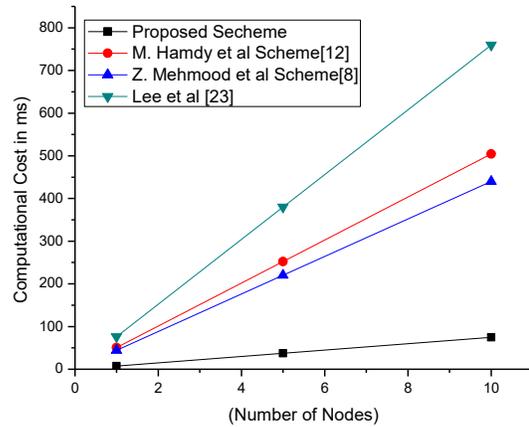


Fig. 4. Computational cost comparison.

B. Communication Overhead

The proposed scheme communication overhead as compared with other existing schemes [8], [12], [23] and the computed values are shown in Tables 3, 4 and 5 and then design graph according to these computed values which are shown in Fig. 5.

Communication overhead of our proposed scheme with schemes [8], [12], [23] is represented in Fig. 5 where our scheme shows 84.2% as compared to scheme [8], 85.7% efficiency than scheme [12] and 78.57% than [23].

TABLE III. COMMUNICATION OVERHEAD COMPARISON WITH Z. MEHMOOD ET AL.

Scheme	Communication Overhead	Communication Overhead Reduction in Percent
Z. Mehmood et al. [8]	(1024+192)bits	$\frac{1216-192}{1216} \% = 84.2 \%$
Proposed	(128+16+48) bits	

TABLE IV. COMMUNICATION OVERHEAD COMPARISON WITH M. HAMDY ET AL.

Scheme	Communication Overhead	Communication Overhead Reduction in Percent
M. Hamdy et al. [12]	(1024+320)bits	$\frac{1344-192}{1344} \% = 85.7 \%$
Proposed	(128+16+48) bits	

TABLE V. COMMUNICATION OVERHEAD COMPARISON WITH LEE ET AL.

Scheme	Communication Overhead	Communication Overhead Reduction in Percent
Lee et al. [23]	2(160+160+128)bits	$\frac{896-192}{896} \% = 78.57 \%$
Proposed	(128+16+48) bits	

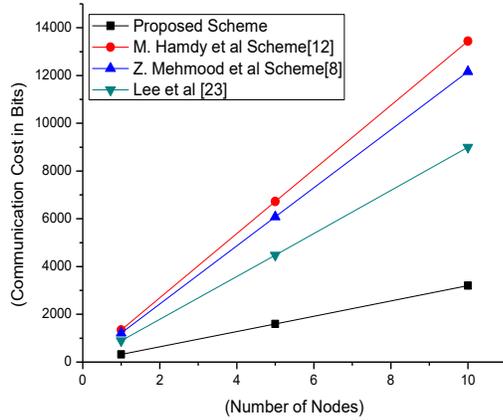


Fig. 5. Communication overhead comparison.

C. Memory Requirement for Key Storage

The proposed scheme memory for key storage as compared with other existing schemes [8], [12], [23] and the computed values are shown in Tables 6, 7 and 8 and then design graph according to these computed values which is shown in Fig. 6.

The NIST standard key size for algorithms AES, DES, RSA, ECC is given in Table 2. Fig. 6 represent analysis of memory requirement our proposed scheme with schemes [8], [12], [23]. Our proposed scheme reduces 75% as compare to scheme [8], 80% memory requirements as compare to scheme [12] and 28.57 than [23].

TABLE VI. STORAGE COMPARISON OF PROPOSED SCHEME AND Z. MEHMOOD ET AL.

Schemes	Key stored	Approximate key size in bits	Percent reduction in memory storage
Z. Mehmood et al. [8]	$k_{pi}, k_{Cj}, e_{gw}$	128+128+1024	$\frac{1280-320}{1280}\% = 75\%$
Proposed	$M_{sk}, S_k, ID, nonce$	128+128+48+16	

TABLE VII. STORAGE COMPARISON OF PROPOSED SCHEME AND M. HAMDY ET AL.

Schemes	Key stored	Approximate key size in bits	Percent reduction in memory storage
M. Hamdy et al. [12]	$d_{si}, p_{si}, e, k_{pi}, k_{Cj}$	160+160+1024+128+128	$\frac{1600-320}{1600}\% = 80\%$
Proposed	$M_{sk}, S_k, ID, nonce$	128+128+48+16	

TABLE VIII. STORAGE COMPARISON OF PROPOSED SCHEME AND LEE ET AL

Schemes	Key stored	Approximate key size in bits	Percent reduction in memory storage
Lee et al. [23]	$k_{pi}, k_{Cj}, e_{gw}$	160+160+128	$\frac{448-320}{448}\% = 28.57\%$
Proposed	$M_{sk}, S_k, ID, nonce$	128+128+48+16	

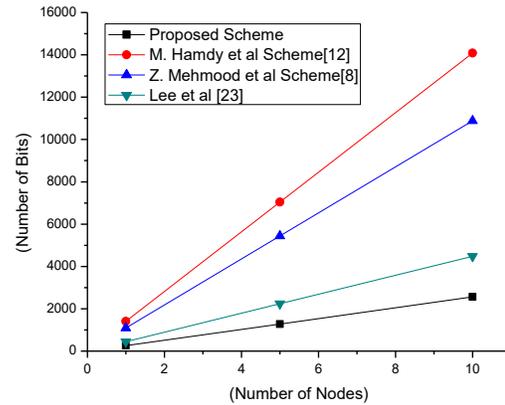


Fig. 6. Memory requirement for key storage.

D. Key Agreement and Authentication Delay

The delay in authentication and key agreement of the proposed scheme in comparison with existing schemes [8], [12], [23] is shown in graph Fig. 7 where the delay of our proposed scheme is very less and negligible.

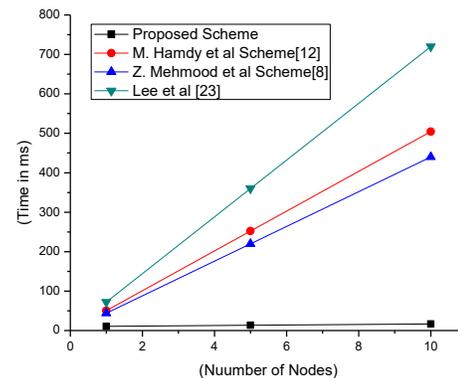


Fig. 7. Key agreement and authentication delay.

E. Energy Consumption for Authentication and key Agreement

The communication distance of our proposed scheme is less than 100 meters as per the standard size of the ward as the distance in our scheme  $d < d_0$  so we use free space model

$\epsilon = \epsilon_{fs} = 10 \text{ pJ/bit/m}^2$  where  $\epsilon_{fs}$  is the amplifier energy factor of the free space model. Graph in Fig. 8 shows that our scheme is quite better than the existing schemes [8], [12], [23].

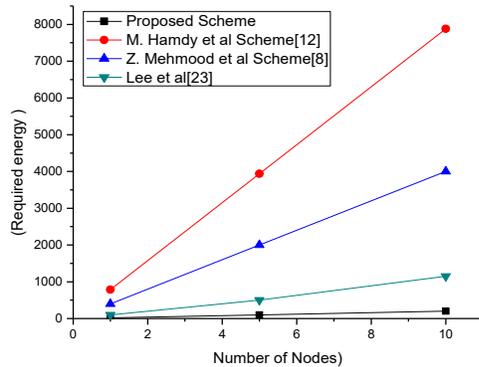


Fig. 8. Energy consumption for authentication and key agreement.

### IX. CONCLUSION

In this paper, an efficient key agreement and nodes authentication scheme is presented which is compared with other solutions to prove the efficiency of our proposed scheme. Our proposed three stages solution not only protects patient data from unauthorized elements but also overcome the weaknesses of the existing schemes and thus proves its suitability for the resource constrained environment of BSNs. The comparison of the existing three schemes and our scheme has shown that our scheme leads in efficiency as 90.99% in computation cost as compared to M. Hamdy et al., 89.67% as compare to Z. Mehmood et al. and 69.98% as compared to Lee et al. 85.7% in communication overhead than M. Hamdy et al., 84.2% than Z. Mehmood et al. and 78.57% than Lee et al. in storage, 80% than M. Hamdy et al., 75% than Z. Mehmood et al. and 28.57% than Lee et al.

### REFERENCES

- [1]. Sahana and I. S. Misra, "Implementation of RSA security protocol for sensor network security: Design and network lifetime analysis," 2011 2nd Int. Conf. Wirel. Commun. Veh. Technol. Inf. Theory Aerosp. Electron. Syst. Technol. (Wireless VITAE), pp. 1–5, 2011.
- [2]. K. S. K. Jingwei Liu, "Hybrid security mechanisms for wireless body area networks," 2010 Second Int. Conf. Ubiquitous Futur. Networks, pp. 98–103, 2010.
- [3]. P. M. Khan, A. Husain and K. S. Kwak, "Medical Applications of Wireless Body Area Networks," Int. J. Digit. Content Technol. its Appl., vol. 3, no. 3, pp. 185–193, 2009.
- [4]. S. L. Keoh, E. Lupu, and M. Sloman, "Securing body sensor networks: Sensor association and key management," 7th Annu. IEEE Int. Conf. Pervasive Comput. Commun. PerCom 2009, 2009.
- [5]. M. Li, S. Yu, W. Lou, and K. Ren, "Group device pairing based secure sensor association and key management for body area networks," Proc. - IEEE INFOCOM, pp. 1–9, 2010.
- [6]. S. L. Keoh, "Efficient Group Key Management and Authentication for Body Sensor Networks," 2011 IEEE Int. Conf. Commun., pp. 1–6, 2011.
- [7]. D. He, S. Chan, Y. Zhang, and H. Yang, "Lightweight and confidential data discovery and dissemination for wireless body area networks," IEEE J. Biomed. Heal. Informatics, vol. 18, no. 2, pp. 440–448, 2014.
- [8]. Z. Mehmood, Nizamuddin, S. Ashraf Ch., W. Nasar, and A. Ghani, "An efficient key agreement with rekeying for secured body sensor networks," 2012 2nd Int. Conf. Digit. Inf. Process. Commun. ICDIPC 2012, pp. 164–167, 2012.
- [9]. N. U. Amin, M. Asad, and S. A. Chaudhry, "An authenticated key agreement with rekeying for secured body sensor networks based on hybrid cryptosystem," Proc. 2012 9th IEEE Int. Conf. Networking, Sens. Control, pp. 118–121, 2012.
- [10]. G. Wu, L. Yao, B. Liu, K. Yao, and J. Wang, "A biometric key establishment protocol for body area networks," Int. J. Distrib. Sens. Networks, vol. 2011, 2011.
- [11]. X. H. Le, M. Khalid, R. Sankar, and S. Lee, "An Efficient Mutual Authentication and Access Control Scheme for Wireless Sensor Networks in Healthcare," J. Networks, vol. 6, no. 3, pp. 355–364, 2011.
- [12]. M. H. Eldefrawy, M. K. Khan, and K. Alghathbar, "A key agreement algorithm with rekeying for wireless sensor networks using public key cryptography," Anti-Counterfeiting Secur. Identif. Commun. (ASID), 2010 Int. Conf., pp. 1–6, 2010.
- [13]. P. Kumar, S. G. Lee, and H. J. Lee, "A user authentication for healthcare application using wireless medical sensor networks," Proc.- 2011 IEEE Int. Conf. HPCC 2011 - 2011 IEEE Int. Work. FTDCS 2011 -Workshops 2011 Int. Conf. UIC 2011- Work. 2011 Int. Conf. ATC 2011, vol. 1, pp. 647–652, 2011.
- [14]. D. Liu, P. Ning, and R. Li, "Establishing pairwise keys in distributed sensor networks," ACM Trans. Inf. Syst. Secur., vol. 8, no. 1, pp. 41–77, 2005.
- [15]. A. Boukerche and Y. Ren, "The design of a secure key management system for mobile ad hoc networks," 2008 33rd IEEE Conf. Local Comput. Networks, pp. 320–327, 2008.
- [16]. R. Sampangi, S. Dey, S. Urs, and S. Sampalli, "a Security Suite for Wireless Body Area Networks," Int. J. Netw. Secur. Its Appl., vol. 4, no. 1, pp. 97–116, 2012.
- [17]. M. Meingast, T. Roosta, and S. Sastry, "Security and privacy issues with health care information technology," Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 1, pp. 5453–5458, 2006.
- [18]. C. C. Y. Poon, Y. T. Zhang, and S. Di Bao, "A novel biometrics method to secure wireless body area sensor networks for telemedicine and M-health," IEEE Commun. Mag., vol. 44, no. April, pp. 73–81, 2006.
- [19]. C. Jiang, B. Li, and H. Xu, "An Efficient Scheme for User Authentication in Wireless Sensor Networks," 21st Int. Conf. Adv. Inf. Netw. Appl. Work., pp. 438–442, 2007.
- [20]. K. Malasri and L. Wang, "Addressing security in medical sensor networks," Proc. 1st ACM SIGMOBILE Int. Work. Syst. Netw. Support Healthc. Assist. living Environ. - Heal. '07, p. 7, 2007.
- [21]. K. Malasri and L. Wang, "Design and implementation of a secure wireless mote-based medical sensor network," Sensors, vol. 9, pp. 6273–6297, 2009.
- [22]. F. N. Wu, I. H. Li, and I. E. Liao, "A Traffic Load-Aware Energy Efficient Protocol for Wireless Sensor Networks," vol. 2008, pp. 10–12, 2008.
- [23]. Lee, Y. S., Alasaarela, E., & Lee, H. J, "An Efficient Encryption Scheme using Elliptic Curve Cryptography (ECC) with Symmetric Algorithm for Healthcare System," Int. J. Sec and Its Applica, vol.8, no. 3, pp.63-70.
- [24]. N. Gura, A. Patel, A. Wander, H. Eberle, and S. C. Shantz, "Comparing elliptic curve cryptography and RSA on 8-bit CPUs," in Cryptographic Hardware and Embedded Systems (Lecture Notes in Computer Science), vol. 3156. New York, NY, USA: Springer-Verlag, 2004, pp. 119–132.
- [25]. Li, Fagen, and Pan Xiong. "Practical secure communication for integrating wireless sensor networks into the internet of things." Sensors Journal, IEEE13, no. 10, 2013, pp.3677-3684.
- [26]. N. Gura, A. Patel, A. Wander, H. Eberle, and S. C. Shantz, "Comparing elliptic curve cryptography and RSA on 8-bit CPUs," Lect. notes Comput. Sci., pp. 119–132, 2004.
- [27]. S. Singh, S. K. Maakar, and S. Kumar, "A Performance Analysis of DES and RSA Cryptography," Int. J. Emerg. Trends Technol. Comput. Sci., vol. 2, no. 3, pp. 418–423, 2013.
- [28]. P. Szczechowiak, A. Kargl, M. Scott, and M. Collier, "On the application of pairing based cryptography to wireless sensor networks," in Proc.2nd ACM Conf. Wireless Netw. Security, Zurich, Switzerland, 2012, pp. 1–12.

# Design of Efficient Pipelined Router Architecture for 3D Network on Chip

Bouraoui Chemli

Electronics and Microelectronics Laboratory  
Faculty of Sciences of Monastir  
Monastir University, Monastir, Tunisia

Abdelkrim Zitouni

College of Education in Jubail  
University of Dammam,  
Jubail Industrial, KSA

Alexandre Coelho

TIMA Laboratory, INPG  
University of Grenoble-Alpes  
Grenoble, France

Raoul Velazco

TIMA Laboratory, INPG  
University of Grenoble-Alpes Grenoble,  
France

**Abstract**—As a relevant communication structure for integrated circuits, Network-on-Chip (NoC) architecture has attracted a range of research topics. Compared to conventional bus technology, NoC provides higher scalability and enhances the system performance for future System-on-Chip (SoC). Divergently, we presented the packet-switching router design for 2D NoC which supports 2D mesh topology. Despite the offered benefits compared to conventional bus technology, NoC architecture faces some limitations such as high cost communication, high power consumption and inefficient router pipeline usage. One of the proposed solutions is 3D design. In this context, we suggest router architecture for 3D mesh NoC, a natural extension of our prior 2D router design. The proposal uses the wormhole switching and employs the turn mod negative-first routing algorithm. Thus, deadlocks are avoided and dynamic arbiter are implemented to deal with the Quality of Service (QoS) expected by the network. We also adduce an optimization technique for the router pipeline stages. We prototyped the proposal on FPGA and synthesized under Synopsys tool using the 28 nm technology. Results are delivered and compared with other famous works in terms of maximal clock frequency, area, power consumption and estimated peak performance.

**Keywords**—3D network on chip; router optimization; turn model; parallel communication; router pipeline stages

## I. INTRODUCTION

During the last decade, the evolution of technology has shrunk the dimension of transistor and made possible its integration of billions on the same chip. Thus, this increasing number of transistor densities allows the integration of countless cores on a single chip. Therefore, it requires a powerful on-chip interconnection scheme to satisfy the communication between the large numbers of cores on chip [1]. As this communication plays a major role in determining the system performance, traditional on-chip interconnection schemes are no longer suitable for multi-Processor System on Chip (MPSoC) due to their lack of parallelism integration, scalability and resource management [2]. Recently, Network on Chip (NoC) has been introduced as the best candidate to handle the on-chip communication requirements overcoming the limitations of traditional interconnections [3]. NoC

architectures are mainly composed of the router, which allows packets' distribution along the network, the network interface which grants network access and the link which permits the connection of NoC components. Based on a scalable architecture, NoC enables high bandwidth and overalls the system performances [4]. Concurrently, transistor densities keep increasing. They in fact render the integration of hundreds of cores on a planner chip not satisfying for future applications. The latter are getting more complex demanding a higher performance system to handle parallel computing and provide higher bandwidth. At the same time, semiconductor industries are exploiting three dimensional integrated circuits (3D IC) which provide short global interconnects, lower power consumption, and higher performance [5]. Bringing together 2D NoC architecture with 3D IC technology, makes the design of 3D NoC possible which is a stack multiple die in the vertical axis that are interconnected through silicon via (TSV) [6]. Compared to 2D NoC, 3D NoC offers higher performance and higher package density. Thus, 3D NoC satisfies the on-chip communication requirements for future MPSoC.

In this paper, we propose extensible, flexible and efficient router architecture and its implementation for 2D and 3D mesh topologies. We present optimized router pipeline stages in order to reduce their dependencies and improve the router efficiency. The proposal adopts wormhole switching techniques, turns model negative-first routing algorithm to avoid deadlocks and a dynamic arbiter to improve the Quality of Service(QoS) expected by the 3D NoC. In order to evaluate the performance and the hardware cost, we prototype the proposal on FPGA and compare it with other designs.

The rest of the paper is organized as follows. Section 2 deals with related work. Section 3 gives an overview of the 2D router architecture. Section 4 tackles the optimized pipeline stages of the 3D router in detail. Section 5 provides the results evaluation to conclude the dissertation in Section 6.

## II. RELATED WORK

Many works have been proposed in literature addressing the on-chip interconnection design challenges. One of the

brilliant solutions is the extension from 2D NoC to the 3D NoC architectures. The proposed 3D NoC design used 3D mesh topology due to its simplicity, regularity, scalability as it is a direct extension of the 2D mesh topology. In [7], the authors investigated resilience and adaptivity against fault on 3D NoC. They proposed a fault-tolerant routing algorithm named 4NP-First. Compared to stochastic random walk routing scheme, their turn model-based-routing algorithm shows better robustness against fault. However, their architecture implements two virtual channels, one for the transmitted original packet and the second for the redundant packet. Such a technique has a negative impact on the power consumption and the hardware cost. In [8], the authors presented AFRA, a deadlock free routing algorithm that tolerates faults for 3D mesh NoC. When faults are not detected, AFRA sends packets through ZXY. If there is fault detection, however, flits are forwarded through XZXY. This routing scheme shows good performance and robustness against faults. Nonetheless, AFRA focuses only on vertical link faults and ignores horizontal faults. Despite not requiring any additional virtual channels to avoid deadlock, it needs some global information to be stored for some overhead to be added to the router hardware complexity. In [9], the authors proposed 3D mesh NoC. Their scalable architecture adopts wormhole switching and implements look-ahead-routing algorithm. Their hardware implementation on FPGA illustrates a good performance in terms of area and maximal clock frequency. However, the deadlock situation may rise with any adaptive routing algorithm. In order to avoid deadlocks, they have to make their routing algorithm minimal or use virtual channels. In [10], authors investigated on topology and routing algorithm for 3D NoC. They suggested a modified structure of tree topology in order to reduce the degree and the diameter of the network which are vital characteristics for the network topology affecting the system performance. However, future applications require high throughput and low latency which cannot be provided by 3D tree topology. In [11], the authors introduced adaptive router architecture for heterogeneous 3D NoC. They implement a deadlock free adaptive routing algorithm. Compared with homogenous router, they modify the TSV selector and the routing logic blocks which enable hardware cost reduction and performance improvement. Only if the destination node is in a different layer, the TSV selector chooses a valid router as a vertical hub for interlayer routing; otherwise the routing logic would have a similar structure as 2D router. Conversely, heterogeneous NoC has a fixed topology and cannot be customized regarding an application requirement. In [12], authors presented router architecture for symmetric 3D mesh NoC. They implement dimension order XYZ routing algorithm that adopts credit-based flow control and uses virtual channel to avoid deadlock. A priority-based-scheduling is used to support and manage the different levels of QoS. Their results display a low latency and high bandwidth. However, their design suffers from area and power overheads. In [13], we proposed router architecture for 3D mesh based NoC. We implemented turn model negative-first routing algorithm in order to avoid deadlock conflicts. We

adopted a packets priority scheme and round robin arbiter to ensure the QoS expected by the network. However, the proposal suffers from dependency between router pipeline stages which increases the hardware cost affecting the average latency of the network.

In this study, we suggest 3D NoC router based on our preceding design [13], [14]. The proposal implements the Negative-First 3D turn model routing algorithm which employs some routing restrictions to prevent packets from deadlock. It also uses dynamic arbiter to fairly serve packets and enlance the QoS. We go through the optimized router pipeline stages in detail and its impact on reducing the hardware cost and on improving the system performance in terms of bandwidth.

### III. 2D ROUTER ARCHITECTURE

Fig. 1 illustrates the NoC topology which is a 3x3 size mesh using wormhole switching policy and the credit-based flow control. To locate and differentiate between routers in the network, we give every router a unique address defined in XY coordinates. Each router can be connected with maximum of four adjacent routers as well as the local intellectual property (IP). The number of ports per router depends on its position in the network. In order to reduce the chip area and the power consumption, we have to eliminate any unused ports.

In their way to destination, packets must come across three pipeline stages as shown in Fig. 2. The first stage is the routing calculation (RC) in which the destination address is compared to the router address in order to define the next output port. Then, this information is sent to the next stage which is the switch allocation (SA). Based on its arbiter, this stage fairly serves packets to each destination. Finally, the information about the adequate output port is sent to the semi-crossbar traversal stage (ST) ensuring the traversal of packets to their destinations [14].

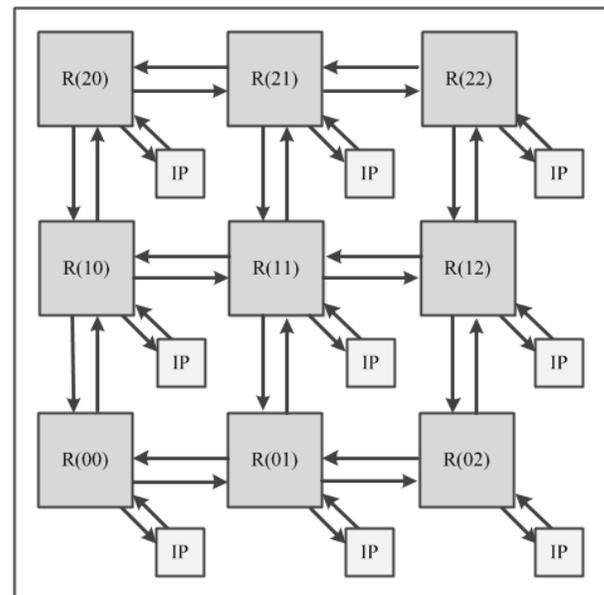


Fig. 1. 3x3 NoC mesh topology.

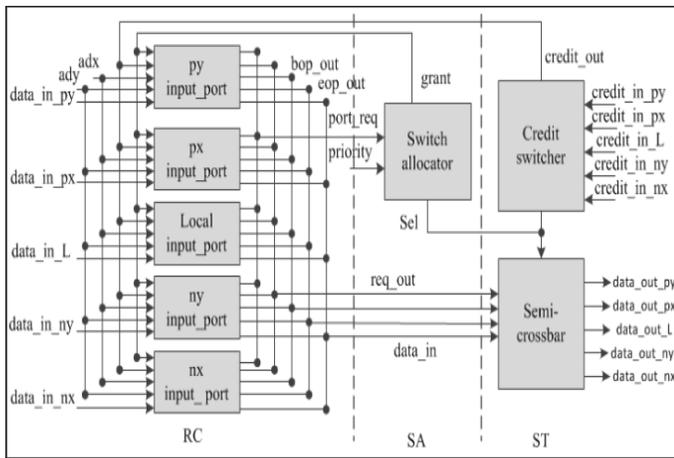


Fig. 2. 2D Router pipeline stages without optimization.

#### IV. 3D ROUTER ARCHITECTURE

##### A. Optimized Router Pipeline Stages

We optimize a previous version of our 3D router design [13]. Fig. 3 shows the 3D router pipeline stages. As for the 2D router, the pipeline stages are the routing calculation, the switch allocation and the semi-crossbar traversal.

We observe in typical pipeline stages a dependency between the routing calculation stage and the switch allocation stage. Each packet must wait for control signals to move from one phase to another. This fact increases the latency of the network. We break those stages' dependency by executing the routing calculation process and the switch allocation concurrently. Furthermore, on the previous routing calculation process, to define the output port, the X address of the flit is first compared to the X address of router, the Y address of the flit is then compared to the Y address of router to finally compare the Z address of the flit to the Z address of router. This makes the routing algorithm more complex and affects the hardware complexity of the router.

Therefore we will compare the flit address with the router address without any decoded comparison. By using those optimizations, we aim to reduce the hardware cost and the communication latency of the network design.

##### B. Topology

As shown in Fig. 4, the NoC topology is a 3x3x3 size mesh. Each router can contain up to seven bidirectional ports. One port is connected to the local IP while the other six are connected to the adjacent routers in each direction of the network (north, east, south, west, up and down). Each router is defined by its XYZ coordinates. We choose the 3D mesh topology because it is the direct extension of 2D mesh topology and also due to several advantages like simplicity of implementation, regularity and scalability over other topologies.

##### C. Communication Flow

The proposal adopts the wormhole switching policy. As shown in Fig. 5, the packet is composed of two types of flits; the header flit and the body flit. The header flit is composed of 32 bits. First, six bits are allocated to the destination address.

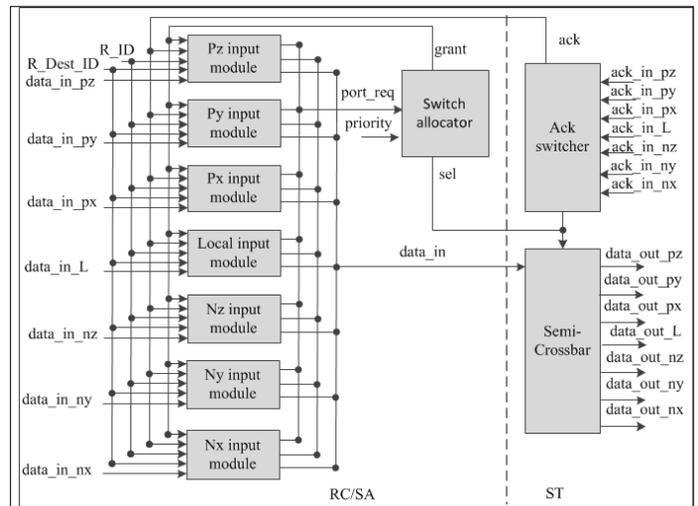


Fig. 3. 3D router pipeline stages with optimization.

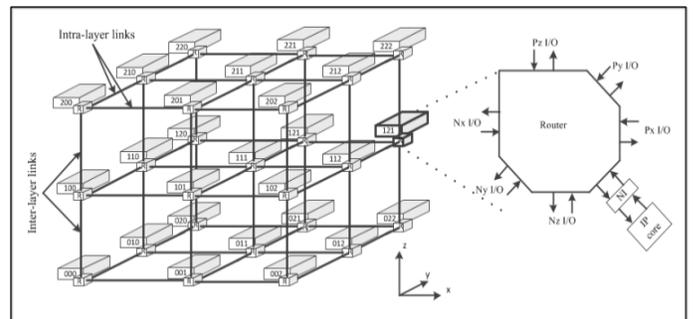


Fig. 4. 3D NoC based on 3x3x3 mesh topology.

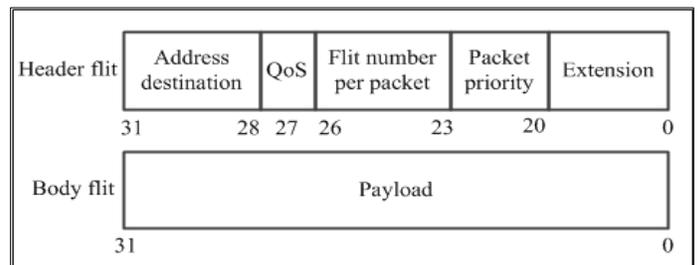


Fig. 5. Packet format.

One bit is then allocated to the quality of service required by the NoC. Next four bits are allocated to the flit number per packet. Next three bits are allocated to the packet priority leaving the rest of the bits to constitute an extension. The body flit is composed of 32 bits data payload. The packet format and the flit size can be changed according to the application specifications.

##### D. Routing algorithm and deadlock avoidance

In [15], Glass presented the turn model for partially adaptive routing algorithm and targeted the mesh topology. This model designs a wormhole routing algorithm without the addition of physical or virtual channels. The principal of this model is to study all turns that can be taken by the packets in the network from source node to destination node as well as the cycles formed by those turns. A turn is referred to as a 90

degree change in the direction of the packet and the cycle is referred to as four turns. Those cycles may enter packets into dependencies waiting named deadlocks leading to the network frailer. Therefore, they eliminate enough turns to prevent cycle's concurrency and make a deadlock-free routing algorithm. Fig. 6 shows an example of deadlock involving four packets. Fig. 6(a) displays a deadlock situation between packets from different plans. Fig. 6(b) presents a deadlock situation between packets belonging to the same plan.

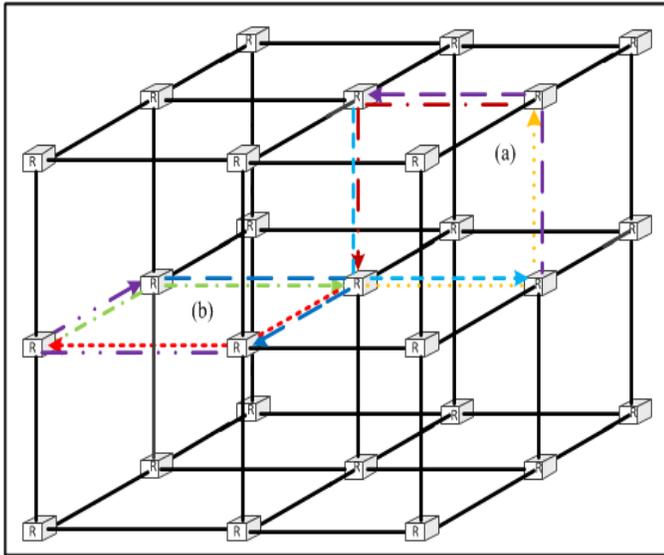


Fig. 6. Deadlock involving four packets from: (a) different plans, (b) same plans.

In 3D mesh, when flits travel between routers to reach their destination, they can pursue six directions: north, east, south, west, up and down. Each flit can make up to 24 turns, 8 turns in each plan (x,y), (x,z) and (y,z). In order to eliminate deadlock, we have to break cycles by prohibiting two turns at each plan. Glass [15] has proposed three turn model routing algorithms which are negative-first, west-first and north-last. We chose the negative-first routing algorithm since it is a simple extension in 3D, symmetric and doesn't require a packet ordering. The analysis of turns in the turn model is based on the XYZ coordinates and the directions of the flits in the network are defined by north, east, south, west, up or down. Thus, to simplify the terminology, the +y is north direction, the +x is east direction, the -y is south direction, the -x is west direction, the +z is up direction and the -z is down direction. The Negative-First routing algorithm routes packets first adaptively along -x, -z and -y and then adaptively along +y, +x and +z. We illustrate in Fig. 7(a) the prohibited turns in different four routers of the network. Regarding its position in the network, each router must eliminate turns from positive direction to negative direction in order to avoid deadlock. As shown in Fig. 7(b), solid lines indicate the allowed turns and the dash lines indicate the prohibited turns in negative-first routing algorithm from each plan of the 3D mesh. When flits are received by the input ports, each input port of each router performs the routing calculation independently of each other. So, the router can handle up to seven flits at the same time. This distributed routing is used to decrease the router latency and the average latency of the network.

In order to define the output port, the routing calculation process compares the destination address of the flit with the current router address and it takes into consideration the restriction turns by the negative-first routing algorithm to avoid deadlock:

- If the destination address is equal to the router address + 1, then the output port will be px, else the output port will be nx.
- If the destination address is equal to the router address + 3, then the output port will be py, else the output port will be ny.
- If the destination address is equal to the router address + 9, then the output port will be pz, else the output port will be nz.

### E. Switch allocation

The router implements distributed arbitration schemes. Thus, the switch allocator contains seven arbiter modules similar to the ones presented in Fig. 8 (one arbiter for each input port). The arbiter controls the connection between input ports and output ports. In order to avoid conflicts, especially when different flits from different input ports demand access to the same output port at the same time, an arbitration scheme is necessary to serve flits fairly. We use a dynamic arbiter that is composed of priority based scheduling, C-element ports and round-robin arbiter. The proposed arbiter prevents and solves conflicts access to the output port based on the priority comparator that compares the incoming flits priorities. It also provides the highest flits priority signals to the C-elements combining those signals with the corresponding requestor to be sure that there is at least one flit demand access to the output port. Then the C-elements send this information to the round-robin arbiter. In this manner, only flits with the highest priority and that demand access to the output port will be served.

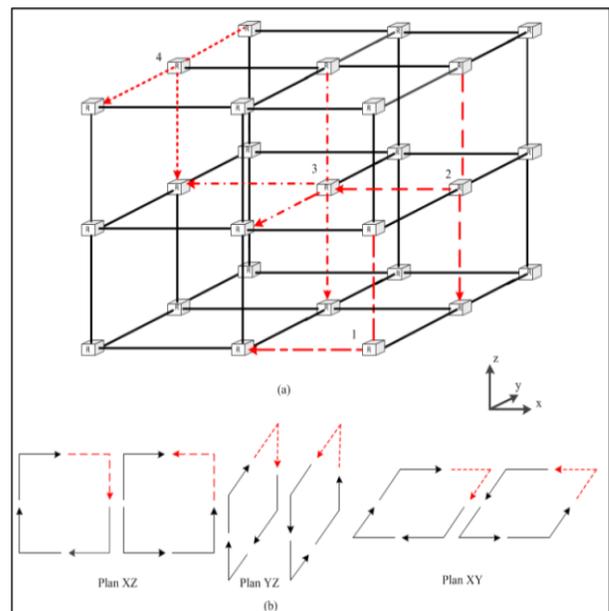


Fig. 7. (a) prohibited turns in four routers of 3D NoC (b) six turns allowed (solid arrows) in negative first routing.

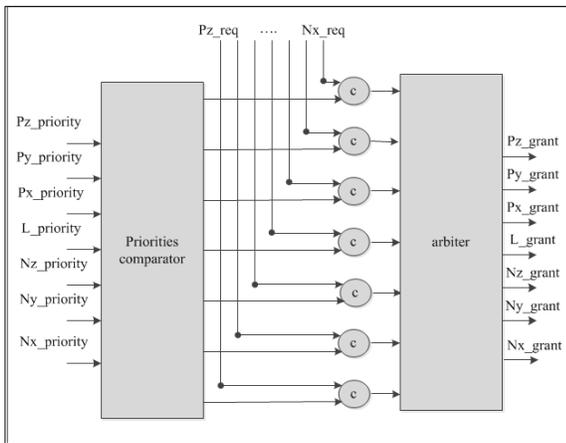


Fig. 8. Arbitre module.

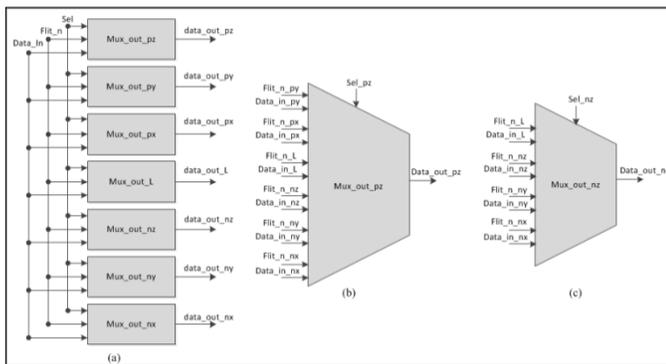


Fig. 9. (a) semi-crossbar circuit, (b) multiplexer of port p, and (c) multiplexer of port n.

F. Semi-Crossbar traversal

The final pipeline stage is the semi-crossbar which is similar to a bridge that interconnects the output port of the current router with the input port of the next router. The semi-crossbar waits for information signal about the selected output port from the switch allocator. Regarding this information, the semi-crossbar establishes an interconnection and sends flits to the adequate output port. The flit number per packet signal informs the semi-crossbar that all flits are transmitted and the channel is free to be used by another flits transmission. The semi-crossbar is based on multiplexer circuit. It uses seven multiplexers, one for each output port as presented in Fig. 9. Regarding the negative-first routing algorithm restrictions, we need to use two types of multiplexers, one for ports type p which can receive flits from any input port as shown in Fig. 9(b) and the other for ports type n which can only receive flits from input port n or from local port as shown in Fig. 9(c).

V. EXPERIMENTAL RESULTS

This section starts with an overview of the hardware complexity between the 2D and the proposed 3D routers designs. Then, it provides a Synopsis of different router implantations. Finally, it presents a results' comparison of the proposed design with other designs. The proposed router is designed and simulated in VHDL language at RTL level. The implementation and evaluation results targeted both FPGA and ASIC technologies and are provided in terms of maximal clock

frequency, area, power consumption and the estimated peak performance.

A. FPGA Based Design

FPGA implementation of the proposed has been performed on Xilinx Virtex5 XC5VFX70T FPGA board using Xilinx ISE 13.1 design software. The parameters used to simulate 2D and 3D routers designs are presented in Table 1. Table 2 presents a comparison of the hardware evaluation results for both designs. The results indicate that the 3D router is 1.29 times faster than the 2D router. The estimated peak performance of the 3D router is 1.32 times greater than the 2D router. Compared to the 2D router, the hardware cost and the power consumption are decreased by 61.2% and 61.7% respectively. Thanks to the optimization used in the router pipeline stages design, results confirm an important improvement on the hardware complexity.

TABLE I. SIMULATION PARAMETERS

Router Parameters	2D router	3D Router
Flit size (bit)	32	32
Switching	wormhole	wormhole
Flow control	Credit based	Handshaking
Scheduling	Dynamic arbitre	Dynamic arbitre

TABLE II. COMPARISON BETWEEN 2D AND 3D ROUTERS HARDWARE COMPLEXITIES

Designs	2D router	3D Router
Frequency (MHz)	164	146
Area (Slice)	1200	2800
Power (mW)	13	32
Estimated peak performance (Gbits/s)	53	46

Table 3 shows an overview of the different router implementations. The number of ports of the router depends on its position in the network. Hence, we have four routers of different port numbers: the 7-Port router which is located at the center of the cube and similar to the router R (1,1,1) of Fig. 4. We also have the 6-Port router which is situated at the center of the cube faces and similar to the router R (1,1,0) of Fig. 4. The 5-Port router which is to be found at the middle of the cube edges and similar to the router R (1,0,1) of Fig. 4. The 4-Ports router which is positioned at the vertex of the cube and similar to the router R (0,0,0) of Fig. 4. As can be seen by this table, the maximal clock frequency is decreased when the number of ports per router rises. This decreasing number of frequency is caused by the growth of the arbitre scheduling phase. Other metrics to evaluate the proposal is the estimated peak performance per router which depends on the maximal clock frequency, the flit size and the number of cycles to transmit one flit:

$$PP_{perport} = (Fmax / T) * flit\ size$$

As the maximal clock frequency decreases, the estimated peak performance falls because it is related to the frequency of the design. The hardware resources and the power consumption

are increased when the number of ports per router rises. These increasing numbers can be explained by the growth of the router complexity.

TABLE III. RESULTS OF ROUTERS IMPLEMENTATION ON FPGA

Designs	4-Port Router	5-Port Router	6-Port Router	7-Port Router
Frequency (MHz)	188	164	156	146
Area (Slice)	900	1200	2000	2800
Power (mW)	10	13	23	32
Estimated peak performance (Gbits/s)	60	53	50	46

TABLE IV. PROPOSED ROUTER RESULTS FOR DIFFERENT VIRTEX FPGA DEVICES

FPGA	Vitex-4	Virtex-5	Virtex-6
Topology	3D Mesh	3D Mesh	3D Mesh
Number of ports	7	7	7
Routing algorithm	Negative-first	Negative-first	Negative-first
Frequency (MHz)	93	146	158
Area (Slice)	2101	2764	1355
Estimated Peak Performance per port (Gbits/s)	29	46	50

In order to compare the proposed router with other designs, we implemented it in three different FPGA technologies as illustrated in Table 4. Results of other routers prototyped in FPGA are provided in terms of maximal clock frequency and area as shown in Table 5. The authors of [11] describe router architecture for 3D heterogenous NoC. Their architecture uses adaptive routing and was implemented in Virtex-6 FPGA. Previously in [13] we described router architecture for 3D mesh-based NoC. The design adopts the negative-first turn model routing algorithm and has been implemented in Virtex-5 FPGA. The authors of [16] describe buffer-less router architecture for 3D NoC. Their architecture uses minimal routing and has been implemented in Virtex-4 FPGA. The results demonstrate that area wise, the proposal outperforms all other designs. The proposal is 1.02, 2.82 and 5.49 times smaller than routers of [11], [13] and [16] respectively. The results show that when speaking of maximal clock frequency, the proposal is 1.37 times faster than router of [13]. This also proves that the proposal underperforms the routers of [11] and [16] in terms of maximal clock frequency because we use a disturbed routing and arbitration. However, it allows the router to handle up to seven packets at the same time. The proposal increases the throughput while maintaining an area/speed trade-off. It even gives a better performance when we use advanced FPGA technology like Virtex-7. The estimated clock

frequency reaches 168 MHz and the estimated peak performance extends to 53.76 Gbits/s.

TABLE V. PERFORMANCE COMPARISON OF THE PROPOSED ROUTER WITH OTHER STATE OF THE ART ONES

Design	[11]	[13]	[16]
Topology	3D Mesh	3D Mesh	3D Mesh
Number of ports	7	7	7
Routing algorithm	Adaptive	Negative-first	Minimal
Frequency (MHz)	327	107	250
Area (Slice)	1391	7800	11550
FPGA device	Virtex-6	Virtex-5	Virtex-4

TABLE VI. AREA AND POWER CONSUMPTION OF THE 2D AND THE 3D ROUTERS IN ST FD-SOI 28 NM TECHNOLOGY

Router	Area ( $\mu\text{m}^2$ )	Leak. Power (mW)	Dyn. Power (mW)
3D	5029.82	0.1207	3.8051
2D	2197.81	0.0414	1.5385

### B. ASIC Based Design

In order to evaluate area overhead and power consumption, the 2D and the 3D routers are synthesized by the Synopsys Design Vision tool. This tool uses the FD-SOI 28 nm technology assuming an operating point of 1GHz and a supply voltage of 1V. The resulting area, leakage and dynamic power consumption estimations of each router were extracted from the synthesized circuit and summarized in Table 6. Compared to the 2D router, the 3D router shows a augmentation of about 2.2 times for area, and about 2.53 times for power, due to the additional hardware requirements of the 3D router.

Table 7 illustrates a comparison in terms of area and power consumption of the proposed 3D router with two other routers presented in [17] and [18]. Our choice has been fixed on these routers due to their remarkable performance. The authors of [17] describe robust router architecture that implements an adaptive routing algorithm ensuring fault tolerance both in router components and network links. It also provides high throughput by avoiding deadlocks without any use of virtual channels. The authors of [18] designate router architecture for vertically/partially connected 3D NoC based on stacked 2D Mesh topology. Their router implements Elevator-First routing algorithm that avoids deadlocks by using only two virtual channels in the plane. The results show that in terms of area, the proposed 3D router is 18.27 and 14.33 times smaller than the router of [17] and [18], respectively.

The results indicate that the proposed 3D router is characterized by the best leakage power consumption with a reduction of about 76.3% relatively to [17]. Dynamic power

consumption wise, the proposed router outperforms the 2D router of [17] with a reduction of about 28.8%. The augmentation relative to the 3D router of [18] is only about 15.06%. This upsurge can only be explained by the use of dynamic arbiter that needs more computation than the round-robin arbiter used in [18].

TABLE VII. COMPARISON IN TERMS OF AREA AND POWER CONSUMPTION WITH OTHER ROUTERS

Router	Area ( $\mu\text{m}^2$ )	Leak. Power (mW)	Dyn. Power (mW)
[17]	94189.68	0.511	5.346
[18]	72087.00	-	3.000
<b>3D</b>	5029.82	0.120	3.805

## VI. CONCLUSION

In this dissertation, we adduced the router design for 3D mesh NoC topology which establishes an extension of our former work. We presented 3D NoC router architecture in detail and labelled its pipeline stages optimization, prototyped its architecture on FPGA and synthesized under Synopsys tool using the 28 nm technology. We assessed the proposal performance in terms of maximal clock frequency, area, power consumption and bandwidth and compared it with other famous works. Evaluation results prove that concerning clock frequency, the proposal is 1.37 times faster than our preceding work but it underperforms the routers of [11] and [16] because we use a disturbed routing and arbitration. Hardware cost wise, the proposal is 18.27 and 14.33 times smaller than the router of [17] and [18], respectively. Therefore, it has also been revealed that the best performance is exemplified in a 76.3% reduction of the leakage power consumption with low dynamic power consumption. Additionally, the proposal validates high performance improvement when compared to both the 2D router and 3D router designs thanks to the optimization method. Results have as well verified the capacity of the proposal to handle cost/performance trade-off for 3D NoC.

## REFERENCES

[1] Elhaji, Majdi, et al. "System level modeling methodology of NoC design from UML-MARTE to VHDL." *Design Automation for Embedded Systems* 16.4 (2012): 161-187.

[2] Chemli, Bouraoui, and Abdelkrim Zitouni. "Design and evaluation of optimized router pipeline stages for network on chip." *Image Processing, Applications and Systems (IPAS), 2016 International. IEEE, 2016.*

[3] Tatas, K., Siozios, K., Soudris, D., & Jantsch, A. (2014). *Designing 2D and 3D network-on-chip architectures*. Springer.

[4] Chemli, Bouraoui, and Abdelkrim Zitouni. "Low Cost Network on Chip Router Design for Torus Topology." *IJCSNS* 17.5 (2017): 287.

[5] Azarkhish, Erfan, Igor Loi, and Luca Benini. "A case for three-dimensional stacking of tightly coupled data memories over multi-core clusters using low-latency interconnects." *IET Computers & Digital Techniques* 7, no. 5 (2013): 191-199.

[6] Sheibanyrad, Abbas, Frédéric Petrot, and Axel Jantsch. *3D integration for NoC-based SoC Architectures*. Springer, 2011.

[7] Pasricha, Sudeep, and Yong Zou. "A low overhead fault tolerant routing scheme for 3D Networks-on-Chip." In *Quality Electronic Design (ISQED), 2011 12th International Symposium on*, pp. 1-8. IEEE, 2011.

[8] Akbari, Sara, Ali Shafiee, Mahmoud Fathy, and Reza Berangi. "AFRA: A low cost high performance reliable routing for 3D mesh NoCs." In *2012 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 332-337. IEEE, 2012.

[9] Jain, Arpit, Rakesh Dwivedi, Adesh Kumar, and Sanjeev Sharma. "Scalable Design and Synthesis of 3D Mesh Network on Chip." In *Proceeding of International Conference on Intelligent Communication, Control and Devices*, pp. 661-666. Springer Singapore, 2017.

[10] Khan, Mohammad Yahya, Sapna Tyagi, and Mohammad Ayoub Khan. "Tree-Based 3-D Topology for Network-On-Chip." *World Applied Sciences Journal* 30, no. 7 (2014): 844-851.

[11] Agyeman, Michael Opoku, Ali Ahmadinia, and Alireza Shahrabi. "Efficient routing techniques in heterogeneous 3d networks-on-chip." *Parallel Computing* 39, no. 9 (2013): 389-407.

[12] Salah, Yahia, Yahia Said, Mohsen Ben Jemaa, Salah Dhahri, and Mohamed Atri. "Cost/performance evaluation for a 3D symmetric NoC router." In *Image Processing, Applications and Systems Conference (IPAS), 2014 First International*, pp. 1-7. IEEE, 2014.

[13] Chemli, Bouraoui, and Abdelkrim Zitouni. "A Turn Model Based Router Design for 3D Network on Chip." *World Applied Sciences Journal* 32, no. 8 (2014): 1499-1505.

[14] Chemli, Bouraoui, and Abdelkrim Zitouni. "Design of a Network on Chip router based on turn model." In *Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2015 16th International Conference on*, pp. 85-88. IEEE, 2015.

[15] Glass, Christopher J., and Lionel M. Ni. "Adaptive routing in mesh-connected networks." In *Distributed Computing Systems, 1992., Proceedings of the 12th International Conference on*, pp. 12-19. IEEE, 1992.

[16] Yu, Xiao, Li Li, Yuang Zhang, Hongbing Pan, and Shuzhuan He. "Mass message transmission aware buffer-less packet-circuit switching router for 3D NoC." In *2013 10th IEEE International Conference on Control and Automation (ICCA)*, pp. 983-986. IEEE, 2013.

[17] Alhussien, Abdulaziz, Chifeng Wang, and Nader Bagherzadeh. "Design and evaluation of a high throughput robust router for network-on-chip." *IET computers & digital techniques* 6, no. 3 (2012): 173-179.

[18] Bahmani, Maryam, Abbas Sheibanyrad, Frederic Petrot, Florentine Dubois, and Paolo Durante. "A 3D-NoC router implementation exploiting vertically-partially-connected topologies." In *2012 IEEE Computer Society Annual Symposium on VLSI*, pp. 9-14. IEEE, 2012.

# UHF RFID Reader Antenna using Novel Planar Metamaterial Structure for RFID System

Marwa Zamali, Lotfi Osman, Hedi Ragad  
UR "CSEHF" 13ES37, FST,  
University of Tunis El Manar  
2092 Tunis, Tunisia

Mohamed Latrach  
ESEO Institute of Science & Technology  
49107 Angers,  
France

**Abstract**—An Ultra High Frequency (UHF) half-loop antenna used in Radio Frequency Identification (RFID) systems is proposed with a planar patterned metamaterial structure of compact size. The size of the planar patterned metamaterial structure is  $(0.20\lambda \times 0.20\lambda \times 0.0023\lambda)$  mm<sup>3</sup>. This antenna consists of two metamaterial unit cells having negative permittivity and permeability. Simulation results of input return loss, radiation pattern, and directivity of this antenna are presented using CST software. A comparison between the conventional antenna and the new metamaterial half-loop antenna is also provided. The simulated results show that the metamaterial antenna has a resonance frequency of 0.866 GHz, a realized gain of 1.96 dB, and an efficiency increase of about 20%. Simulation and measurement results are in perfect agreement, which proves that the proposed antenna can operate in the UHF band for RFID systems.

**Keywords**—Loop antenna; metamaterial; miniaturization; RFID system; UHF band

## I. INTRODUCTION

Nowadays, antennas play a very important role in tracking any object using the automatic identification technology. This technique has been increased in many fields like industry, business, food and library management, and so on [1]. In wireless communications, the RFID systems are one of the most promising technologies [2] that emerged after automatic identification to detect any object anywhere [3], [4]. They are now widespread in many industrial applications such as merchant flow tracking, stock, control, and supply chain management [5], [6].

The UHF RFID systems are used in different regions of the world. We will be interested in the European UHF band which is between 865MHz and 868 MHz [7], [8]. The drawback of this band is the need of a large antenna, which led researchers to look for a solution minimizing the size of the antenna while keeping its performance. There are several methods of reducing the size of the antenna such as the use of short circuits (PIFA), slot antennas, and fractal shapes [9]-[12]. However, these techniques do not allow having a good performance in terms of impedance, directivity, and large wide band, hence the interest in the use of metamaterials [13], [14] to decrease the resonance frequency and maintain the performance of the antenna. Metamaterials, which were propounded by Victor Veselago in 1968, are also called left-handed materials (LHM) because the vectors E, H, and K form a left-handed system [15], [16].

This work treats a half-loop antenna coupled with a novel structure of metamaterial to decrease the resonance frequency by 20% and increase the gain by 76%. The obtained results are shown by a numerical software using CST-MWS and validated by experimental results. The presented paper falls into three parts. In the first section, the physical and design for the structure of metamaterial are presented and simulated regarding the S11 and S12 to find the permeability and permittivity. The second section shows the simulation results for the proposed antenna presenting the return loss and the realized gain, that's why in the thirist section; the antenna is coupled with metamaterial to achieve the desired results in terms of the resonance frequency, realized gain and radiation pattern. In the last section, the metamaterial is validated by the experimental result.

## II. PRESENTATION OF THE NEW METAMATERIAL STRUCTURE

The concept of artificial material, i.e. metamaterial, was synthesized in the late 90s. However, in the past 20 years, the interest in metamaterial technology has strongly increased with researchers on superlens and telecommunication environments, including transmission lines and antenna applications. Metamaterials are artificial materials obtained by regulating the interactions of materials with electromagnetic (EM) waves to make them acquire specific properties [17]. The photonic metamaterial patterns to be incorporated on the antenna were made through 4-level fractal structures, which are sometimes called a space filling curves. The fractal pattern is generated by a master line or the first level of the structure. The multi-band functionality and sub-wavelength effect are the two most important features of this S and U-shaped fractal. The main advantages of the fractal cell are its sub-wavelength properties, simple architecture, and wide application. The sub-wavelength allows the system to have a smaller size when compared with the wavelength. Thus, the cell can behave as a compact reflector. In the case of metamaterials, the curl of electric field E and magnetic field H have a power density  $S=E \times H$  and moves in the opposite direction of the wave vector k. The most important two parameters of electromagnetic waves are Electric Permittivity  $\epsilon$  and magnetic permeability  $\mu$ . Fig. 1 shows the geometry of the fractal unitary metamaterial cell. In order to calculate the effective values of permittivity and permeability, the Nicolson-Ross-Weir (NRW) approach is used [18]. The equations are as follows:

$$n = \frac{1}{kd} \cos^{-1} \left[ \frac{1}{2S_{21}} (1 - S_{11}^2 + S_{21}^2) \right] \quad (1)$$

$$Z = \sqrt{\frac{(1+S_{11})^2 - S_{21}^2}{(1-S_{11})^2}} \quad (2)$$

Where, the S parameters result from CST simulation, n and z are related to the relative permittivity and permeability.

$$\epsilon = \frac{n}{Z}, \mu = \frac{n}{Z} \quad (3)$$

Table 1 below presents the dimensions of the proposed metamaterial cell.

Fig. 2 shows the S-parameters of the new metamaterial structure whose band gap ranges from 0.855 to 1.12 GHz.

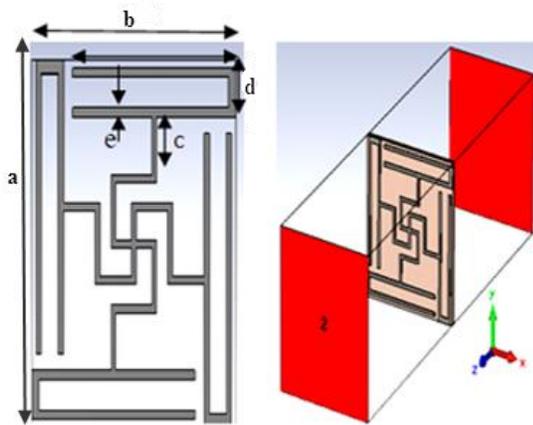


Fig. 1. Geometry of the proposed metamaterial cell.

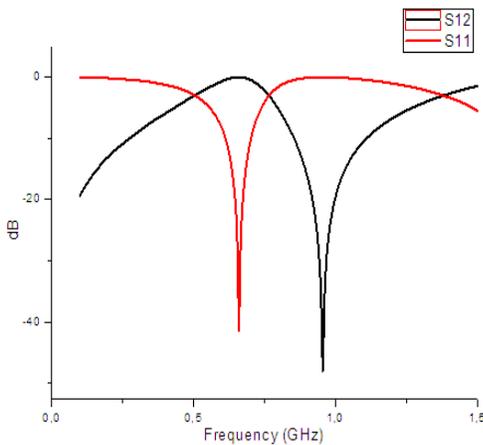


Fig. 2. Parameters of the novel metamaterial structure.

TABLE I. DIMENSIONS OF THE PROPOSED METAMATERIAL CELL

a (mm)	b (mm)	c (mm)	d (mm)	e (mm)
28	22	5	3.75	0.5

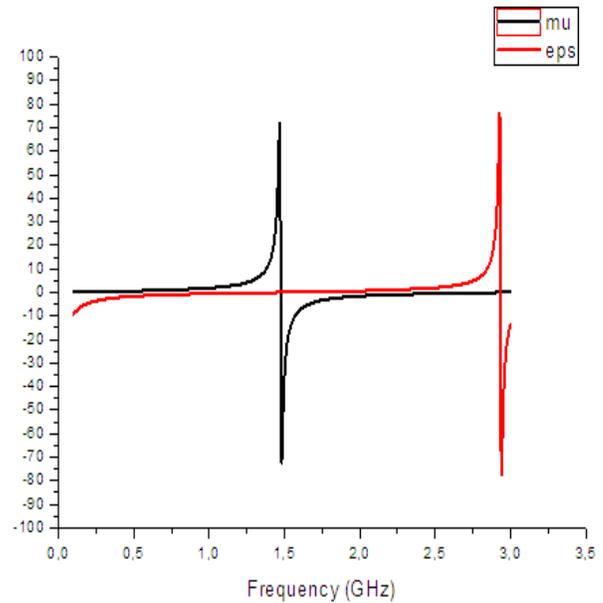


Fig. 3. Real part of the relative permittivity and the permeability.

Fig. 3 presents the real part of the relative permittivity and the permeability. It may be noted that the real part of the permittivity is negative around the resonance frequency.

### III. DESIGN OF THE HALF-LOOP ANTENNA

A half-loop antenna is a closed circuit antenna that consists of a single wire bent into a half circle and mounted on a ground plane. The conductor is fed through the ground plane at one end while the other end is terminated on the ground plane. The half-loop antenna exists in various forms, but in this paper, the circular form will be used. The small size and high efficiency are the advantages of a properly designed half-loop antenna constructed on the UHF frequency band.

#### A. Half Loop Antenna Without Metamaterial

Fig. 4 exhibits the basic structure of the half-loop antenna without metamaterial cells. It consists of a circular half-loop of 30 mm radius R resonating at 1.05 GHz printed on an FR4 substrate ( $\epsilon_r=4.4$ ,  $h=0.8$  mm and  $\tan \delta=0.0008$ ). The simulated frequency and the realized gain are shown in Fig. 5 and 6, respectively.

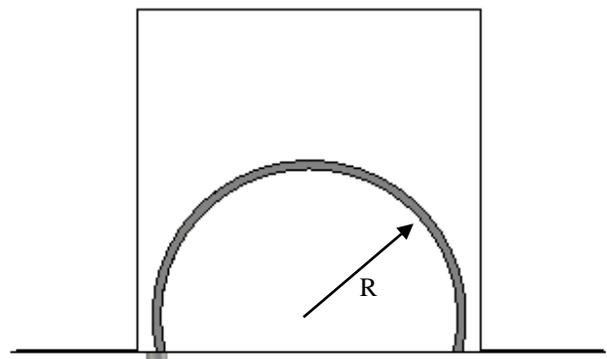


Fig. 4. Design of the proposed antenna.

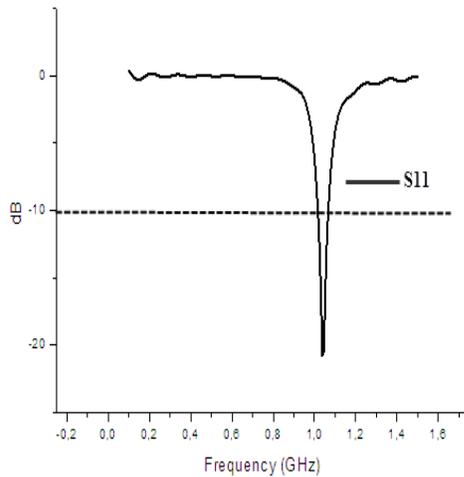


Fig. 5. Reflection coefficient of the antenna without metamaterial.

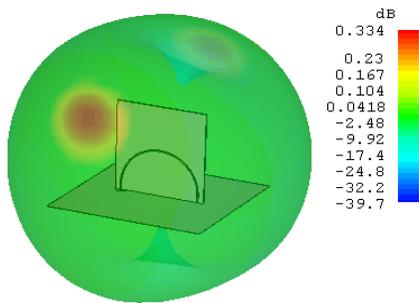


Fig. 6. Realized gain at 0.866 GHz of the antenna without metamaterial.

One important property of antenna is the efficiency. It can be written as the ratio of the total radiated power  $P_r$  to the input power  $P_i$  of the antenna.

$$\eta = \frac{P_r}{P_i} = \frac{Gain}{Directivity} \times 100 \quad (4)$$

The realized gain of the antenna without metamaterial is 0.33 dB, likewise, the directivity is about 2.11 dBi. So, the efficiency of the antenna without metamaterial is 68%.

### B. Half Loop Antenna with Metamaterial

An important fraction of the radiation gets lost in the substrate above the radiant element and on the edges around the structure. This problem is occurred by surface waves and leakage waves which limit the performance of the antenna. To overcome this limitation, we use metamaterial cells to prevent the distribution of the surface waves. The addition of the metamaterial cells serves to reduce the resonance frequency, hence miniaturizing the structure. Adding metamaterial cells in the half-loop antenna modifies the radiation pattern by focusing the energy in one direction [19], [20] and it is completely confined under the radiating element.

The proposed antenna shown in Fig. 7 has a loop radius ( $r$ ) of 30 mm and a width ( $w$ ) of 1.25 mm. This half-loop is integrated on an FR4 substrate of 70mm \*70mm \*0.8mm (the same substrate of the metamaterial cells) that is perpendicular to a ground plane of 120mm \*120mm \*0.1 mm, and is simulated through the use of the CST software.

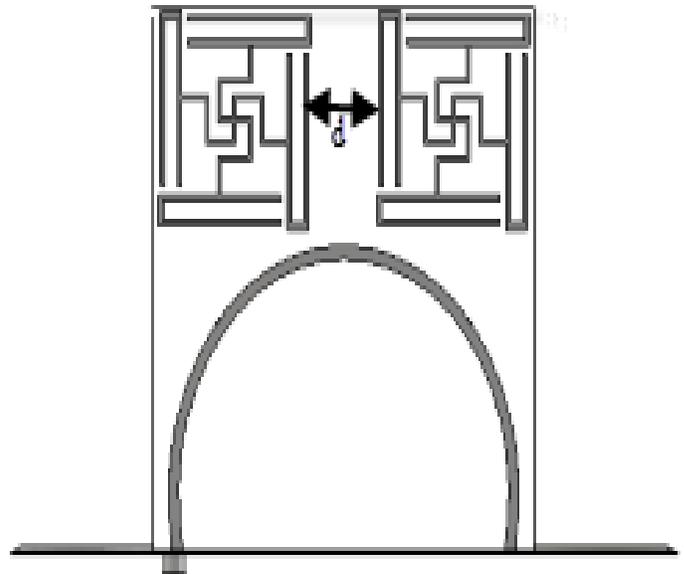


Fig. 7. Design of the metamaterial antenna.

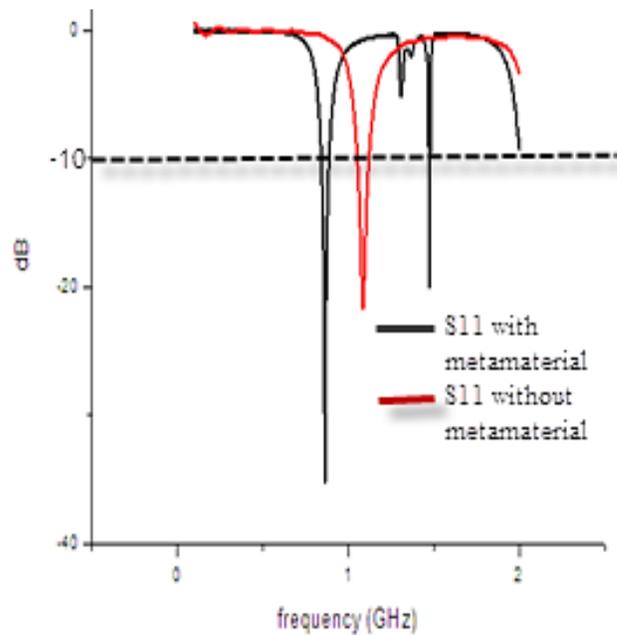


Fig. 8. Reflection coefficient of the antenna with and without metamaterial.

The proposed half-loop antenna with metamaterial cells exhibited better results in terms of return loss, with -36 dB against -22 dB, and diminution of resonance frequency from 1.05 GHz to 0.866 GHz. The S11 parameters for the two designs are shown in Fig. 8.

This figure shows that the use of metamaterials cells allows the decrease of the resonance frequency from 1.05 GHz to 0.866 GHz, i.e. a decrease of about 185 MHz. To understand the performance of distance between the two metamaterial cells, we examined the effects on return loss when varying parameter  $d$ . The simulation of the resonance frequencies for different values of  $d$  is presented in Fig. 9.

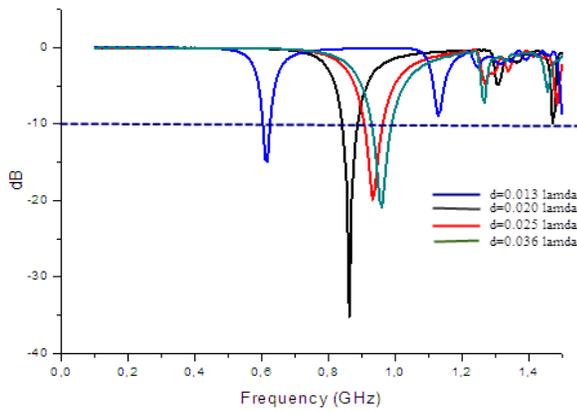


Fig. 9. Effects of the distance d on the return loss.

Fig. 9 shows the effects of distance d on the resonance frequency, we can notice that the resonance frequency is decreased when the distance between both cells is decreased. However, the gain at the low resonance frequency ( $f_r = 0.61$  GHz,  $d = 0.036 \lambda_0$ ) is negative, on which the antenna has been chosen whose distance separating the two cells is approximately 8 mm ( $d = 0.020 \lambda_0$ ). The novel antenna exhibits better results in terms of gain, with a simulated gain of about 1.93 dB with against 0.44 dB without metamaterials.

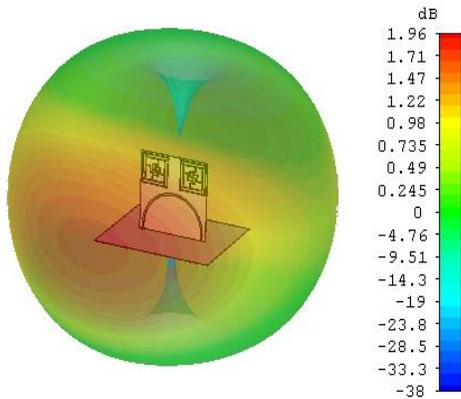


Fig. 10. Realized gain at 0.866 GHz of the metamaterial antenna.

Fig. 10 shows that the realized gain at 866MHz of the metamaterial antenna is about 1.93 dB. The simulated directivity of metamaterial antenna is 2.11 dBi. Theoretically [21], the maximum gain of aperture antenna is expressed as:

$$D_{max} = \frac{4\pi A}{\lambda_0^2} \quad (5)$$

Where,  $\begin{cases} A = L^2 * m \\ \lambda_0 = \frac{c_0}{f_0} \end{cases}$

L is the length of the ground plane, m is the length of the substrate and A is the aperture of the antenna. So  $D_{max} = 2.73$  dBi. The simulated directivity of metamaterial antenna has almost approached the theoretically limit of the antenna's directivity with percentage error (5%).

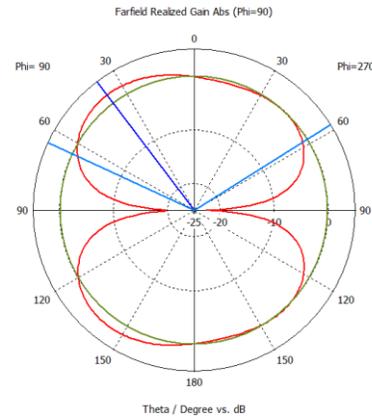


Fig. 11. Simulated radiation pattern in the E-plane at 0.866 GHz.

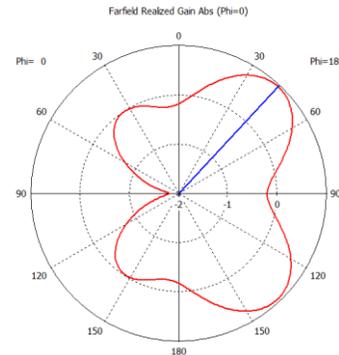


Fig. 12. Simulated radiation pattern in the H-plane at 0.866 GHz.

Fig. 11 and 12 show the simulated radiation pattern in plane E and H, respectively. We can notice that the simulated 3dB beam width for this antenna with metamaterial is decreased by  $57^\circ$  at plane E which implies an increase in directivity. The efficiency of the antenna with metamaterial is about 87%, that's why it may be concluded that thanks to the metamaterials that efficiency has improved.

Table 2 summarizes the characteristic values of the antenna with and without metamaterial in terms of resonance frequency, frequencies range, and bandwidth which are calculated by the following equation:

$$BW\% = \frac{2(f_{max} - f_{min})}{(f_{max} + f_{min})} \times 100 \quad (6)$$

TABLE II. CHARACTERISTICS OF THE ANTENNA WITHOUT AND WITH METAMATERIAL

	Without metamaterial	With metamaterial
$f_r$ (GHz)	1.04	0.866
RL (dB)	21	36
$f_{min}$ (GHz)	1	0.863
$f_{max}$ (GHz)	1.1	0.869
BW (%)	4.7	0.34
Gain (dB)	0.33	1.96
Directivity(dB)	2.11	2.59
Efficiency (%)	68	87

#### IV. SIMULATION AND EXPERIMENTAL RESULTS

The novel antenna with and without metamaterial cells are realized in the Maxwell Laboratory of group ESEO-Angers, France. Fig. 13 illustrates the experimental prototypes whose ground planes have the dimensions of 12cm×12cm.

Simulated and measured S11 parameter of the antenna with and without metamaterial cells are presented in Fig. 14.



Fig. 13. Realized metamaterial antenna.

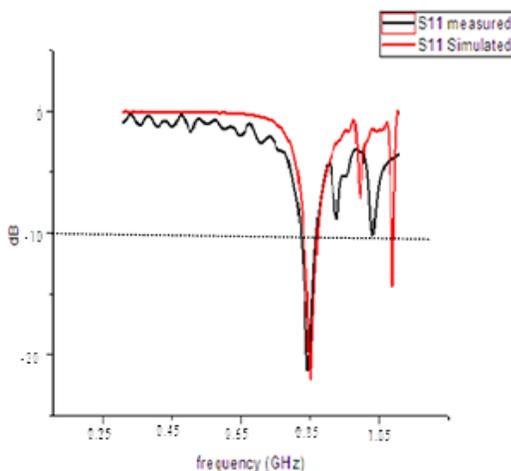


Fig. 14. Simulated and measured |S11| of the metamaterial antenna.

#### V. CONCLUSION

In this paper, a novel metamaterial cell is designed with negative permittivity and permeability. When metamaterial cells are applied to the antenna, this leads to an improvement in the gain and the radiation patterns with a decrease in the resonant frequency. Measurements on the manufactured prototype validated the simulation results carried out under CST software. As perspective to the study, the proposed antenna could be used in libraries for identification, sorting and transportation of library books.

#### ACKNOWLEDGMENT

The authors would like to thank the ESEO (Ecole Supérieure d'Electronique de l'Ouest) of Angers-France for

their support and assistance during the preparation of this research paper.

#### REFERENCES

- [1] Priyanka Grover, Anshul Ahuja, "Radio Frequency Identification based library management system," International Journal of Advanced Computer Science and Applications, Vol. 1, No. 1, July 2010.
- [2] K. Finkenzeller, "Fundamentals and Applications in Contactless Smart Cards and Identification," RFID Handbook, New York: Wiley, 2003.
- [3] Daniel W. Engels, Tom A. Scharfeld, Sanjay E. Sarma, "Review of RFID Technologies," Cambridge, MA USA 02139.
- [4] Arttu Lehto, Jussi Nummela, Leena Ukkonen, Lauri Sydanheimo and Markku Kivikoski, "Passive UHF RFID in Paper Industry: Challenges, Benefits, and the Application Environment," IEEE Transactions on Automation Science and Eng., vol. 6, no. 1, pp. 66-79, Jan. 2009.
- [5] K. Finkenzeller, "Radio-frequency identification fundamentals and applications," RFID Handbook: 2nd edition, Wiley, West Sussex, England, 2004.
- [6] J. Sabri, N. Omrane, T. Hichem, M. Abdelkader, G. Ali, "Miniaturized meander slot for RFID tag with dielectric resonator at 60GHz," International Journal of Advanced Computer Science and Applications, vol. 7, no. 4, 2016.
- [7] K. Finkenzeller, RFID Handbook: Radio-Frequency Identification Fundamentals and Applications. New York: Wiley, 2002.
- [8] V. D. Hunt, A. Puglia, and M. Puglia, "RFID-A Guide to Radio Frequency Identification" Hoboken, NJ: Wiley, 2007.
- [9] S. Pflaum, G. Kossiavas and R. Staraj, "Circularly polarized wire patch antenna for RFID applications," 6th European Conference on Antennas and Propagation (EUCAP), pp. 2997-3000, 2012.
- [10] K. L. Wong, and C.C. Huang, and W.S. Chen, "Printed Ring slot antenna for circular Polarization Propagation," AP-50, 1, pp.75-77, January 2002.
- [11] Marwa Zamali, Lotfi Osman, Hedi Ragad and Mohamed Latrach, "A Novel Miniaturized Circularly Polarized PIFA Antenna for RFID Readers," IEEE 15th Mediterranean Microwave Symposium (MMS), Lecce, Italy, Nov. 30 2015-Dec. 2 2015.
- [12] J.Y. Sze, and C.C. Huang, "Coplanar wave guide- Fed square slot antenna for Broad band circular Polarization Radiation," IEEE transac. on antennas and propagation, vol. 51, no. 8, pp.2141-2144, Aug. 2003.
- [13] Zeev Iluz, Reuven Shavit, and Reuven Bauer, "Microstrip Antenna Phased Array with Electromagnetic Band gap Substrate," IEEE Trans. on Antennas and Propag., vol. 52, No. 6, pp. 1446-1453, June 2004.
- [14] Fan Yang and Yahiya Rahmat-Samii, "Microstrip Antennas Integrated With Electromagnetic Band-Gap (EBG) Structures: A Low Mutual Coupling Design for Array Applications," IEEE Transactions on Antennas and Propagation. vol. 51, no.10, pp.2936-2946, Oct. 2003.
- [15] B.zarghooni, A. Dadgarpour, and T. A. Denidi, "Reconfigurable planar metamaterial unit cell," Phil. IET Microwaves Antenna & propagation, vol. 9, no. 10, pp. 1074-1079, March 2015.
- [16] Ludovic Leger, Thierry Monediere and Bernard Jecko, "Enhancement of gain and radiation bandwidth for a planar 1-D EBG Antenna," IEEE Antennas Wireless Comp. Lett., vol.15, no.9, pp.573-575, Sept. 2005.
- [17] S. A. Ramakrishna, "Physics of negative refractive index materials," Rep. Prog. Phys., 2005, 68, 449-521.
- [18] Richard R. Ziolkowski, "Double Negative Metamaterial Design, Experiments and Application," IEEE on Microwave Theory and Techniques, vol. antenna's miniaturization and high directivity. 51, No. 7, July 2003.
- [19] S.-Y. Chen, R.O. Ouedraogo, A. Temme, A.R. Diaz, E.J. Rothwell, "MNG- Metamaterial- Based Efficient Small Loop Antenna," IEEE APSURSI 2009, pp. 1-4.
- [20] E. Ramanandraibe, M. Latrach, and A. Sharaiha, "A half loop antenna associated with one SRR cell," ICEAA 2013, Italy, September 2013.
- [21] F. Zhu, Qingchun Lin and Jun Hu, "A Directive Patch Antenna with a Metamaterial Cover," Asia-Pacific Microw. Conf. Proc., vol. 1, 2005.

# Factors Influencing Users' Intentions to Use Mobile Government Applications in Saudi Arabia: TAM Applicability

Raed Alotaibi<sup>1,2</sup>, Luke Houghton<sup>1</sup> & Kuldeep Sandhu<sup>1</sup>

<sup>1</sup>Griffith business school, Griffith University  
Brisbane, Australia

<sup>2</sup>College of community, Shaqra University  
Shaqra, Saudi Arabia

**Abstract**—M-government applications in Saudi Arabia are still at an early stage. In this study, a modified technology acceptance model (TAM) was used to identify and measure the factors that influence users' intentions to use m-government applications in Saudi Arabia. This study focuses on the relationships between behavioural intention to use (BIU) and six independent factors: three TAM constructs (perceived usefulness [PU], attitude towards use [ATU], and perceived ease of use [PEU]) and three external factors: perceived trustworthiness [TRU], perceived security [SEC], and awareness [AWAR]). Only PU, ATU and TRU had a significant positive influence on BIU for m-government applications. The results also showed that most participants had a positive attitude towards using m-government applications. Overall, the results demonstrate that the model is suitable in the Saudi m-government context.

**Keywords**—TAM; Saudi Arabia; e-government; m-government applications

## I. INTRODUCTION

In 2005, the Saudi Government established Yesser to provide e-government services with the overall aim of making all government services ready for use by citizens anywhere and at any time by 2010 [1]. Some studies have asserted that e-government initiatives in Saudi Arabia have been delayed, although there has been some progress in implementation [2, 3]. A recent study [4] stated that in developing countries like Saudi Arabia, the adoption of e-government services is poor.

M-government has been defined as the government providing services and information via mobile devices and wireless communication networks such as cellular phones, PDAs and their supporting systems to businesses, citizens, public employees and non-profit organizations [5]. Many studies claim that m-government is like e-government and it is considered a complementary subset [6]-[12]. It has been claimed that in Saudi Arabia, mobile phone use has seen massive growth among citizens [13]. In 2015 there were 53 million mobile subscriptions in Saudi Arabia, representing about 167.5% of the population almost two mobile subscriptions per person [14]. A recent study [15] claimed that most (67%) of the Saudi population use a smart phone. Recent studies [16], [17] claimed that in Saudi Arabia, traditional methods

for

transacting with government sectors are still very common and are preferred by citizens. The studies also asserted that Saudi citizens are yet to completely adopt m-government. Therefore, the Saudi Government can exploit this high use of mobile devices to provide government services via applications. Few m-government applications, however, have actually been adopted in Saudi Arabia. In fact, the use of m-government applications in Saudi Arabia is still in its infancy [18]-[21].

M-government provides government services to users via mobile technology to provide quick service, increase user mobility and enable easy access to services. It is considered a sophisticated way to provide government services to users through mobile devices [6]. It has been suggested that there are four types of m-government: 1) m-government-to-government (mG2G); 2) m-government-to-citizen (mG2C); 3) m-government-to-employee (mG2E); and 4) m-government-to-business (mG2B) [12]. This study focuses on mG2C to identify and measure the factors that influence users' intentions to use m-government applications, as it has been claimed that most interaction today is m-government-to-citizen [22]. Therefore, this study aims to measure some factors, found in literature review, that influence the intentions of Saudis to use m-government applications.

This study's contribution to knowledge is in understanding Saudi citizens' attitudes towards using m-government applications and the factors that influence users' intentions to use those applications. This information will help government decision makers involved in e-government and m-government initiatives. Acknowledging and addressing these factors will support future m-government applications and their implementation. Theoretically, this study examines and evaluates the applicability of a modified technology acceptance model (TAM) in the Saudi m-government context. More specifically, this study will evaluate the TAM's applicability by examining relationships between behavioural intention to use (BIU) and six independent factors: three TAM constructs (perceived usefulness [PU], attitude towards use [ATU] and perceived ease of use [PEU]) and three external factors (perceived trustworthiness [TRU], perceived security [SEC], and awareness [AWAR]). In this way we envisage a practical contribution by government decision makers and a theoretical contribution by TAM researchers.

To conclude this study, there are three aims. They are as follows:

- 1) to measure the influence of external factors (TRU, SEC and AWAR) and TAM constructs (PU, ATU and PEU) on users' intentions to use m-government applications in the Saudi context;
- 2) to measure Saudi citizens' attitudes towards using m-government applications; and
- 3) to propose, examine, develop and validate a model of m-government that suits the Saudi cultural context.

## II. LITERATURE REVIEW

### A. Relationship Between E-Government and M-Government

M-government is considered another method to provide government services to citizens [23]. Many studies report that m-government and e-government are essentially the same thing, but m-government can be considered a sophisticated type of e-government [24], [25]. Some studies consider e-government to be a key component of m-government [26]-[29]. Another study claims that because m-government is an independent method of achieving government objectives through the provision of services and information, it is similar to e-government [30].

### B. Advantages, Goals and Objectives of M-Government

M-government has the advantages of easy infrastructure setup, improved e-government efforts, ease of learning, inclusiveness, remote area access and low cost [12], [19], [30], [31]. It has been stated that the advantages of m-government include: the provision of location-based government services, on-time information delivery, mobility, ubiquity, time savings, ease of use and improved emergency management [32]. A recent study [25] confirmed all the m-government advantages noted in previous studies, but further highlighted two new advantages: international trade benefits and democratic reforms. Althunibat, Alrawashdeh, and Muhairat [6] claimed that the goal of m-government is to attract users to use government services, since m-government is easily accessible for services 24 hours a day, seven days a week. Along the same lines, it has been demonstrated that m-government improves connection methods between citizens and the government, encourages citizens to participate in local community matters, provides e-government services with certain additional features (e.g. timeliness, convenience, etc.) and implements and provides e-government services to citizens in distant locations for which the government has not previously provided services, such as rural areas [33].

### C. M-Government in Saudi Arabia

M-government applications are already used to provide government services in most countries but most of these applications provide limited services, such as tracking systems for stolen vehicles, emergency assistance and weather updates [18]. In Saudi Arabia some services are already provided by m-government, such as Riyadh and Madinah Education, Appointments and Document Tracking, Health Mobile, tracking of Higher Education Information, and Employee Inquiry [34]. Although some applications in Saudi Arabia have been provided, m-government applications are still considered

as being in the first stage of implementation [18]-[21]. Accordingly, researchers seek to identify and measure the factors that influence users' intentions to use m-government applications.

There is a lack of empirical studies that examine and validate the applicability of TAM in the Saudi m-government context. There is also a lack of empirical studies identifying and measuring factors that influence users' intentions to use m-government applications in Saudi Arabia.

### D. Theoretical Framework

TAM is derived from the theory of reasoned action (TRA) that states that beliefs impact on intentions, and intentions impact on actions [35]. There are differences between TAM and TRA. TAM is used for the adoption of new technologies and focuses on users' attitudes, behaviours and perceptions in the adoption of this technology, while TRA is more general and used for various types of cases and adopted behaviours according to perceived positive outcomes [36].

TAM was first introduced in 1986 to measure new technology acceptance by users [37]. As shown in Fig. 1, TAM is expected to measure user acceptance of new technologies based on various factors, including behavioural intention to use, attitude towards use, perceived usefulness and perceived ease of use [38]. According to Davis [39], perceived usefulness can be defined as "the degree to which a person believes that using a particular system would enhance his or her job performance", and perceived ease of use can be defined as "the degree to which a person believes that using a particular system would be free of effort". Behavioural intention is defined as "the strength of one's intention to perform a specified behaviour" [40]. Attitudes can be defined as "an individual's positive or negative feelings (evaluative affect) about performing the target behaviour" [41].

TAM was selected as the research model in this study because it can determine the influence of external factors on attitude, intention to use, and belief [42]. As this study aims to study factors that influence behavioural intention to use (BIU) of m-government applications, TAM was chosen as the best model to predict user behaviour towards new technologies [42] because most other technology acceptance models only focus on technical factors [38]. According to Al-Hujran, Al-dalahmeh, and Aloudat [40], TAM is used to predict and explain users' acceptance of new technologies, proposing that constructs like perceived usefulness and perceived ease of use are key factors in information system (IS) and information technology (IT) acceptance behaviours.

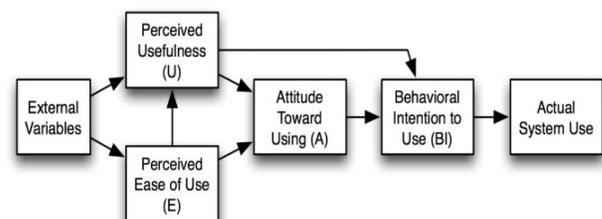


Fig. 1. The Technology Acceptance Model (TAM) [39].

TAM is one of the most effective research models to determine factor acceptance of information systems/technology [43]. TAM is also commonly used to measure user behaviour and to identify the factors influencing users' acceptance of new technologies [38], [44]-[46].

#### E. Research Model and Hypotheses

The original basic TAM was modified to achieve the study's objectives by including three external factors. Therefore, this study focuses on the relationships between behavioural intention to use (BIU) and six independent factors, either TAM constructs (perceived usefulness [PU], attitude towards use [ATU], and perceived ease of use [PEU]) or external factors (perceived trustworthiness [TRU], perceived security [SEC], and awareness [AWAR]) to identify and measure the factors that influence users' intentions to use m-government applications in the Saudi context. The research model for this study is presented in Fig. 2.

1) *Hypotheses for relationships between TAM constructs and BIU*: Based on the original TAM results [39], BIU is positively influenced by PU and ATU. In this study BIU is defined as a user who has a strong intention to use m-government applications. Positive ATU is defined as a user who wants to use m-government applications. PU is defined as a user who thinks that his or her job will be more productive and efficient by using m-government applications. In addition, BIU of e-learning is positively influenced by PEU [47]. In this study, PEU is defined as the degree to which a user believes that using m-government applications does not need much effort. Consequently, the hypotheses regarding TAM constructs in this study are:

H1: Perceived usefulness (PU) will have a positive significant influence on behavioural intention to use (BIU) m-government applications.

H2: Attitude towards use (ATU) will have a positive significant influence on behavioural intention to use (BIU) m-government applications.

H3: Perceived ease of use (PEU) will have a positive significant influence on behavioural intention to use (BIU) m-government applications.

2) *Hypotheses for relationships between the external factors and BIU*: Trust plays a key role in new technology adoption because it impacts on users' intentions [48], [49]. Trust also has a strong positive influence on the adoption of m-

commerce [50]. In addition, trust has a positive effect on users' intentions to use m-government services [20]. Consequently, the following is hypothesised:

H4: Perceived trustworthiness (TRU) will have a positive significant influence on behavioural intention to use (BIU) m-government applications.

In this study, perceived security includes perceived privacy. It has been claimed that security is very important in new technology adoption because users do not use these technologies if they are not perceived to be safe [51]. A previous study [52] found that security and privacy play a key role in m-government adoption. Security also has a positive impact on intention to use tourism m-payment systems [53] and is a key factor in the use of mobile banking [54]. Consequently, the following is hypothesised:

H5: Perceived security (SEC) will have a positive significant influence on behavioural intention to use (BIU) m-government applications.

In Jordan, a lack of awareness among citizens prevented e-government adoption, resulting in the need for increased citizen awareness [55]. Awareness is also an important factor in new technology adoption, especially m-government adoption [33]. A recent study [56] found that a positive relationship between awareness and internet banking adoption. Another study [57] showed that awareness had a positive influence on the intention to adopt e-government in Bahrain. Consequently, the following is hypothesised:

H6: Awareness (AWAR) will have a significant positive influence on behavioural intention to use (BIU) m-government applications.

#### F. Summary

This study will address three existing problems:

1) To date, there have been insufficient empirical studies to identify and measure the factors that influence users' intentions to use m-government applications in Saudi Arabia.

2) To date, there have been insufficient empirical studies to measure Saudi citizens' attitudes towards using m-government applications.

3) To date, there have been insufficient empirical studies to validate and examine the applicability of TAM, with some modifications, in the Saudi m-government context.

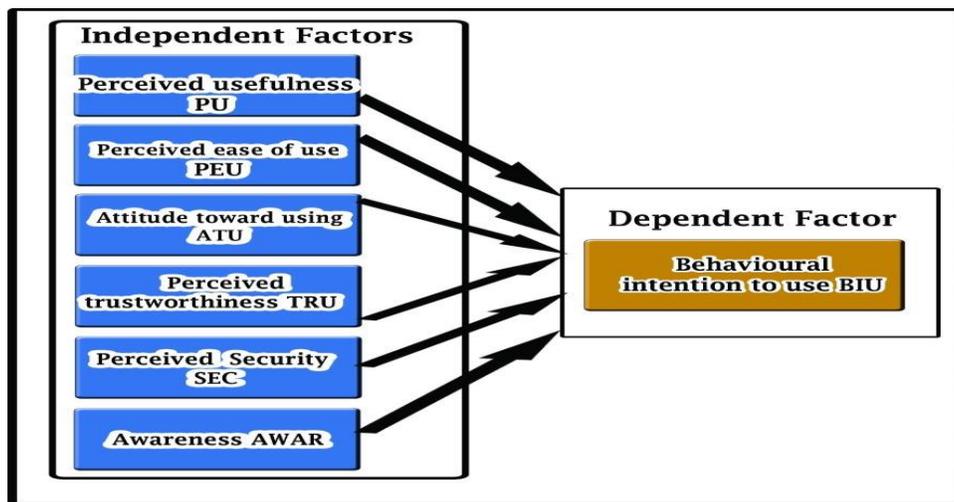


Fig. 2. Research model for this study.

### III. RESEARCH METHOD

This study used quantitative methods to determine the impacts of the independent factors on BIU of m-government applications in Saudi Arabia. The researcher used a self-administered questionnaire to collect data from participants. Researchers often use questionnaires to identify factors that influence the adoption and acceptance of new technology, taking advantage of the low cost and limited time required [58]. Questionnaires also improve ease of access to participants and increase the likelihood of receiving truthful responses [33].

#### A. Questionnaire

Balnaves and Caputi [59] warned that words used in questionnaires can affect results; thus, researchers should be careful about their wording. In particular, they should use easy and clear language to ensure participants' understanding [40]. The study questionnaire was validated by its use and adoption in previous studies [20], [38], [39], [60]-[62] and was modified to suit the m-government context. To ensure clarity and avoid errors, the questionnaire was first tested by English language experts. As the respondents were native Arabic speakers, they received an Arabic copy of the questionnaire, which was sent to expert translators to ensure accuracy and clarity of the responses.

#### B. Instrument

The questionnaire consisted of four parts. Part one focussed on demographic characteristics. Parts two and three listed items related to TAM model constructs and external factors. Parts two and three also sought to measure participants' attitudes toward using m-government applications in Saudi Arabia and to measure the factors that influence users' intentions to use m-government applications. Finally, part four allowed participants to make their own comments and offered research results to those who left an email address. All TAM constructs and external factors were measured at seven levels on a Likert-type scale. Participants were asked to choose from 7 = Strongly Agree, 6 = Moderately Agree, 5 = Slightly Agree, 4 = Neutral, 3 = Slightly Disagree, 2 = Moderately Disagree and 1 = strongly Disagree.

#### C. Data Collection

In this study, the target population was Saudi citizens (male and female) who have internet access. This study used probability sampling (random sampling), which is defined as a technique in which "there is an equal chance to be selected and participate between units (e.g. persons, cases) in the population and that selections are made independently" [63]. However, given that the number of Saudi citizens was more than 20 million in 2015 [64], it was difficult to give all Saudi citizens an equal chance to participate in this study due to time limitations and associated costs. Consequently, researchers focused on Saudi citizens using email groups, Twitter, social networks and Facebook, and distributed an online link to a questionnaire to these users. This approach was used to obtain the minimum sample size required for this study, which was calculated to be 384 participants.

This allowed the researcher to collect data from a large population while providing clear instructions on how to complete the questionnaire. To improve the response rate, the researcher adopted some techniques recommended in [65], in particular, that the questionnaire is as brief as possible, the first page of the questionnaire introduces the researchers and explains the significance of the study and its objectives. Each week, the researcher sent follow-up reminders about the questionnaire to participants through email and social networks.

More than 1600 questionnaires were distributed and 1,152 (72%) of the 1600 questionnaires were returned. Of those, 370 (23%) questionnaires were incomplete and 782 (48%) included full responses. The full response rate of 48% was considered adequate because a 30% response rate is considered acceptable when analysing surveys [66].

#### D. Ethical Considerations

The study was conducted anonymously, transparently, and with the promise of no harm. The participants had the right of easy access to research results and the choice (non-compulsory) of whether to participate. They were also under no obligation to fully complete the questionnaire. Institutional ethical approval was provided by Griffith University.

#### IV. DATA ANALYSIS

##### A. Demographics

As shown in Table 1, the vast majority of participants who completed this questionnaire were male 613 (78.4 %) compared to 169 (21.6%) females.

TABLE. I. DEMOGRAPHIC INFORMATION FOR PARTICIPANTS WHO FULLY COMPLETED THE QUESTIONNAIRE

Information		Number of participants	Percentage of sample
Gender	Male	613	78.4%
	Female	169	21.6%
	Total number of participants	782	100%
Age	< 30	274	35%
	30 < 40	362	46.3%
	40 < 50	112	14.3%
	50 < 60	30	3.8%
	60 or more	4	0.5%
	Total number of participants	782	100%
Where participants resided	City	638	81.6%
	Village	144	18.4%
	Total number of participants	782	100%
Qualifications	High School	105	13.4%
	Diploma	107	13.7%
	Bachelor	375	48%
	Postgraduate	187	23.9%
	Other	8	1%
Total number of participants	782	100%	
Occupation	Students	97	12.4%
	Working in a government organization	500	63.9%
	Working in the private sector	98	12.5%
	Businessperson	18	2.3%
	Other	69	8.8%
	Total number of participants	782	100%
Experience with m-government applications	Have not used m-government applications	187	23.9%
	Less than one year	157	20.1%
	More than one year/less than two years	125	16%
	More than two years/less than three years	137	17.5%
	Three years or more	176	22.5%
	Total number of participants	782	100%

Nearly half of the respondents were aged 30–39 (46.3%, 362). In contrast, the smallest age group of respondents was those aged 60 or more (0.5%, 4).

Most participants (81.6%, 638) lived in cities. More than 70% (562) had university qualifications while almost all others had attended high school (13.4%, 105) or had completed a diploma (13.7%, 107).

Most (64%, 500) worked in government organisations, with students (12.4%, 97) and private sector workers (12.5%, 98) being common among participants. Very few business people (2.3%, 18) completed the questionnaire. About 75% (595) of participants had previously used m-government applications. Table 1 presents the rest of the demographic information.

##### B. Measurement Scale Analysis

1) *Validity*: Construct validity is defined as the extent to which items in an instrument reflect the theoretical construct. Construct validity is also defined as the extent to which one can measure the concept that should be measured [67]. In the same vein, it has been claimed that validity seeks to measure the extent to which the instrument achieves its goals [68], [69]. Factor analysis is a popular analytic tool to measure construct validity [69].

In this study, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted to measure construct validity by splitting the sample randomly, using SPSS, into two samples. It has been claimed that “This approach allows for cross validation of the final factor structure in a subsample that is relatively independent from efforts to refine the item pool. This reduces capitalizing on sample-specific variance” [70]. One-half of the sample (n = 388) was used in an EFA to explore the scale’s underlying factor structure. It was then used to measure the reliability of the instrument using Cronbach’s alpha. The other half (n = 394) was used in a CFA to test the goodness-of-fit of the revealed factor structures [71].

##### a) Exploratory Factor analysis (EFA)

Exploratory factor analysis (EFA) is used for several purposes, including developing and testing instruments [72]. According to Field [73], “this technique has three main uses: (1) to understand the structure of a set of variables, (2) to construct a questionnaire to measure an underlying variable, (3) to reduce a data set to a more manageable size while retaining as much of the original information as possible”. While all factors in this study were adopted from previous studies based on a literature review, EFA was considered appropriate because these factors have not been previously applied empirically in an m-government context in Saudi Arabia.

Kaiser-Meyer-Olkin (KMO) was examined to assess the adequacy of sampling. It has been confirmed that the values of KMO correlation for EFA analysis results are adequate if above 0.6 to 0.7 [68]. Furthermore, Field [74] asserted that KMO values less than 0.5 indicate unacceptable sampling adequacy while those greater than 0.9 indicate excellent sampling adequacy. Consistent with this, de Vaus [75] asserted that KMO values should be above 0.5.

EFA requires two important steps, namely, 1) factor extraction; and 2) factor rotation. Therefore, a Principal component analysis (PCA) was used as extraction method to determine the factors that explain the structure of the variables [76] and to obtain an empirical abstraction of the number of factors and to identify the factors in the data [77]. A Varimax rotation was conducted to maximise the orthogonality (independence) of separate factors [78]. The Varimax rotation also provides clearer separation of factors [79]. There are some criteria regarding factor loadings after Varimax rotation. It has been asserted that factor loadings less than 0.4 should be discarded because they are too low [73], [79]. In this study, any factor loading value less than 0.5 was excluded to assure that all items have practical significance [79].

EFA was conducted in this study with one split sample (n = 388) with the following results. The KMO value for an EFA that includes all constructs was 0.932, which indicates excellent sampling adequacy for this study. Moreover, Bartlett’s Test of Sphericity (Approx. Chi-Square = 10163.185) was highly significant (p < 0.001) and is consistent with there being a strong relationship between items included in the analysis [73]. These findings support the factorability of the EFA conducted for these factors [73], [79], [80]. The seven factors in the rotated component matrix: perceived usefulness PU, attitude toward using ATU, perceived ease of use PEU, behavioural intention to use BIU, perceived trustworthiness TRU, perceived security SEC and awareness AWAR are presented in Table 2.

TABLE. II. ROTATED COMPONENT MATRIX

	Component						
	1	2	3	4	5	6	7
PU1		0.586					
PU2		0.740					
PU3		0.775					
PU4		0.813					
PU5		0.752					
PU6		0.688					
ATU1						0.835	
ATU2						0.804	
ATU3						0.859	
PEU1	0.732						
PEU2	0.750						
PEU3	0.798						
PEU4	0.739						
PEU5	0.755						
PEU6	0.669						
BIU1							0.740
BIU2							0.694
BIU3							0.597
TRU1				0.697			
TRU2				0.664			
TRU3				0.765			
TRU4				0.714			
TRU5				0.697			
SEC1			0.761				
SEC2			0.839				
SEC3			0.850				
SEC4			0.747				
SEC5			0.810				
AWAR1					0.846		
AWAR2					0.856		
AWAR3					0.862		

In summary, the EFA test returned a seven-components solution that explained 76.746% of the cumulative variance, with a KMO measure of sampling of 0.932, which is consistent with these items being highly suitable for factor analysis.

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

*b) Reliability*

According to Drost [81], “reliability is consistency of measurement or stability of measurement over a variety of conditions in which basically the same results should be obtained”. After identifying the scale structure via EFA, the reliability of the instrument was evaluated using Cronbach’s alpha on the half-sample used for the EFA (n = 388). Sekaran [66] asserted that internal consistency is a very common measure of reliability in information systems (IS). Internal consistency relates to consistent responses for items to scale a single measurement [82]. Reliability can be divided into four ranges: values of up to 0.50 have low reliability; from 0.50 to 0.70 moderate reliability; from 0.70 to 0.90 high reliability; and from 0.90 and above excellent reliability [83]. It has been noted that Cronbach’s alpha values should be 0.7 or higher to conclude that the internal consistency is reliable [79]. The results presented in Table 3 indicate that all values were higher than 0.70 and most were nearly 0.9, which is considered excellent reliability. Furthermore, when all items were entered at the same time, the overall reliability for this instrument was 0.947, which is considered excellent. Therefore, the results of this study can be considered reliable in the Saudi m-government context.

*a) Overview SEM*

A Structural Equation Model (SEM) comprises a collection of statistical methods that can be used to clarify and analyse relationships between variables [84]. SEM has two components, the measurement model and the structural model [79], [82], [85]. SEM can also test theoretical models [79]. Moreover, SEM tests hypotheses between variables by measuring the magnitude of the path of the coefficient between variables [86]. Gefen, Straub, and Boudreau [87] recommended the use of SEM in IT/IS studies, especially in behavioural studies.

TABLE. III. RELIABILITY COEFFICIENT VALUES

Constructs	Number of items	Cronbach Alpha reliability	Comments
Perceived Usefulness (P U)	6	0.927	Excellent reliability
Attitude toward using (ATU)	3	0.907	Excellent reliability
Perceived Ease of Use (PEU)	6	0.900	Excellent reliability
Behavioural intention to use (BIU)	3	0.816	High reliability
perceived trustworthiness (TRU)	5	0.894	Excellent reliability
perceived security (SEC)	5	0.914	Excellent reliability
Awareness (AWAR)	3	0.962	Excellent reliability
Overall reliability	<b>31</b>	0.947	Excellent reliability

SEM describes the relationships between variables clearly using graphic diagrams [82]. Confirmatory factor analysis (CFA) is used to measure the model by assessing the indicators used to measure the latent variables [85]. In this study, the SEM procedure, Confirmatory Factor Analysis, was conducted via SPSS AMOS 24. It measured the validity of the factor structure of half the sample (n = 394) by using CFA. SPSS AMOS 24 was then used to assess relationships between variables via a structural model that used the entire sample (n = 782), as reported later.

*b) Measurement model using CFA*

CFA is a SEM analysis technique [84], [88]. The main reason for using CFA is to measure the construct validity of the hypothesised factor structure [79]. Harrington [89] claimed that there are four reasons for conducting CFA: testing method effects, construct validation, psychometric evaluation of measures, and testing measurement invariance. CFA is also a suitable technique to measure the validity of scales [90]. CFA focuses on the construct's validity and the model's overall fit. It assesses the measurement theory by using empirical evidence of the validity of items [84]. Two aims of CFA are to examine the relationship between a group of continuous latent variables and a group of observed variables [91]. In addition, CFA is conducted to define the goodness-of-fit between collected data and a model used in another study [92]. Finally, CFA is widely used to analyse latent variables [93].

Fit indices used to assess the fit of the measurement model include: CMIN (minimum discrepancy), Chi-Square ( $\chi^2$ ), Goodness-of-fit (GFI) and adjusted goodness of fit (AGFI), Comparative fit index (CFI), Incremental fit index (IFI), and the Root Mean Square Error Approximation (RMSEA) [84]. Table 4 summarizes all criteria for model assessment.

AMOS 24 was used to conduct CFA for all constructs to assess the model fit using half of the study sample (n = 394) (Fig. 3). The results reveal that the model has acceptable fit values and is valid (Table 5). Consequently, the model is fit and valid in the Saudi m-government context.

*C. Model Assessment*

*1) Assessment of the hypotheses:* The structural equation model (SEM) for the entire sample (n = 782) examined six hypotheses via Maximum Likelihood, by calculating the significance and the strength of each pathway Fig. 4. Only the significant results are reported here.

TABLE. IV. SUMMARY OF CRITERIA FOR MODEL ASSESSMENT

Fit indices	Criteria	References
$\chi^2/df$ (CMIN/df)	<3 is a good fit; <5 is an acceptable fit	[84], [88], [94]
RMSEA	<0.05 is an excellent fit; <0.08 is a good fit; <0.1 is an acceptable fit.	[88], [95]
GFI, AGFI, IFI and CFI	AGFI >0.80 is a good fit; GFI, AGFI, IFI and CFI >0.95 is an excellent fit; > 0.90 is a good fit; > 0.80 is an acceptable fit.	[84], [87], [96]-[98]

TABLE. V. RESULTS OF THE MODEL GOODNESS-OF-FIT INDICES BY CFA

Fit indices	Result	Comment
$\chi^2/df$ (CMIN/df)	1.987	Good fit
RMSEA	0.050	Excellent fit
GFI	0.886	Acceptable fit
AGFI	0.857	Good fit
IFI	0.960	Excellent fit
CFI	0.959	Excellent fit

The results revealed some non-significant paths, including those involving perceived ease of use, perceived security and awareness. Therefore, H3, H5 and H6 are not supported in the Saudi m-government context. Significant paths between factors and BIU were found for perceived usefulness, attitude toward using and perceived trustworthiness. Therefore, hypotheses H1, H2 and H4 are confirmed and supported in the Saudi m-government context:

*a) Perceived usefulness (PU) has a significant positive influence on Behavioural intention to use ( $\beta = 0.266$ ,  $P < 0.001$ ) and, accordingly, H1 is supported.*

*b) Attitude toward using (ATU) has a significant positive influence on Behavioural intention to use ( $\beta = 0.484$ ,  $P < 0.001$ ) and, accordingly, H2 is supported.*

*c) Perceived trustworthiness (TRU) has a significant positive influence on Behavioural intention to use ( $\beta = 0.247$ ,  $P < 0.001$ ) and, accordingly, H4 is supported.*

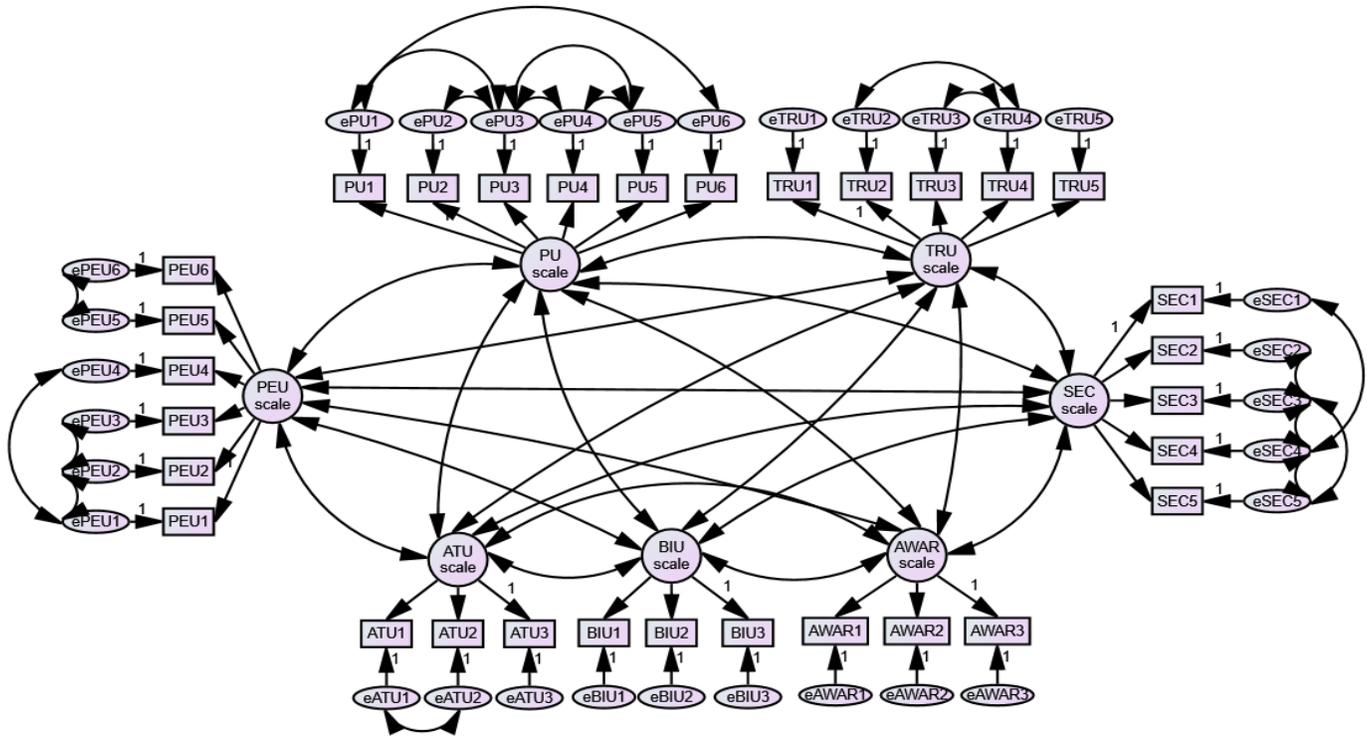


Fig. 3. CFA for all constructs.

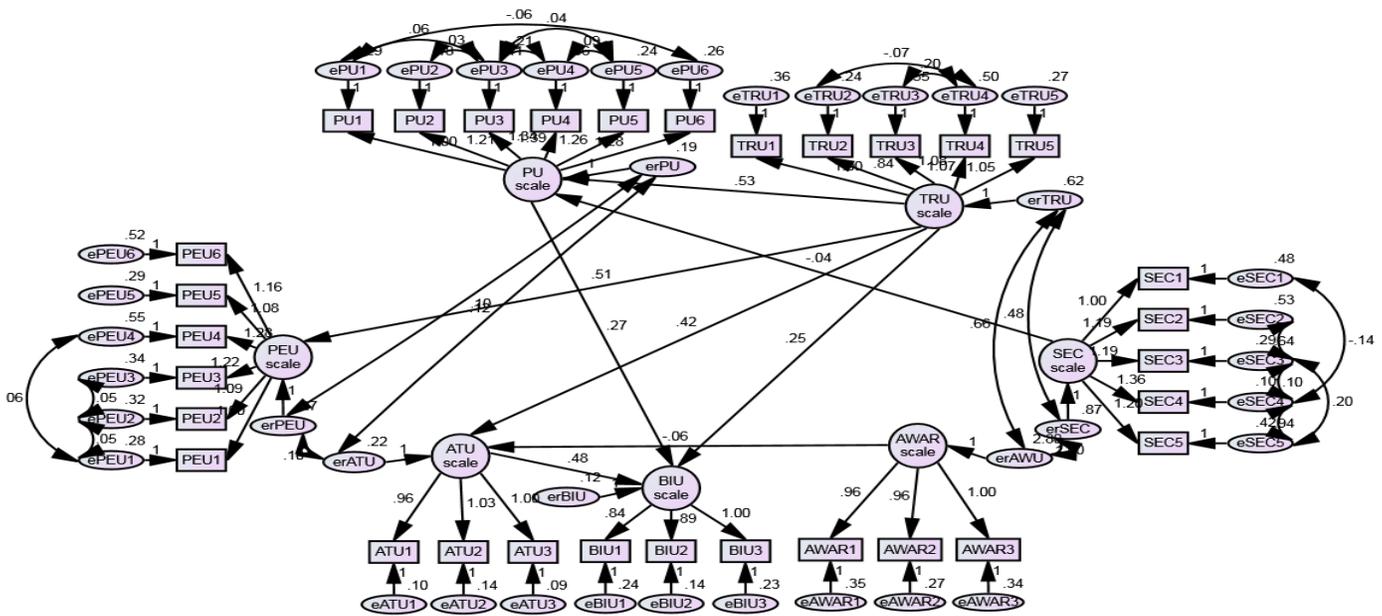


Fig. 4. Structural model for standardized path coefficients.

## V. DISCUSSION

The results of this study demonstrated that all TAM construct factors, except perceived ease of use, have a

significant positive influence on BIU which is compatible with the original TAM outcomes [39]. This unusual outcome is also supported by [42], where PEU did not affect intention to use e-learning in Korea. Another study [99] found that PEU had no

influence on intention to use e-commerce in Indonesia. Other studies have stated that ease of use was not a strong factor for measuring the intention to use new systems [39], [100]. Davis [39] stated that ease of use is better for predicting usefulness rather than intention to use. These results indicate that users accepted the difficulty in using m-government applications which have already been released. In other words, they do not care about ease of use, they are focused on obtaining the services they need, regardless of whether this is easy or not. It goes without saying that decision makers in Saudi government who are providers of m-government applications for citizens, especially Yesser, should provide the services via applications in a user friendly way with clear and simple to follow instructions to encourage the users to use and operate these applications. In fact, ease of use of applications is considered an advantage in general, although the result of the current research revealed that it was not considered important. Also, decision makers in the Saudi government who provide m-government applications for citizens, especially Yesser, should consider perceived usefulness when providing these applications because the results of this and previous studies show this factor to be very important to users.

The preliminary analysis and data screening revealed that participants tended to give high ratings to items, usually in the moderately agree (6) to strongly agree (7) interval, particularly for items related to ATU. For this reason, both the items and the associated scale scores were somewhat negatively skewed. Closer observation of the ratings suggested that few of the participants were prepared to give negative ratings, and relatively few were prepared to even give a neutral rating. In effect, the majority of the participants were using a four-point scale that utilised the four main positive responses (4 = Neutral, 5 = Slightly Agree, 6 = Moderately Agree and 7 = Strongly Agree). Previous results showed that ATU has a significant positive influence on BIU. This means that the sample population has a positive attitude toward using m-government applications and intends to use these applications in the future. Therefore, the majority of participants believe that using m-government applications is a good and positive idea. This outcome is supported by a study [101] which found that the majority of participants (male and female) wanted to use government services via mobile devices in Saudi Arabia. Results of the current study show that most participants have a positive attitude toward using m-government applications. This suggests that citizens will use m-government without any resistance to change and will support the government decision to provide these services. Therefore, decision makers in Saudi government, especially Yesser, who are providers of m-government applications for citizens, should continue to expand the provision of services via applications.

The outcomes of this study revealed that perceived trustworthiness has a positive significant influence on BIU of m-government applications. This suggests that most participants in this study trust m-government applications and their benefits because these applications have been released by the government. This outcome is consistent with the literature showing that when trust in the internet and government increases, increases in the intention to use e-government services also occur [35]. Carter and Weerakkody [102] found

that trust has a positive impact on e-government services in the UK. The study by Alrowili, Alotaibi and Alharbi [20] supports this outcome, finding that perceived trust had a direct positive effect on users' intentions to use m-government services in Saudi Arabia. The recommendation for decision makers in the Saudi Government who are providers of m-government applications for citizens, especially Yesser, is that they should try to build trust with Saudi citizens to encourage them to use services via m-government applications. Also, applications designers and developers should build trustworthy applications, because trust has been shown to be a very important factor for citizens in their use of m-government application.

Perceived security had no influence on BIU. This outcome might indicate that participants do not care about security because they trust applications released by the government. Security may also not be an issue for the majority of participants who intend to use m-government applications, because they have been using the internet for a long time and are aware of security and privacy problems [103]. This outcome is supported by several studies. Ratten [104] found that privacy concerns do not impact on purchase intentions of cloud computing services in India and the USA. A later study [105], found that security is not much of a concern for consumers who intend to adopt cloud computing in the USA or Australia. More recently, it was found that increased security does not lead to increased intention to use online channels for purchase in Jordan [103]. The recommendation for decision makers in the Saudi government who are providers of m-government applications for citizens, especially Yesser, is that they should continue to be concerned about this factor and try to increase the level of perceived security in applications, although this is not a concern in the current research study. Also, they should provide services via secure applications to encourage users to use these applications. In fact, increased perceived security in applications is considered to be an advantage. Therefore, application designers and developers should build these applications, maintaining their security. Also, the Saudi government should issue a law to protect users if they have problems regarding security and privacy of using m-government applications.

The outcomes of this study revealed that awareness does not influence BIU. This outcome is incompatible with some previous empirical studies. For example, it has been found that awareness has a positive influence on intention to adopt m-banking [106]. However, the results of this study indicate that the majority of participants have a low level of awareness about m-government applications because they do not have knowledge or information about these applications. This may indicate that government sectors do not raise awareness about their services to citizens. Therefore, several other studies have suggested that government should increase the rate of awareness among citizens by using media such social networks, television advertisements and newspapers [58], [107]. The recommendation for decision makers in the Saudi government who are providers of m-government applications for citizens, especially Yesser, is that they should act on the concern about this factor and try to distribute information and knowledge for citizens about m-government applications and their benefits via the media to inform citizens about these apps.

## VI. CONCLUSION AND IMPLICATIONS

The applicability of the modified TAM model to fulfill the study objectives was evaluated by measuring the influence of six independent factors (PU, ATU, PEU, TRU, SEC and AWAR) on users' intentions to use m-government applications in Saudi Arabia.

The study identified the influential factors as being ATU, PU and TRU. The model was modified to include these factors and tests revealed that the modified TAM model was suitable and fit for the Saudi m-government context. The results also showed that most participants had a positive attitude towards using m-government applications.

The results of this study will help decision makers involved in e-government and m-government initiatives in Saudi Arabia to adapt m-government applications properly and successfully by acknowledging and addressing the key influential factors. The modified model can be used in other countries to understand factors that influence users' intention to use m-government applications.

## VII. DIRECTIONS FOR FUTURE WORK

This study adopted the TAM model, with some modifications, to fulfil the study objectives by measuring the influence of six factors (PU, ATU, PEU, trust, security and awareness) on users' intention to use m-government applications in Saudi Arabia. In future research it would be better to incorporate TAM with other models such as DOI in order to examine the applicability of the incorporated model in the Saudi m-government context. This would lead to measuring and considering other factors which were not measured in this study but that may influence users' intention to use m-government applications. Since this study focused on mG2C systems to explore and measure the factors that influence users' intentions to use m-government applications from citizens' perspectives, it has disregarded other types of m-government such as m-government-to-government (mG2G), m-government-to-employee (mG2E) and m-government-to-business (mG2B). Future research should focus on one or more of these other types to explore and measure the factors that influence behavioural intention to use m-government applications. Changing the context between countries may also be a useful way to further develop and validate this model.

### REFERENCES

- [1] O. Alfarraj, Factors influencing the development of e-government in Saudi Arabia: a qualitative investigation of the developers perspectives, School of Information and Communication Technology, Griffith University, Brisbane, Australia, 2013.
- [2] O. Alfarraj and T. Alhussain, "Making Sense of E-Government development in Saudi Arabia: A Qualitative Investigation," paper presented at The Eighth International Conference on Forensic Computer Science, 2013.
- [3] O. Alfarraj, T. Alhussain, and A. Abugabah, "Identifying the Factors Influencing the Development of eGovernment in Saudi Arabia: The Employment of Grounded Theory Techniques," International Journal of Information and Education Technology, vol. 3, no. 3, p. 319, 2013.
- [4] S. Alghamdi and N. Beloff, "Innovative Framework for e-Government adoption in Saudi Arabia: A Study from the business sector perspective," International Journal of Advanced Computer Science and Applications, vol. 7, no. 1, pp. 655-664, 2016.
- [5] J. Moon, From E-Government to M-Government?: Emerging Practices in the Use of Mobile Technology by State Governments. E-Government Series Texas, USA: IBM Center for The Business of Government, 2004.
- [6] Althunibat, T. Alrawashdeh, and M. Muhairat, "The acceptance of using m-government services in Jordan," Journal of Theoretical and Applied Information Technology, vol. 63, no. 3, pp. 733-740, 2014..
- [7] T. El Kiki, and E. Lawrence, "Government as a mobile enterprise Real-time, ubiquitous government," in Information Technology: New Generations, Third International Conference on , 2006. IEEE.
- [8] O. Östberg, "A Swedish View on 'Mobile Government,'" paper presented at the International Symposium on E & M-Government, Seoul, Korea, <http://www.statskontoret.se/upload/Publikationer/2003/2003128.pdf. 2003.
- [9] M. Kumar, and O. Sinha, "M-government-mobile technology for e-government," paper presented at the International conference on e-government, India, 2007.
- [10] Misra, "Make M-Government an Integral Part of E-Government: An Agenda for Action," Compendium, pp. 78-86, 2009.
- [11] H. Sheng, and S. Trimi, "M-government: technologies, applications and challenges," Electronic Government, An International Journal, 5, no. 1, pp. 1-18, 2008.
- [12] Mengistu, H. Zo, and J. Rho, "M-Government: Opportunities and challenges to deliver mobile government services in developing countries," paper presented at the Fourth International Conference on Computer Sciences and Convergence Information Technology, 2009. IEEE.
- [13] T. Ahmad, T, et al., "CURRENT REVIEW OF ICT AND M-GOVERNMENT SERVICES IN SAUDI ARABIA," International Journal of Computer Engineering and Applications, vol. VII, no. II, pp. 71-77, 2014.
- [14] CITC, Annual report. Communication and Information technology commission Saudi Arabia, 2015.
- [15] M. Khan, "Exploring the Push and Pull Drivers in M-Government Framework that Influence Acceptance of Services on Mobile Devices," International Journal of Computer Science and Network Security (IJCSNS), vol. 16, no. 2, p. 23, 2016.
- [16] A. Baabdullah, O. Nasseef, and A. Alalwan, Consumer adoption of mobile government in the Kingdom of Saudi Arabia: the role of usefulness, ease of use, perceived risk and innovativeness, in IFIP International Federation for Information Processing. 2016. Switzerland: Springer International Publishing.
- [17] A. Babullah, Y. Dwivedi, and M. Williams, "Saudi Citizens' Perceptions on Mobile Government (mGov) Adoption Factors," in UK Academy for information systems conference proceedings 2015. UK: Association for information systems AIS electronic library (AISeL).
- [18] T. Alhussain, Factors Influencing the Adoption of Biometric Authentication in Mobile Government Security, in school of information and communication technology science. 2012, Griffith University: Brisbane, Australia.
- [19] A. Alsenaidy, and T. Ahmad, A review of current state m-government in Saudi Arabia. 2012.
- [20] T. Alrowili, M. Alotaibi, and M. Alharbi, "Predicting citizens' acceptance of M-government services in Saudi Arabia an empirical investigation," paper in the Proceedings of the 2015 Annual IEEE Systems Conference (SysCon), Marriott Pinnacle Vancouver, Vancouver, BC, Canada: IEEE, 2015.
- [21] S. Alotaibi, and D. Roussinov, "A conceptual model for examining mobile government adoption in Saudi Arabia," in the Proceedings of The 15th European Conference on eGovernment ECEG, University of Portsmouth, 2015. Barcelona, Spain.
- [22] M. Ntaliani, C. Costopoulou, and S. Karetso, "Mobile government: A challenge for agriculture," Government Information Quarterly, vol. 25, no. 4, pp. 699-716, 2008.
- [23] A. Al-Hadidi, and Y. Rezgui, "Adoption and diffusion of m-Government: Challenges and Future Directions for Research," in Collaborative Networks for a Sustainable World.;Springer, 2010, pp. 88-94.

- [24] A. Nava, and I. Dávila, "M-Government for Digital Cities: Value Added Public Services," in The Proceedings of the 1st European Mobile Government Conference, UK: Mobile Government Consortium International Pub., 2005.
- [25] K. Assar, "M-government in Saudi Arabia," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 5, no. 1, pp. 76-83, 2015.
- [26] Y. Kim, et al., Architecture for implementing the mobile government services in Korea," in Conceptual modeling for advanced application domains. 2004, Springer. p. 601-612.
- [27] A. Abanumy and P. Mayhew, "M-government implications for e-government in developing countries: The case of Saudi Arabia," in EURO mGOV. 2005. Brighton, UK.
- [28] L. Antovski, and M. Gusev, "M-government framework," in Euro mGov. 2005. Brighton, UK.
- [29] H. Scholl, "The mobility paradigm in government theory and practice: A strategic framework," in Euro mGov. 2005. Brighton, UK.
- [30] A. Jahanshahi, et al., "Comprehensive model of mobile government in Iran," Indian Journal of Science and Technology, vol. 4, no. 9, pp. 1188-1197, 2011.
- [31] I. Snellen, and M. Thaens, "From e-government to m-government: towards a new paradigm in public administration," Working document, Erasmus University, Rotterdam, 2008.
- [32] M. Alomari, H. Elrehail, and H. Al Shibly, "Mobile-Government: Challenges and Opportunities Jordan as Case study," International Journal of Business and Social Science, vol. 4, no. 12, pp. 244-250, 2013.
- [33] A. Al-Hadidi, "Exploratory study on adoption and diffusion of m-government services in the sultanate of oman," in School of Engineering. 2010, CARDIFF UNIVERSITY (UNITED KINGDOM).
- [34] T. Ahmad, et al., "Current review of ICT and m-government services in Saudi Arabia," Int. J. Comput. Eng. Appl, vol. 7, no. 2, 2014.
- [35] L. Carter, and F. Bélanger, "The utilization of e-government services: citizen trust, innovation and acceptance factors," Information systems journal, vol. 15, no.1, pp. 5-25, 2005.
- [36] B. Dickerson, "Motivation factors impacting employee acceptance of new technology," in School of Business and Technology. 2013, Capella University: USA.
- [37] F. Davis, "A technology acceptance model for empirically testing new end-user information systems: Theory and results," in School of Management. 1986, Massachusetts Institute of Technology: USA.
- [38] S. Alharbi, and S. Drew, "Using the technology acceptance model in understanding academics' behavioural intention to use learning management systems," (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 5, no.1, pp. 143-155, 2014.
- [39] F. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," MIS Quarterly, vol. 13, no. 3, pp. 319-340, 1989.
- [40] O. Al-Hujran, M. Al-dalameh, and A. Aloudat, "The role of national culture on citizen adoption of eGovernment services: An empirical study," Electronic Journal of E-government, vol. 9, no. 2, pp. 93-106, 2011.
- [41] Al-Adwan, A., A. Al-Adwan, and J. Smedley, Exploring students acceptance of e-learning using Technology Acceptance Model in Jordanian universities. International Journal of Education and Development using Information and Communication Technology, vol. 9, no. 2, pp. 4-18, 2013.
- [42] S. Park, "An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use e-Learning," Educational Technology & Society, vol. 12, no. 3, pp. 150-162, 2009.
- [43] P. Chau, "An empirical assessment of a modified technology acceptance model," Journal of management information systems, vol. 13, , pp. 185-204, 1996.
- [44] Q. Ma, and L. Liu, "The no. 2 acceptance model: A meta-analysis of empirical findings," Journal of Organizational and End User Computing (JOEUC), vol. 16, no.1, pp. 59-72, 2004.
- [45] J. Moon, and Y. Kim, "Extending the TAM for a World-Wide-Web context," Information & Management, vol. 38, no. 4), pp. 217-230, 2001.
- [46] K. Al-Busaidi, and H. Al-Shihi, "Instructors' Acceptance of Learning Management Systems: A Theoretical Framework," Communications of the IBIMA 2010. 2010.
- [47] C. Ong, J. Lai, and Y. Wang, "Factors affecting engineers' acceptance of asynchronous e-learning systems in high-tech companies," Information & management, vol.41, no. 6), pp. 795-804, 2004.
- [48] P. Palvia, "The role of trust in e-commerce relational exchange: A unified model," Information & Management, vol. 46, no. 4, pp. 213-220, 2009.
- [49] D. Gefen, and D. Straub, "Managing user trust in B2C e-services, E-service Journal, vol. 2, no. 2, pp. 7-24, 2003.
- [50] T. Wei, et al., "What drives Malaysian m-commerce adoption? An empirical analysis," Industrial Management & Data Systems, vol. 109, no. 3, pp. 370-388, 2009.
- [51] T. Teo, S. Srivastava, and L. Jiang, "Trust and electronic government success: An empirical study," Journal of Management Information Systems, vol. 25, no. 3, pp. 99-132, 2008.
- [52] A. Chang, and P. Kannan, "Preparing for wireless and mobile technologies in government," E-, pp. 345-393, 2003.
- [53] R. Peng, L. Xiong, and Z. Yang, "Exploring tourist adoption of tourism mobile payment: An empirical analysis," Journal of Theoretical and Applied Electronic Commerce Research, vol. 7, no. 1, pp. 21-33, 2012.
- [54] M. Mahad, et al., "Factor Affecting Mobile Adoption Companies in Malaysia," International Journal of Economics and Financial 5, pp. 84-91, 2015.
- [55] M. Alomari, "Predictors for Successful E-government Adoption in the Hashemite Kingdom of Jordan: The Deployment of an Empirical Evaluation Based on Citizen-Centric Perspectives", in International Business and Asian studies. 2011, Griffith University: Brisbane, Australia.
- [56] S. Kumar and S. Madhumohan, "Internet banking adoption in India," Journal of Indian Business Research, vol. 6, no. 2, pp. 155-169, 2014.
- [57] M. Meftah, B. Gharleghi, and B. Samadi, "Adoption of E-Government among Bahraini Citizens," Asian Social Science, vol. 11, no. 4, pp. 141-149, 2015.
- [58] I. Abunadi, "Influence of Culture on e-Government Acceptance in Saudi Arabia," in School of Information and Communication Technology. 2012, Griffith University: Brisbane, Australia.
- [59] M. Balnaves, and P. Caputi, Introduction to Quantitative Research Methods: An Investigative Approach. London: SAGE Publications, 2001.
- [60] R. Alotaibi, K. Sandhu, and L. Houghton, "A Study of Service Users' Attitudes towards E-Government Initiatives in the Kingdom Of Saudi Arabia," (IJCSIT) International Journal of Computer Science and Information Technologies, vol. 5, no. 6, pp. 6892-6901, 2014.
- [61] H. Al-Busaidi, "A model of intention to use mobile government services," in School of Management and Information Systems, Faculty of Business and Law. 2012, Victoria University: Australia.
- [62] M. AL-Majali, "No more traditional stock market exchange: a study of internet trading service (ITS) in Jordan," Journal of Internet Banking and Commerce, vol. 17, no. 1, pp. 1-19, 2012.
- [63] C. Teddlie, and F. Yu, "Mixed methods sampling a typology with examples," Journal of mixed methods research, vol. 1, no. 1, pp. 77-100, 2007.
- [64] Stats. General Information about The Kingdom of Saudi Arabia. 2016 [cited 25 December 2016]; Available from: <http://www.stats.gov.sa/en/4025>.
- [65] H. Alsaghier, "An investigation of critical factors affecting citizen trust in e-government: empirical evidence from Saudi Arabia," in Information Communication, Technology. 2010, Griffith University. : Brisbane, Australia.
- [66] U. Sekaran, Research methods for business: a skill-building approach, vol. 4. New York; Great Britain: Wiley, 2003.

- [67] Bagozzi, R., Y. Yi, and L. Phillips, Assessing construct validity in organizational research. *Administrative science quarterly*, 1991. 36(3): p. 421-458.
- [68] Netemeyer, R.G., W.O. Bearden, and S. Sharma, *Scaling Procedures: Issues and Applications*. 2003, Thousand Oaks, CA: SAGE Publications.
- [69] Turocy, P.S., Survey research in athletic training: the scientific method of development and implementation. *Journal of Athletic Train*, 2002. 37(suppl 4): p. 174-179.
- [70] Dik, B., et al., Development and validation of the calling and vocation questionnaire (CVQ) and brief calling scale (BCS). *Journal of Career Assessment*, 2012. 20(3): p. 1-22.
- [71] Ng, S., Validation of the 10-item Chinese perceived stress scale in elderly service workers: one-factor versus two-factor structure. *BMC psychology*, 2013. 1(9): p. 1-8.
- [72] Osborne, J. and A. Costello, *Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical assessment, research & evaluation*, 2005. 10(7): p. 1-9.
- [73] Field, A., *Discovering statistics using SPSS*. 2nd ed. 2005, London: Sage Publication.
- [74] Field, A., *Discovering Statistics using IBM SPSS Statistics*. 3rd ed. 2009: SAGE Publications.
- [75] de Vaus, D., *Surveys in Social Research*. 2002, St.Leonards, NSW: Allen & Unwin.
- [76] M. Alnatheer, Understanding and measuring information security culture in developing countries: case of Saudi Arabia, in *Faculty of Science and Technology*. 2012, Queensland University of Technology: Brisbane, Queensland, Australia.
- [77] M. Alshehri, Using the UTAUT Model to Determine Factors Affecting Acceptance and Use of E-government Services in the Kingdom of Saudi Arabia, in *School of Information and Communication Technology*. 2012, Griffith University: Brisbane, Australia.
- [78] B.G. Tabachnick, and L.S. Fidell, *Using Multivariate Statistics*. 5th ed. Boston: Pearson, 2007.
- [79] J.F. Hair, et al., *Multivariate Data Analysis*. 6th ed. Upper Saddle River: Pearson, 2006.
- [80] J. Pallant, *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS*. 2nd ed. Crows Nest, NSW: Allen & Unwin, 2005.
- [81] E. Drost, "Validity and reliability in social science research," *Education Research and Perspectives*, vol. 38, no. 1, pp. 105, 2011.
- [82] R.B. Kline, *Principles and Practice of Structural Equation Modeling*. 2nd ed. New York: Guildwood, 2005.
- [83] P.R. Hinton, et al., *SPSS Explained*. East Sussex, England: Routledge, 2004.
- [84] J.F. Hair, et al., *Multivariate data analysis: a global perspective*. 7th ed. New Jersey: Pearson, 2010.
- [85] R.H. Hoyle, *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks: SAGE Publications, 1995.
- [86] B.M. Byrne, *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum Associates, 2001.
- [87] D. Gefen, D. Straub, and M. Boudreau, "Structural equation modeling and regression: Guidelines for research practice," *Communications of the association for information systems*, vol. 4, no. 7, pp. 1-77, 2000.
- [88] B.M. Byrne, *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*. 2nd ed. New York: Taylor & Francis Group, 2010.
- [89] D. Harrington, *Confirmatory Factor Analysis*. Oxford, New York: Oxford University Press, USA, 2009.
- [90] A. Bhattacharjee, and G. Premkumar, "Understanding changes in belief and attitude toward information technology usage: a theoretical model and longitudinal test," *MIS quarterly*, vol. 28, no. 2, pp. 229-254, 2004.
- [91] F. Baker, *Item response theory: parameter estimation techniques*. Vol. 176. 2004, New York: Marcel Dekker.
- [92] M. Weitzner, et al., "Developing a care giver quality-of-life instrument. Preliminary steps," *Cancer practice*, vol. 5, no. 1, pp. 25-31, 1996.
- [93] W. Chin, and P. Todd, "On the use, usefulness, and ease of use of structural equation modeling in MIS research: a note of caution," *MIS quarterly*, vol. 19, no. 2, pp. 237-246, 1995.
- [94] T.A. Brown, *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press, 2006.
- [95] D. Hooper, J. Coughlan, and M. Mullen, "Structural equation modelling: Guidelines for determining model fit," *The Electronic Journal of Business Research Methods*, vol. 6, no. 1, pp. 53-60, 2008.
- [96] P. Barrett, "Structural equation modelling: Adjudging model fit," *Personality and Individual Differences*, vol. 42, no. 5, pp. 815-824, 2007.
- [97] Dawes, J., M. Faulkner, and B. Sharp, "Business orientation scales: development and psychometric assessment," paper presented at the 27th EMAC Conference. Stockholm, Sweden, 1998.
- [98] P. Greenspoon, and D. Saklofske, "Confirmatory factor analysis of the multidimensional students' life satisfaction scale," *Personality and Individual Differences*, vol. 25, no. 5, pp. 965-971, 1998.
- [99] M. Luthfihadi, and W. Dhewanto, "Technology Acceptance of E-commerce in Indonesia," *International Journal of Engineering Innovation and Management*, vol. 3, pp. 9-18, 2013.
- [100] E. Karahanna, and D. Straub, "The psychological origins of perceived usefulness and ease-of-use," *Information & Management*, vol. 35, no. 4, pp. 237-250, 1999.
- [101] T. Alhussain, and S. Drew, "Towards Secure M-Government Applications: A survey study in the Kingdom of Saudi Arabia," paper presented at the 2010 International Conference on Intelligent Network and Computing (ICINC 2010), Kuala Lumpur, Malaysia IEEE, 2010.
- [102] L. Carter, L. and V. Weerakkody, "E-government adoption: A cultural comparison," *Information Systems Frontiers*, vol. 10, no. 4, pp. 473-482, 2008.
- [103] K. Faqih, "An empirical analysis of factors predicting the behavioral intention to adopt Internet shopping technology among non-shoppers in a developing country context: Does gender matter?" *Journal of Retailing and Consumer Services*, vol. 30, pp. 140-164, 2016.
- [104] V. Ratten, "Indian and US consumer purchase intentions of cloud computing services," *Journal of Indian Business Research*, vol. 6, no. 2, pp. 170-188, 2014.
- [105] V. Ratten, "International consumer attitudes toward cloud computing: A social cognitive theory and technology acceptance model perspective," *Thunderbird International Business Review*, vol. 57, no. 3, pp. 217-228, 2015.
- [106] R. Safeena, N. Hundewale, and A. Kamani, "Customer's Adoption of Mobile-Commerce A Study on Emerging Economy," *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 1, no. 3, pp. 228-233, 2011.
- [107] T. Al-Tourki, et al., "E-government in Saudi Arabia: Barriers, Challenges and its Role of Development," *e-government*, vol. 48, no. 5, 2012.

# Image and AES Inspired Hex Symbols Steganography (IAIS) for Anti-Forensic Artifacts

Somyia M. Abu Asbeh

Software Engineering Dept.  
Princess Sumaya University for  
Technology  
Amman, Jordan

Sarah M. Hammoudeh

Faculty of Medical and Human  
Sciences  
University of Manchester  
Manchester, UK

Arab M. Hammoudeh

College of Medicine  
University of Sharjah  
Sharjah, UAE

**Abstract**—Technology (including mobiles and computers) has become a basic, indispensable need in our daily life. With an initial purpose of achieving basic functions such as communication, technology has evolved into a virtual gate to the whole world connecting individuals through social media and various websites and applications. Most importantly, technology became the reservoir of our personal information and important, sensitive data. This has led to increased risks of security breaches and data thefts demanding countermeasure approaches. One of these approaches is Steganography. Steganography is a data hiding approach that allows for invisible, relatively safe communication. Several forms of steganography have been developed, among which are Image steganography and our previously developed AES Inspired Steganography. In this paper we propose a new variation in which we combine both of these approaches, Image and AES Inspired Steganography (IAIS). This approach proposes hiding the hex symbol format of the encrypted secret data into a carrier image file. The image file is converted to a hexadecimal representation in which the hex symbol could be embedded without applying any noticeable changes to the original image. Deciphering the hidden information requires secret keys agreed upon by the communicating parties confidentially. These carrier files can be exchanged among mobile devices and/or computers. Comparisons between the original cover images and the cover images with the hidden text have shown that no changes occurred in the colour histogram of the images. However, the noise test has shown that exposure to noise can affect the hexadecimal content of the image, hence the embedded hex symbol representation of the secret text.

**Keywords**—Mobile forensics; anti-forensics; data hiding; steganography; AES; AIS

## I. INTRODUCTION

Since the rise of the Internet, information security became an important aspect in the field of information technology and communication. Cryptography was created as a technique for ensuring the secrecy of communicated information. Various approaches of data encryption and decryption were developed to prevent access by third parties. Unfortunately, discovering the existence of an encrypted text would increase the risk of access by third party. Therefore, steganography was developed to conceal the existence of the secret text.

The basic concept of Steganography is to achieve invisible communication through hiding secret information in other information files of different formats. Current technological

approaches of steganography use digital data files as carrier files and networks as high speed delivery channels [1]. In this paper, we have used both encryption and steganography to ensure a high level of information security against intrusions and hacking attempts. This approach will utilize our previously developed AES inspired hex symbol steganography [2] [3] in combination with the image steganography.

## II. RELATED WORK

This section will be discussing some previously published approaches of cryptographic algorithms and data hiding.

### A. Cryptographic algorithms

Generally, cryptographic algorithms can be classified into three basic classes; hash algorithm, symmetric key algorithms, and asymmetric key algorithms. They are defined by the number of cryptographic keys used in conjunction with the algorithm [4].

1) *Cryptographic Hash Algorithms*: No keys are required in this algorithm. They generate a relatively small digest (hash value, hash code, or hash sum) from a (possibly) large input in a way that is fundamentally difficult to reverse (i.e., hard to find an input that will produce a given output). Hash functions are used as building blocks for key management, namely; providing data authentication and integrity services (e.g. generating a hashed message authentication code, HMAC), message compression for digital signature generation and verification, generating deterministic random numbers [5]. Examples of this type are SH, SH-1, MD5, etc.

2) *Symmetric Key Algorithms*: A secret key is pre-shared between the sender and the receiver over a secure channel before the communication. The word “symmetric” refers to the fact that both sender and receiver use the same key to encrypt and decrypt the information, respectively [6]. Examples of the widely used modern symmetric cryptosystem are DES, 3DES, AES, Blowfish, RC5, RC2, CAST and IDEA.

3) *Asymmetric Key Algorithms*: Asymmetric algorithms rely on a pair of keys; one key is used for encryption and a different but related key is used for decryption, with the characteristic of being computationally infeasible to determine the decryption key given only knowledge of the cryptographic algorithm and the encryption key [7].

### B. Data Hiding

Data hiding tools have been developed to secretly embed and hide undiscoverable data through multiple approaches. These approaches include transferring data to other portable storage devices and then wiping the data from the phone or computer; making data “invisible” and concealing their existence; embedding data in multimedia (image, audio and video) files; and altering file extensions. Although some of these approaches, such as altering file extensions, are relatively old, evidence has shown that they can still bypass some forensic analysis methods. For instance, an experiment has shown that changing the extension of an .mp4 file to .pdf allowed for hiding it from the evidence tree generator, FTK imager, without applying any changes to its location [8].

#### 1) Least Significant Bit Encoding (LSB)

The principle of this technique is to find and replace the least significant bits in the image frames of the cover video. If each pixel in the gray carrier image consists of 8 bits for example, then the LSB of this pixel is replaced by a bit from the secret message. However, for an RGB (Red, Green and Blue) carrier image, each of the red, green, and blue components which are each of 8-bits lengths can be utilized for the same purpose by replay the LSB in them by bits of the secret message [9].

Thus the secret message to be hidden in the carrier image undergoes two processes, the first is conversion to ASCII code and the second is conversion to binary representation with 8 bits for each of the words.

#### 2) A Hash Based LSB Technique (HLSB)

Dasguptal et al. [10] reported a hash based LSB Techniques in spatial domain. It is a utilization of an algorithm portrayed with AVI (Audio Video Interleave) file as a cover medium. A video stream (AVI) composed of collection of frames and the secret data is concealed in these frames as payload. The information of the cover video (AVI) such as number of frames (n), frame speed (fp/sec), frame height (H) and frame width (W) are extracted from the header.

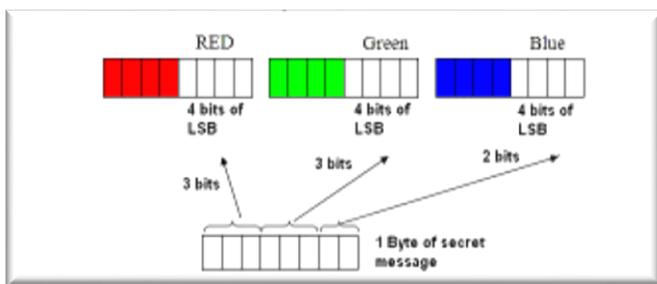


Fig. 1. Secret data embedded in RGB pixels of carrier frame [10].

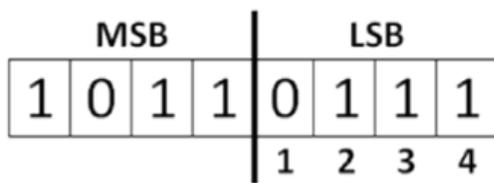


Fig. 2. The hiding positions of the four bits (4) available bits of LSB's side.

The cover video is then divided and separated into frames. The size of the message is irrelevant in video steganography due to the fact that the message could be embedded in multiple frames. The advised technique conceals 8 bits of secret data at a time in LSB of RGB pixel value of the carrier frames in 3, 3, 2 order respectively. Such that out of eight 8-bits of message 3-bits are inserted in R pixel, 3-bits are inserted in G pixel and remaining 2-bits are inserted in B pixel, as shown in Fig. 1. This distribution pattern is applied because the chromatic influence of the blue pixel color component to the human eye is more than that of the red and green pixel components.

The hash based LSB technique would also increase the payload. Moreover, the fact that the variation between the colours is small makes its detection by the human eye a very strenuous task. Using a hash function of this form, the hiding positions of the eight bits out of the four available bits of LSB as shown in Fig. 2 could be obtained

$$k = p \dots \dots \dots (1)$$

Where,  $k$  is LSB bit position within the pixel,  $p$  stands for the position of each hidden image pixel and  $n$  represents the number of bits of LSB which is 4 for the present case.

A stego video is then formed after concealing data in multiple frames of the carrier video by grouping these frames together. That video will be, used as normal sequence of streaming frames.

The reverse steps could be followed by the intended user to uncover and extract the secret data. The decoding process of a setgo video consists of breaking the video into frames again after going through the header information. The data of the secret message is regenerated, by utilizing the same hash function known to the intended user.

#### 3) Neighbourhood Pixel Information

Hossain et al. [11] proposed three effective steganography tools that use the neighborhood information to calculate the quantity of data which could be embedded in a cover image input pixel without causing a noticeable change. The smooth and complicated areas of an image are determined by the neighborhood relationship. According to this concept, small quantities of secret data are hidden in the smooth areas, and large quantities are hidden in the complicated ones. This whole concept is built on psycho visual repetition in grey scale digital images; the edged parts can withstand more change in comparison to the smooth ones.

### III. THE PROPOSED IMAGE AND AES INSPIRED STEGANOGRAPHY (IAIS)

#### A. Image and AES Inspired Steganography (IAIS) Design

In this paper, we will be introducing a new data hiding and encryption method in which we combine our previously developed AES Inspired Steganography (AIS) with image steganography. This combination would allow for an increased level of security and concealment of the hidden secret data. This Steganography approach will include the use of a table to specify the hiding location of the secret data. In general, this method consists of multiple steps of encryption applied to the secretly hidden message. First, the message is embedded into a

hex symbols carrier file, which is divided into embedding matrices and cipher key matrices according to varied patterns chosen by the communicating parties. This approach was discussed in detail in our previous paper on the AES Inspired Steganography [2]. The hex symbol carrier file as a whole is then embedded into the chosen cover image as shown in Fig. 3. Furthermore, throughout the steganography process, encryptions and rearrangements are continuously applied allowing for increased security measures.

**B. Image and AES Inspired Steganography (IAIS) Algorithm**

Initially, the communicating parties (i.e. sender and receiver) agree upon certain tables that will be used as keys for embedding and extracting of the secret message content. Secret content could be embedded at different locations in the image including the header, the tail or both or could be attached to the end of the image and embedding of AIS processed hex symbol within the image. The method of embedding the AIS encrypted text at each of these locations will be described next.

**C. Attachment of AIS processed hex symbol to the end of the image**

Once AIS is applied to a secret text, a cover image is chosen (e.g. Fig. 4(a)). The chosen picture is then converted to a hexadecimal file using WinHex. The created AIS hex symbol in then embedded in the carrier image by attaching it to the end of the hexadecimal format of the carrier image. A deciphering key is embedded as well in the hexadecimal file in the location agreed upon by the communicating parties. Afterwards, the resulting hexadecimal file is converted back to the image format which can be sent to the second communicating party Fig. 4(b). The second communicating party receives the cover image containing the hex symbol representation of the secret text and the decryption key. They key consists of two parts: 1) a hex symbol code book (in alphabetic codes); and 2) an index number determining the location and size of the hidden hex symbol in the image (Fig. 5).

**D. Embedment of AIS processed hex symbol in the image’s header, tail, or both**

Similarly, a cover image is chosen after the creation of the hex symbol form of the hidden secret text (Fig. 6(a), 7(a)) and is converted to a hexadecimal file using WinHex. The hex symbol is then embedded in the header or tail of the cover image’s hexadecimal format or both, simultaneously.

A deciphering key is embedded as well in the hexadecimal file in the location agreed upon by the communicating parties. The cover image carrying the hidden text and the key will then be converted back to the image format from the hexadecimal format and sent to the receiving party (Fig. 6(b), 7(b)).

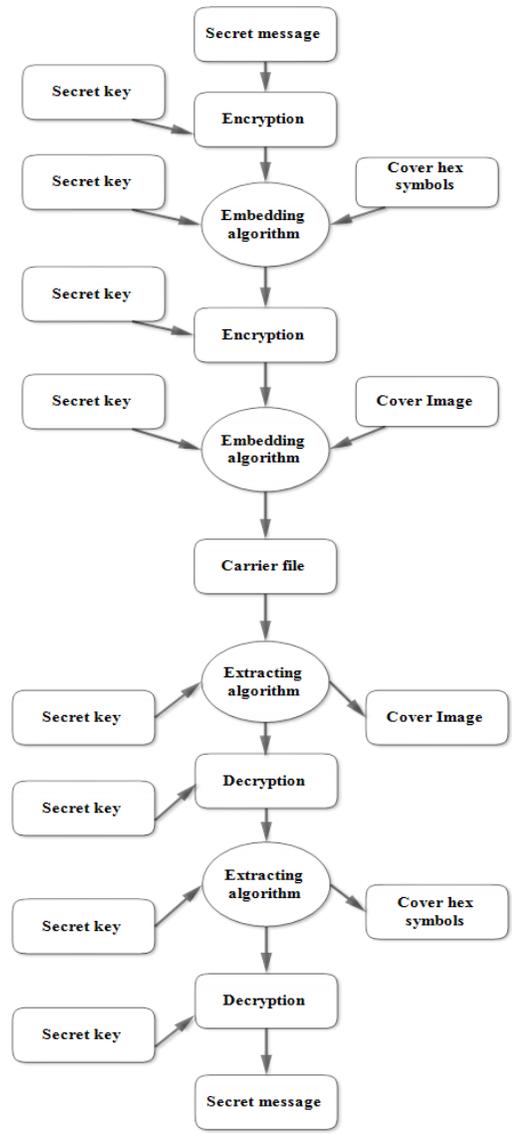


Fig. 3. The general Image and AES Inspired Steganography (IAIS) process.

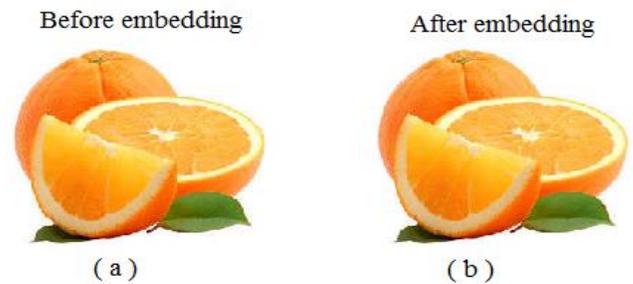


Fig. 4. Cover image before and after attaching the AIS encrypted text to the end of the image.

Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
16		00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F
Index	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
32		00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F
...	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	

Fig. 5. Example of the shared key index book.





Fig. 11. Snapshot of the arrangement (collage) of the chosen cover images.

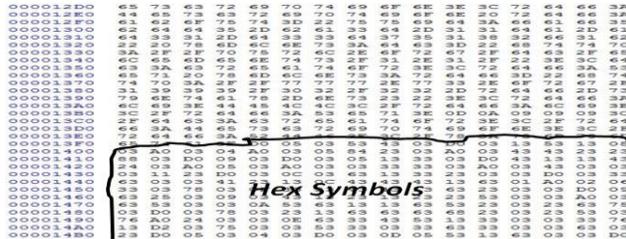


Fig. 12. Hex symbols in image.



Fig. 13. Carrier file containing hex symbols after the image steganography process.

## V. ANALYSIS AND DISCUSSION

### A. Robustness

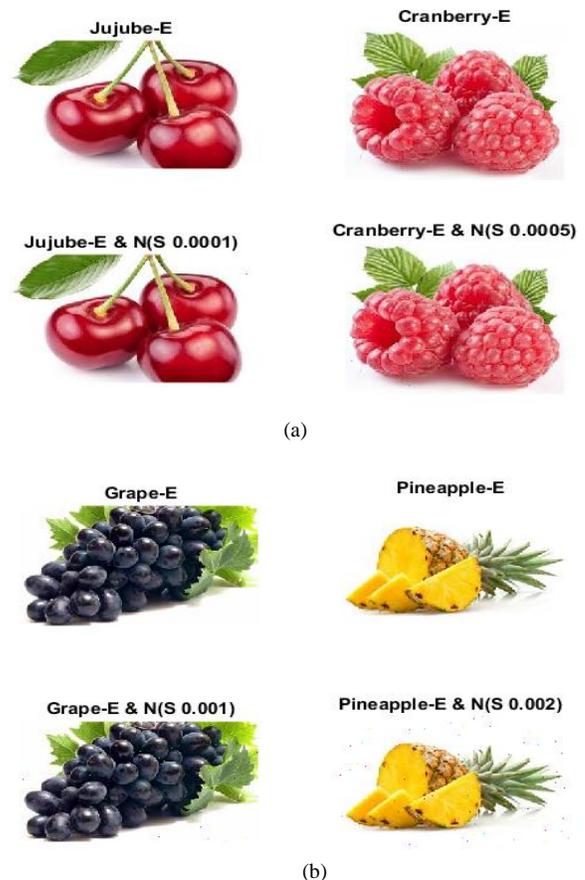
The stego-files have been found to be resistant against changes in size and content when compressed with WinRAR or ZIP file format and then processed for message extraction. This resistance indicates a kind of robustness against processing procedure that could be applied to the carrier file such as compression. Also, the robustness means the resistance to external effects such as noise or other image processing. Experiments of adding noise to the stego image were conducted as below. Two type of noise, namely Gaussian noise and Salt and Pepper noise were introduced to stego images and the mean square error (MSE) and peak signal to noise ratio (PSNR) were calculated for different percentages of noise. These results are listed in Fig. 14 and Fig. 15. The noise effect on the images of the data listed in Fig. 14 and 15 is shown in the Fig. 16 and 17 for both Salt pepper noise and Gaussian. The MSE and PSNR were also plotted in Fig. 18 and 19.

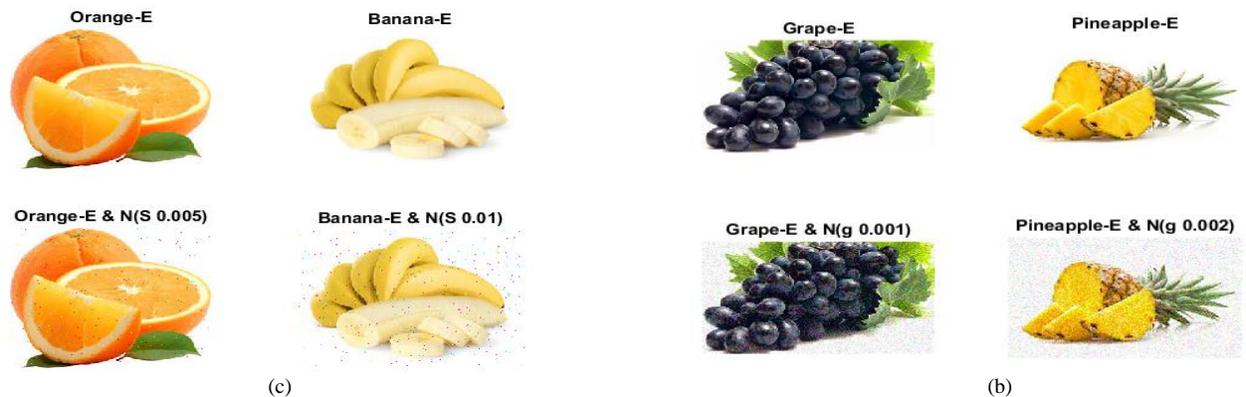
Num	Pic name	Noise percentage	SNR	PSNR	MSE
1	jujube-E	0.0001	42.4780	45.2887	1.9240
2	Cranberry-E	0.0005	34.6368	37.3139	12.0696
3	Grape-E	0.001	28.5604	33.1917	31.1826
4	Pineapple-E	0.002	29.2412	30.3633	59.8073
5	Orange-E	0.005	24.8986	27.0396	128.5646
6	Banana-E	0.01	22.7457	23.6880	278.1515
7	Apple-E	0.02	17.9092	20.8372	536.2466
8	Kiwi-E	0.03	16.7490	18.9780	822.7686
9	Pomegranate-E	0.06	13.9033	16.3126	1.5199e+03
10	Strawberries-E	0.1	11.8404	13.9782	2.6017e+03

Fig. 14. Calculated the (MSE) and (PSNR) for Salt & pepper noise.

Num	Pic name	Noise percentage	SNR	PSNR	MSE
1	jujube-E	0.0001	18.3941	21.2049	492.7141
2	Cranberry-E	0.0005	18.2903	20.9674	520.4081
3	Grape-E	0.001	16.3100	20.9412	523.5536
4	Pineapple-E	0.002	20.9047	22.0267	407.7655
5	Orange-E	0.005	19.3355	21.4765	462.8409
6	Banana-E	0.01	21.1186	22.0609	404.5679
7	Apple-E	0.02	18.4884	21.4164	469.2909
8	Kiwi-E	0.03	19.5140	21.7430	435.2949
9	Pomegranate-E	0.06	18.4390	20.8482	534.8790
10	Strawberries-E	0.1	17.6281	19.7659	686.2553

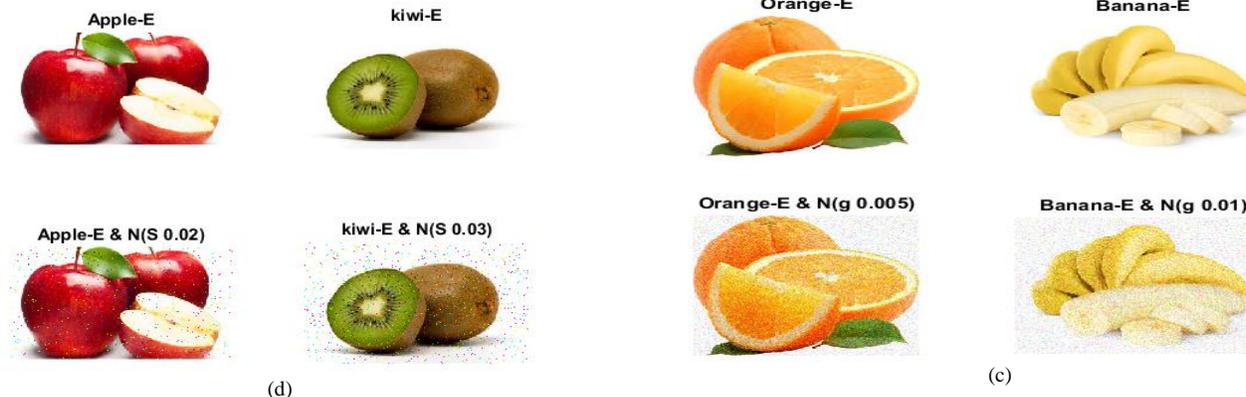
Fig. 15. Calculated the (MSE) and (PSNR) for Gaussian noise.





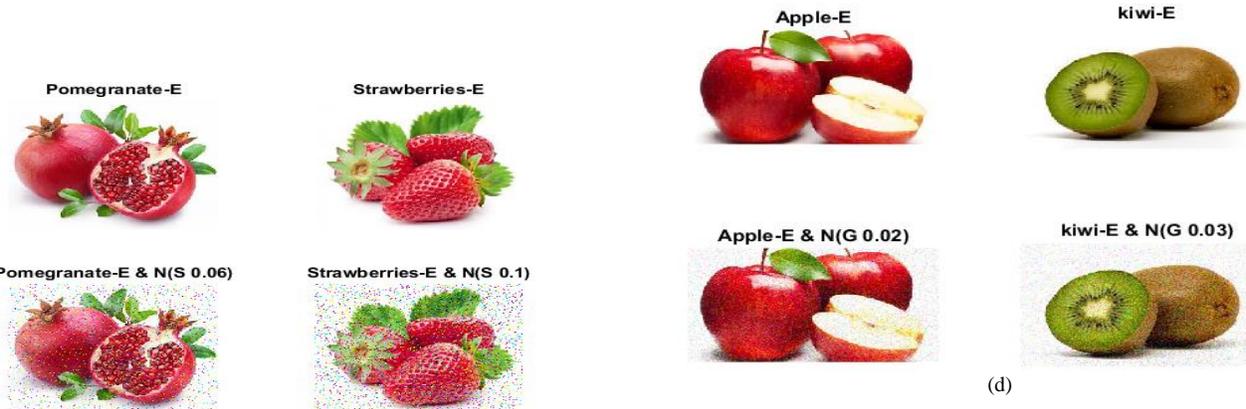
(c)

(b)



(d)

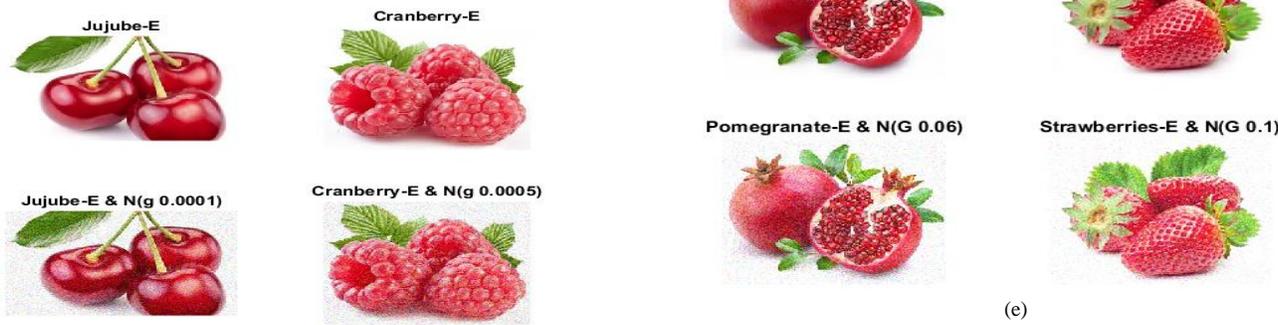
(c)



(e)

(d)

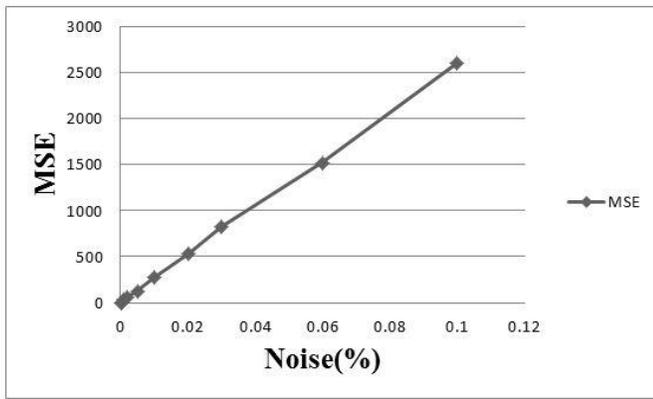
Fig. 16. The visual effect of salt and pepper noise on the images.



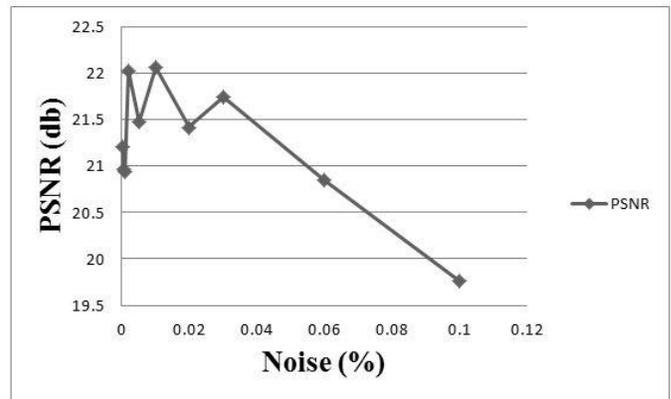
(a)

(e)

Fig. 17. The visual effect of Gaussian noise on the images.



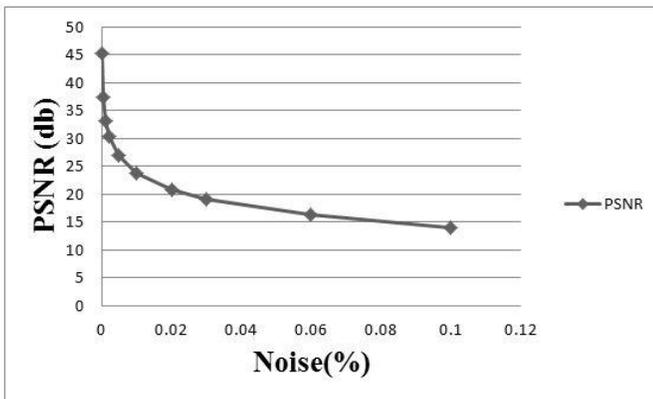
(a)



(b)

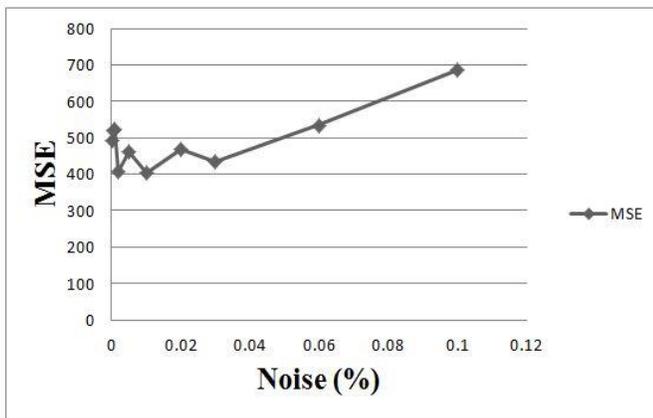
Fig. 19. The MSE and PSNR for different percentages of Gaussian noise.

To visualize the effect of adding the encrypted secret message of the HSA to the carrier image, the histogram for the images with and without the embedded message were compared and shown in Fig. 20. It was found that no difference was noticed in the histogram. This indicates a good advantage for the hex symbols technique.

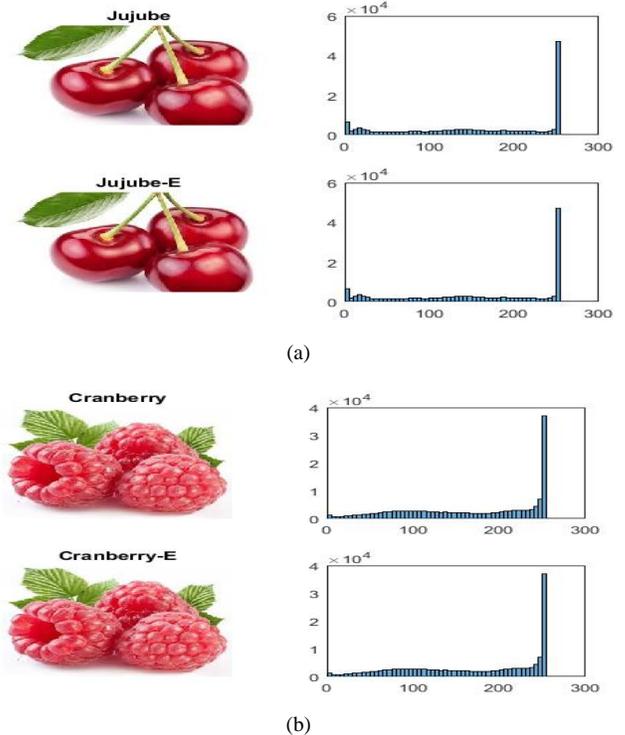


(b)

Fig. 18. The MSE and PSNR for different percentages of salt & pepper noise.



(a)



## VI. CONCLUSION AND FUTURE WORK

In this paper we proposed a new variation of steganography approaches in which we combined image steganography with our previously proposed AES Inspired Steganography (IAIS). This approach represents an upgrade of the obstacles' level against potential external breaches. Hence, this will allow for increasing the information security while transferring secret information through the internet or private networks.

Future developments could be conducted to increase the complexity and capacity of the steganography approach allowing handling larger amounts of secret information. Furthermore, the approach could be incorporated in different areas of application including medical information and records security.

### REFERENCES

- [1] T. Morkel, J. Eloff and M. Olivier, "An Overview of Image Steganography," *Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005)*, Sandton, South Africa, June/July 2005 (Published electronically) .
- [2] S. Abu Asbeh, S. Hammoudeh and A. Hammoudeh, "AES Inspired Hex Symbols Steganography for Anti-Forensic Artifacts on Android Devices", *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 7, no. 5, 2016.
- [3] S. Abu Asbeh, H. Al-Sewadi, S. Hammoudeh and A. Hammoudeh, "Hex Symbols Algorithm for Anti-Forensic Artifacts On Android Devices," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 7, no. 4, 2016.
- [4] E. Barker, "Recommendation for Key Management - Part 1: General", 4<sup>th</sup> Rev., *NIST Special Publication 800-57*, 2016.
- [5] S. Galbraith, *Mathematics of public key cryptography*, 1st ed. Cambridge: Cambridge University Press, 2012.
- [6] W. Stallings, *Computer Organization and Architecture; Designing for Performance*. 8th Ed., Prentice Hall, 2010.
- [7] W. Stallings. *Cryptography and Network Security: Principles and Practices*. 5th Ed, Prentice Hall, 2010.
- [8] A. Jain, and G. Chhabra, "Anti-Forensics Techniques: An Analytical Review", *IEEE Seventh International Conference on Contemporary Computing (IC3)*, pp. 412-418, 7-9 Aug. 2014.
- [9] A. Swathi and S. Jilani, "Video Steganography by LSB Substitution Using Different Polynomial Equations", *International Journal Of Computational Engineering Research*, vol. 2, no. 5, pp. 1620-1623, 2012.
- [10] K. Dasgupta, J. Mandal and P. Dutta, "Hash Based Least Significant Bit Technique for Video Steganography (HLSB)", *International Journal of Security, Privacy and Trust Management (IJSPTM)*, vol. 1, no. 2, 2012.
- [11] M. Hossain, S. Al Haque, and F. Sharmin, "Variable Rate Steganography in Gray Scale Digital Images Using Neighborhood Pixel Information", *The International Arab Journal of Information Technology (IAJIT)*, vol. 7, no. 1, pp.34-38, 2010.

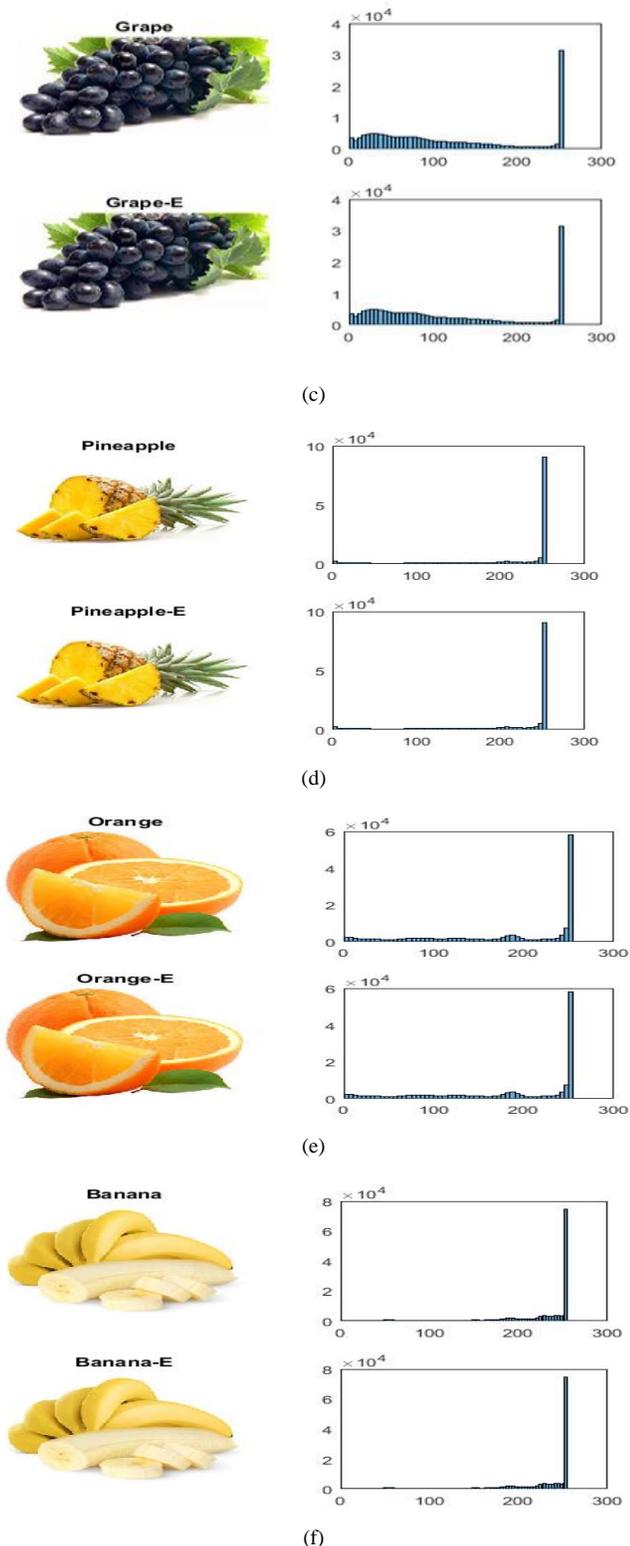


Fig. 20. Histogram comparisons before and after embedding.

# Network Traffic Classification using Machine Learning Techniques over Software Defined Networks

Mohammad Reza Parsaei, Mohammad Javad Sobouti, Seyed Raouf khayami, Reza Javidan  
Department of Computer Engineering and Information Technology  
Shiraz University of Technology (SUTECH), Shiraz, Iran

**Abstract**—Nowadays Internet does not provide an exchange of information between applications and networks, which may result in poor application performance. Concepts such as application-aware networking or network-aware application programming try to overcome these limitations. The introduction of Software-Defined Networking (SDN) opens a path towards the realization of an enhanced interaction between networks and applications. SDN is an innovative and programmable networking architecture, representing the direction of the future network evolution. Accurate traffic classification over SDN is of fundamental importance to numerous other network activities, from security monitoring to accounting, and from Quality of Service (QoS) to providing operators with useful forecasts for long-term provisioning. In this paper, four variants of Neural Network estimator are used to categorize traffic by application. The proposed method is evaluated in the four scenarios: feedforward; Multilayer Perceptron (MLP); NARX (Levenberg-Marquardt) and NARX (Naïve Bayes). These scenarios respectively provide accuracy of 95.6%, 97%, 97% and 97.6%.

**Keywords**—Software defined networks; openflow; traffic classification; neural network; multilayer perceptron

## I. INTRODUCTION

SDN is a new paradigm in telecommunications and computer networks. The main goal of SDN is to meet challenges existing in IP-based networks, such as complex management. In today's networks, administrators must apply many overwhelming changes to the network configurations in case of a little change in network policies, rules or topology, testing new protocols, to have a dynamic network management [1]-[4]. SDN as a comprehensive concept separates data plan (which is responsible for forwarding data packets) and control plan (which is responsible for routing, traffic engineering, and management policies) to confront limitations and challenges of today's networking [5]-[8]. Fig. 1 shows SDN architecture. OpenFlow (OF) protocol is one of the most important and practical communication protocols, which enables controller to interact with the network switches. This protocol is a standard interface that is used mostly in SDN. OF switches contain one or multiple flow tables with flow entries. Each entry consists of paired rules and actions. The tables are filled by the controller. Each rule consists of fields related to headers data, such as source and destination MAC and IP addresses, port numbers and other necessary information. Each action determines instruction(s) to be executed on the packet in order to match the entry's rule [9]-[10].

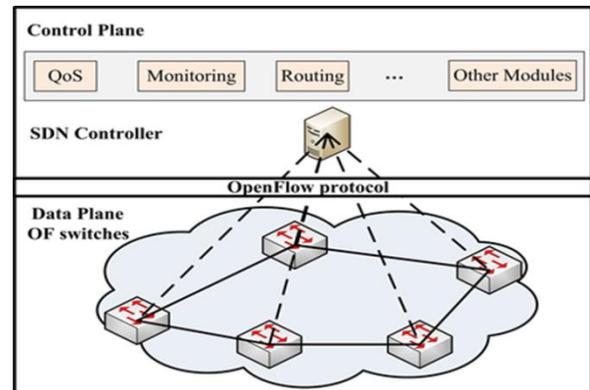


Fig. 1. SDN architecture.

Separation of data plan and control plan, gives ability to network administrators to make programmable policies and easily manage data plan via the controller [11]. SDN also makes it easy to have a dynamic management, configuration, troubleshooting and even testing new protocols and ideas in the network without troubles [12]. An important case in network management for having high availability and efficiency is traffic classification. There are methods for applying traffic classification in networks [13]:

- Using port numbers to determine application and application layer protocols. However, these methods are not completely accurate.
- Deep Packet Inspection (DPI) is used. These methods have high accuracy, however, there are some issues regarding its implementation while dynamic ports and encrypted traffics are not supported in current networks yet. It also causes high overhead to the system and violates user privacy.

These methods have their own problems, so, researches have been recently focusing on machine learning techniques, which take advantage of statistical properties for traffic classification.

Although there are many challenges in current networks for traffic classification, global view of controllers in SDN improves network management while its concept is simple and easy to use for extracting statistical data of network traffic from switches [14].

Hanigan et al. [15] used traffic classification methods based on DPI to inspect flows over SDN to distinguish application protocols in runtime. Their goal is to enable controller to distinguish and isolate different application flows, managing and programming flows to guaranty QoS for delay sensitive applications. When the network is loaded, a great part of the controllers' processing resources must be dedicated to DPI tools. Thus, the performance of the entire network is affected.

Arsalan et al. [16] proposed a framework to determine the application type of existing flows in a wireless network, which consist of several mobile devices connected to an OF switch. In control plan, a machine learning-based trainer receives the information. On the other hand, the OF switch gathers the properties of different flows and sends them to the control layer for creating a model for application layer recognition. After model creation, when a host joins the network, the OF switch sends the device flows properties to the traffic classification model, based on the machine learning technique. Then, the application of source flow is determined. The traffic classification model is based on C5.0 decision tree algorithm.

Jang et al. [17] proposed a method that uses flows properties, gathered in a dataset, as K-means algorithm input, in learning phase, for clustering. These clusters are used to implement a traffic classification model. The clusters with similar features are aggregated based on the information obtained from the content of the packets. Although the accuracy rate of this method is 89 percent, it reduces the need of investigating the packet contents and accurate diagnosis of encrypted packets. Most researches present traffic classification on a set of statistics from stored flows in an offline manner. The high time complexity and processing overhead are two challenges of online traffic classification. Current approaches also cause an overwhelming overhead on the system. The goal of this paper is to classify the traffic over SDN using information in the header of packets received from OF switches and statistics in the controller. By considering protocol capabilities on extracting flows statistics and neural networks variant such as feedforward, MLP, NARX (Levenberg-Marquardt) and NARX (Naïve Bayes), a framework for online traffic classification based on application layer protocol is proposed. The overall accuracy of this model in traffic classification for feed-forward, MLP, NARX (Levenberg-Marquardt) and NARX (Naïve Bayes) algorithms with value of 95.6%, 97%, 97% and 97.6%, respectively.

The highest accuracy of the previous methods was 94% [18] but the accuracy of the proposed method is 97.6%. Advantages of this method over current methods are low processing overhead, low network overhead and low runtime execution. The following sections of this paper are as follows: Section 2 presents the proposed method for online traffic classification in SDN. Section 3 provides the implementation and performance evaluation of proposed method. Finally, discuss and analyze the results in Section 4.

### I. THE PROPOSED METHOD

In this paper, a new method is used for classification of network traffic proportional to SDN architecture.

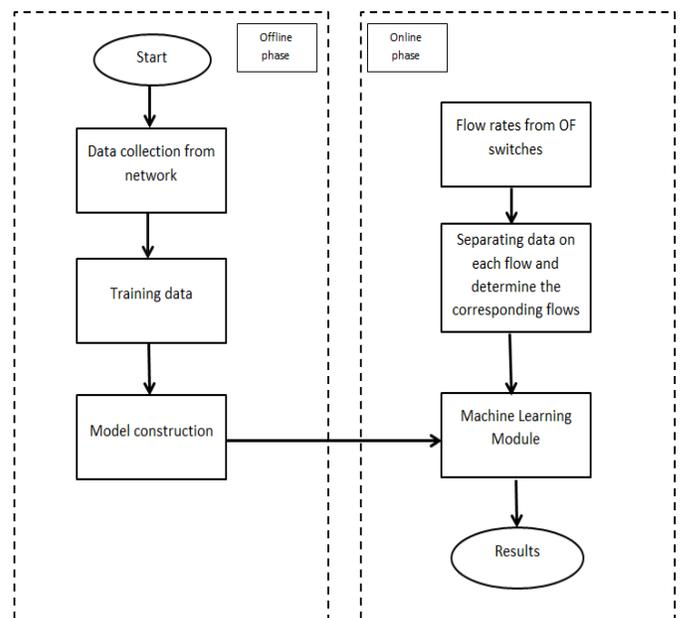


Fig. 2. The outline of the proposed method.

In such networks, all switches are connected to a central controller that may be lower cost than current networks. The protocol of each flow can be identified by classification of traffic based on application layer in the level of control. Fig. 2 shows the outline of the proposed approach.

The proposed method consists of offline and online phases that are as follows:

#### A. Offline phase

In offline phase, making data collection and model classifier are discussed. This means that floodlight uses a web-based graphical interface to connect users to the controller. We use below URI to receive raw data of the required set of training data to get statistics on all existing traffic flows in the switches:

<http://localhost:8080/wm/core/switch/All/flow/Json>.

Information is extracted for each flow by the mentioned URI. The structure of API data exchange is JSON. They are identifiable after calling up information in the browser. The flows are separated for using of massive data that obtained from all flows in switches.

For this purpose used five attribute (IP source, IP destination, the source port, the destination port, transport layer protocol), which specifies a unique flow. In the next step, after separating flows from each other, associated return flows with them that are constituent a 2-way flow are combined and done pre-processing on them for making a sample (a raw of sets of training data). Eventually, each row of the training data set that represents a two-way flow, contains the characteristics (see Table 1). Last item the APPpro (application layer protocol) represents a class variable in the training data set that is completed manually.

TABLE I. CHARACTERISTICS ASSOCIATED WITH EACH FLOW IN TRAINING DATA SET

Attribute	Description
srcIp	Source IP
dstIp	Destination IP
SrcPort	Transport Layer port in source
DstPort	Transport Layer port in destination
reqPro	Transport layer protocol , flow
respro	Transport layer protocol , backflow
reqAvgSz	The average size of the flow pockets
resAvgSz	The average size of the backflow pockets
reqPktperSec	Mean number of pockets per second on the flow
resPktperSec	Mean number of pockets per second on the backflow
reqBytperSec	Mean number of bytes per second on the flow
resBytperSec	Mean number of bytes per second on the backflow
APPpro	Application layer protocol

A classification model is used after classification algorithm on training data set, which is used to create traffic classification module [19-21] for floodlight controller.

### B. Online phases

In online phase, ML module that is added to floodlight controller classifies the network traffic operation by the help of developed model in offline phase. This means that received statics flows in the switch and the application of layer protocol obtains each of them. The results obtained this module and management of bandwidth, security and management issues in order to supply QOS goals provided for javaAPI and restAPI. They order both of them are used to communicate with other modules and applications. ML module evaluated in the four scenarios which following algorithm is used:

#### 1) Feed-forward Neural Network

In feed-forward neural network, connections between units do not form a cycle so it is different from recurrent network. Feed-forward network is the first and most simple type of neural network algorithms. The flow of information always moves forward from input to output. A supervised technique called backpropagation is used to improve its performance. It propagates backward from output to input in the network, decreases errors and optimizes performance by correcting the weight of edges, connected to nodes. The weights can be corrected by Gradient Descent method using (1) for calculating the change of each edge's weight.

$$\Delta W_i = -\eta \partial E / \partial W_i \quad (1)$$

In this equation  $\eta$  is the learning rate which its value is considered equal to 0.1. Also, the expected value is obtained from (2), in which  $d$  is the target value and  $y$  is the perceptron's output.

$$\begin{aligned} \text{Expected value} &= \partial / \partial w_i \ 1/2 \ S_d (t_d - o_d)^2 \\ &= \partial / \partial w_i \ 1/2 \ S_d (t_d - S_i w_i x_i)^2 \quad (2) \\ &= S_d (t_d - o_d) (-x_i) \end{aligned}$$

#### 2) Multilayer Perceptron (MLP)

MLP is a type of neural network, which maps a set of input data to one or more output, based on learning from previous samples. Because of its strong nonlinear approximation behavior, MLP is the most useful model in neural networks and is used almost in every scientific field. An MLP is consisting of multi layers of nodes in a directive graph, in which each layer is fully connected to the next layer. MLP also uses backpropagation for training network. MLP is the modified version of perceptron and can recognize nonlinear data [22]. Using gradient descent, find changes in each weight according to (3), where  $v_i$  is the output of the previous neuron and is calculated by (4):

$$\Delta w_{ji}(n) = -\eta \frac{\partial \varepsilon(n)}{\partial v_j(n)} y_i(n) \quad (3)$$

$$y(v_i) = \tanh(v_i) \text{ and } y(v_i) = (1 + e^{-v_i})^{-1} \quad (4)$$

Here  $V_i$  is the weighted sum of the input synapses.  $\varepsilon(n)$  in (3) is calculated through (5) and the error in output node  $j$  in the  $n$ th training example is calculated by (6), where  $d$  is the target value and  $y$  is the value produced by the perceptron:

$$\varepsilon(n) = \frac{1}{2} \sum_j e_j^2(n) \quad (5)$$

$$e_j(n) = d_i(n) - y_j(n) \quad (6)$$

#### 3) Non-linear Autoregressive Exogenous Multilayer Perceptron (NARX)

The NARX model is generally used for predicting time series by approximation of nonlinear relationships between exogenous variables and the predictor variable, as defined in (7):

$$\begin{aligned} y(t) &= f(x(t-1), x(t-2), \dots, x(t-d)); \\ y(t-1), y(t-2), \dots, y(t-d) \end{aligned} \quad (7)$$

$y(t)$  is the predictor variable and  $x(t)$  denotes the Exogenous time series. Usually function  $f$  is a nonlinear polynomial. For modeling function  $f$  in NARX, it is also possible to create a dynamic MLP network by assuming in time  $t$ ,  $d$  is the previous variable of predictor variable and the predictor variable should be accessible. This configuration is based on delay and is without feedback, called open loop. This model is used for one-step-ahead predictions, because the estimation is based on previous knowledge of real past values for target series and is not based on prediction that produces errors in results. In this model two functions are used for learning:

- Levenberg-Marquardt Algorithm

One of the learning functions of this method is retrieved from Levenberg-Marquardt Algorithm. This algorithm is a way to find the minimum of a nonlinear polynomial function and is a standard method for solving Minimum Square of nonlinear function problem. This algorithm is used to minimize square curve fitness problem. The  $\beta$  parameter in the curve model of  $f(x, \beta)$  is from experimental data set of dependent and

independent variables  $(x_i, y_i)$ , which sum of their square of derivation is minimum, as shown in (8):

$$\hat{\beta} = \arg \min S(\beta) \equiv \arg \min \sum_{i=1}^m [y_i - f(x_i, \beta)]^2 \quad (8)$$

In each step of iteration, the  $\beta$  parameter vector is replaced with a new approximate value of  $\beta + \delta$ . For  $\delta$  calculation, the functions  $f$  are estimated by linearization as in (9). In this equation  $J_i$  is the gradient of function  $f$  with respect to  $\beta$  which calculated as in (10):

$$f(x_i, \beta + \delta) \approx f(x_i, \beta) + J_i \delta \quad (9)$$

$$J_i = \frac{\partial f(X_i, \beta)}{\partial \beta} \quad (10)$$

The sum of squares  $S(\beta)$  at its minimum has a zero gradient with respect to  $\beta$ . Equation (11) shows the first approximation of  $f(x_i, \beta + \delta)$  and its vector notations are in (12):

$$S(\beta + \delta) \approx \sum_{i=1}^m (y_i - f(x_i, \beta) - J_i \delta)^2 \quad (11)$$

$$\begin{aligned} S(\beta + \delta) &\approx \|y - f(\beta) - J\delta\|^2 = \\ &[y - f(\beta) - J\delta]^T [y - f(\beta) - J\delta] = \\ &[y - f(\beta)]^T [y - f(\beta)] - \\ &[y - f(\beta)]^T J\delta - (J\delta)^T [y - f(\beta)] + \\ &\delta^T J^T J\delta = [y - f(\beta)]^T [y - f(\beta)] - \\ &2[y - f(\beta)]^T J\delta + \delta^T J^T J\delta. \end{aligned} \quad (12)$$

Equation (13) is obtained from setting the derivation of  $S(\beta + \delta)$  to zero:

$$(J^T J)\delta = J^T [y - f(\beta)] \quad (13)$$

In this equation,  $J_i$  is the  $i$ th row of the Jacobian matrix and  $Y_i$  and  $f(x_i, \beta)$  are the  $i$ th component of  $Y$  and  $f$  vectors respectively. Levenberg's contribution is to replace this equation by a "damped version" as show in (14):

$$(J^T J + \lambda I)\delta = J^T [y - f(\beta)] \quad (14)$$

Besides, Marquardt replaced the identity matrix "I" with diagonal matrix consist of diagonal elements of  $J^T J$ , resulting in Levenberg-Marquardt algorithm as in (15):

$$[J^T J + \lambda \text{diag}(J^T J)]\delta = J^T [y - f(\beta)] \quad (15)$$

- Naïve Bayes

Another learning function is from Naïve Bayes classifiers family, which do classifies using Bays probability theorem. Naïve Bayes is a common technique for creating classifiers. One of its positive points is its efficiency in solving probability problems. It is also very flexible and needs only a few learning

data for estimating the required parameters [23]. The mathematical form of Bayes theorem is as (16):

$$p(C_k | x) = \frac{p(C_k)p(x | C_k)}{p(x)} \quad (16)$$

## II. PERFORMANCE EVALUATION

A SDN network with one software switch, one controller and two hosts, is configured for our experiments. In order to implement software defined networks, OVS (open v Switch) was considered as OF switch and Floodlight was considered as the controller. Open v Switch is an open source; apache 2.0-licensed virtual switch which is developed under Linux kernel. Floodlight is an open source, apache licensed controller, which developed by using Java platform. A topology as shown in Fig. 3 is set for performance evaluation.



Fig. 3. Scenario topology which evaluated in the laboratory.

TABLE II. SYSTEM DESCRIPTION

Controller and hosts	
CPU	Core i7
RAM	16 GB
OS	Ubuntu 14.10
Controller version	Floodlight v 9.0
Switch	
CPU	Core i5
RAM	8 GB
OS	Ubuntu 14.10
OVS version	2.0.1

There are four systems in this figure, from left to right: the first system is an OF switch (it is converted to OF switch by installing ovs on it). The second and third systems are the first and second clients. The last system is the controller. The systems' configurations are shown in Table 2.

After setting up the topology, the host establishes FTP and HTTP connections for downloading files and video streams, text messaging, peer to peer connection with BitTorrent client, downloading and uploading file parts to/from other hosts simultaneously and aggregates statistics related to these flows.

So, the case study protocols for this scenario are FTP, HTTP, instant messaging, video streaming and peer to peer protocol. Sampling of the flows is done 20 times in 1 second intervals.

In all scenarios, sampling from switch statistics (extracting required data from switch's flow tables) is done 6000 times to collect about 65000 data records. After data aggregation and preprocessing of flows statistics, a data set with 600 records is obtained.

After that the different flow class variable is manually set in the learning data set, and finally the data is converted to CSV format, for creating traffic classification model. Final Model is converted to java code for using as a traffic classifier module in online phase. Results in details are provided in the Fig. 4 to 7.

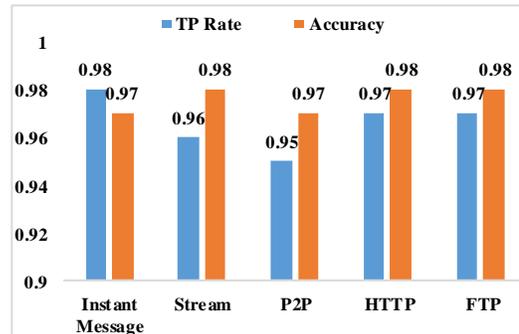


Fig. 7. The result of the NARX (Naïve Bayes) algorithm.

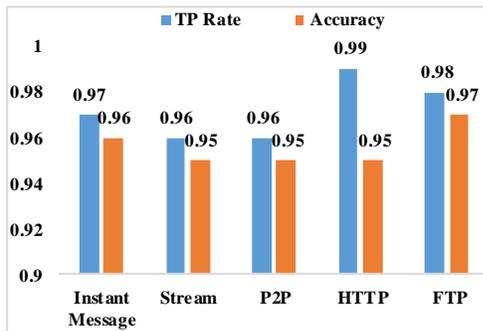


Fig. 4. The result of the Feedforward algorithm.

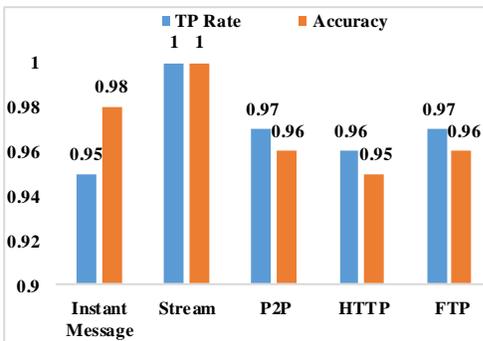


Fig. 5. The result of the MLP algorithm.

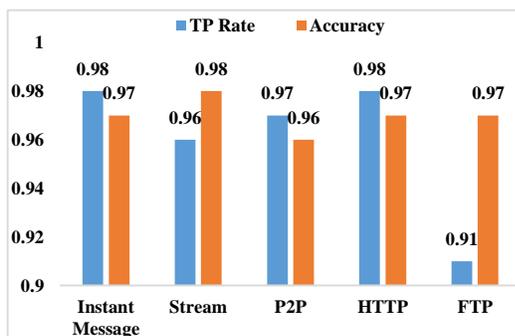


Fig. 6. The result of the NARX (Levenberg-Marquardt) algorithm.

### III. CONCLUSION

The proposed method in this paper is used for application recognition of flows resources with the help of SDN and data mining techniques based on machine learning. Applying traffic classification techniques to the OF network makes the network be application-aware, and enables the network to know flow's requirements. Due to maintaining a global view of the network, the controller could dynamically allocate bandwidth to flows on demand and thus improve their QoS and the analysis and prediction of traffic patterns in network make the controller further optimize resource allocation. This method mainly focused on minimizing controllers' processing overhead and network traffic overhead for network traffic classification. The accuracy of tested class variables in our experiments for feedforward, MLP, NARX (Levenberg-Marquardt) and NARX (Naïve Bayes) algorithms are 95.6%, 97%, 97% and 97.6% respectively. The highest accuracy of the previous methods was 94% but the accuracy of the proposed method is 97.6%. Also, the proposed method does not impose any processing overhead to the controller because unlike the base method, packets' contents are not checked. Our on-going and future works include implementations on different device platforms (iOS, Windows, Linux) and detection of flows belonging to a new application which is not part of the trained classifier.

### REFERENCES

- [1] C. Trois, M. D. Del Fabro, L. C. de Bona, M. Martinello, "A survey on SDN programming languages: toward a taxonomy," IEEE Communications Surveys and Tutorials, vol. 18, no. 4, pp. 2687-2712, 2016.
- [2] B. A. A. Nunes, M. Mendonca, X. N. Nguyen, K. Obraczka, T. Turletti, "A survey of software-defined networking: Past, present, and future of programmable networks," IEEE Communications Surveys and Tutorials, vol. 16, no. 3, pp. 1617-1634, 2014.
- [3] T. Huang, F. R. Yu, C. Zhang, J. Liu, J. Zhang, J. Liu, "A Survey on Large-scale Software Defined Networking (SDN) Testbeds: Approaches and Challenges," IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 891-917, 2016.
- [4] T. Benson, A. Akella, D. A. Maltz, "Unraveling the Complexity of Network Management," In NSDI, pp. 335-348, 2009.
- [5] S. Rowshanrad, M. R. Parsaei, M. Keshtgari, "IMPLEMENTING NDN USING SDN: A REVIEW ON METHODS AND APPLICATIONS," IIUM Engineering Journal, vol. 17, no. 2, pp. 11-20, 2016.
- [6] M. R. Parsaei, R. Javidan, A. Fatemifar, S. Einavipour, "Providing Multimedia QoS Methods over Software Defined Networks: A Comprehensive Review," International Journal of Computer Applications, vol. 168, no. 9, pp. 55-59, 2017.

- [7] A. Mendiola, J. Astorga, E. Jacob, M. Higuero, "A survey on the contributions of Software-Defined Networking to Traffic Engineering," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 918-953, 2017.
- [8] M. Karakus, A. Durresi, "Quality of Service (QoS) in Software Defined Networking (SDN): A survey," *Journal of Network and Computer Applications*, vol. 80, pp. 200-218, 2016.
- [9] M. Dusi, R. Bifulco, F. Gringoli, F. Schneider, "Reactive logic in software-defined networking: Measuring flow-table requirements," *IEEE International Conference on Wireless Communications and Mobile Computing (IWCMC)*, pp. 340-345, 2014.
- [10] F. Hu, Q. Hao, K. Bao, "A survey on software-defined network and openflow: From concept to implementation," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2181-2206, 2014.
- [11] H. Kim, N. Feamster, N. "Improving network management with software defined networking," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 114-119, 2013.
- [12] M. R. Parsaei, S. H. Khalilian, R. Javidan, "A Comparative Study on Fault Tolerance Methods in IP Networks versus Software Defined Networks," *International Academic Journal of Science and Engineering*. Vol. 3, no. 4, pp. 146-154, 2016.
- [13] T. J. Parvat, P. Chandra, "A Novel approach to deep packet inspection for intrusion detection," *Procedia Computer Science*, vol. 45, pp. 506-513, 2015.
- [14] N. Williams, S. Zander, G. Armitage, "A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 5, pp. 5-16, 2006.
- [15] H. Cui, Y. Zhu, Y. Yao, L. Yufeng, Y. Liu, "Design of intelligent capabilities in SDN," *IEEE International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE)*, pp. 1-5, 2014.
- [16] Z. Arslan, A. Alemdaroglu, B. Canberk, "A traffic-aware controller design for next generation software defined networks," *IEEE International Conference on Communications and Networking (BlackSeaCom)*, pp. 167-171, 2013.
- [17] J. Zhang, Y. Xiang, W. Zhou, Y. Wang, "Unsupervised traffic classification using flow statistical properties and IP packet payload," *Journal of Computer and System Sciences*, vol. 79, no. 5, pp. 573-585, 2013.
- [18] Z. A. Qazi, J. Lee, T. Jin, G. Bellala, M. Arndt, G. Noubir, "Application-awareness in SDN," *ACM SIGCOMM computer communication review*, vol. 43, no. 4, pp. 487-488, 2013.
- [19] M. R. Parsaei, S. M. Rostami, R. Javidan, "A Hybrid Data Mining Approach for Intrusion Detection on Imbalanced NSL-KDD Dataset," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 6, pp. 20-25, 2016.
- [20] S. S. Parsa, M. Sourizaei, M. M. Dehshibi, R. E. Shateri, M. R. Parsaei, "Coarse-grained correspondence-based ancient Sasanian coin classification by fusion of local features and sparse representation-based classifier," *Multimedia Tools and Applications*, vol. 76, no. 14, pp. 15535-15560, 2017.
- [21] A. Nabaei, M. Hamian, M. R. Parsaei, R. Safdari, T. Samad-Soltani, H. Zarrabi, A. Ghassemi, "Topologies and performance of intelligent algorithms: a comprehensive review," *Artificial Intelligence Review*, pp. 1-25, 2016. doi:10.1007/s10462-016-9517-3.
- [22] H. Komijani, S. Rezaei Hassanabadi, M. R. Parsaei, S. Maleki, "Radial Basis Function Neural Network for Electrochemical Impedance Prediction at Presence of Corrosion Inhibitor," *Periodica Polytechnica Chemical Engineering*, vol. 61, no. 2, pp. 128-132, 2017.
- [23] M. R. Parsaei, R. Taheri, R. Javidan, "Perusing the effect of discretization of data on accuracy of predicting Naïve Bayes algorithm," *Journal of Current Research in Science*, (1), pp. 457-462, 2016.

# A New Approach for Leukemia Identification based on Cepstral Analysis and Wavelet Transform

Amira Samy Talaat Abou Taleb  
Computers and Systems Department  
Electronic Research Institute, Cairo, Egypt

Amir F. Atiya  
Department of Computer Engineering  
Cairo University, Giza, Egypt

**Abstract**—This paper implements a new leukemia identification method which depends on Mel frequency cepstral coefficient (MFCC) feature extraction and wavelet transform. Leukemia identification is a measurement of blood cell features for detecting the blood cancer of a patient. Blood cell feature extraction is based on transforming the blood cell two dimensional (2D) image into one dimensional (1D) signal and thereafter extracting MFCCs from such signal. Furthermore, discrete wavelet transform (DWT) of the 1D blood cell signals are used for extracting extra MFCCs features to assist the identification procedure. In addition, Wavelet transform with denoising is used to reduce noise and increase classification accuracy. Feature matching/classification of the blood cell to be a normal cell or leukemia cell is performed in the proposed method using five different classifiers. Experimental results of leukemia identification method show that the proposed method is very good with wavelet transform and robust in the presence of noise.

**Keywords**—MFCC; feature extraction; classification; identification system; leukemia

## I. INTRODUCTION

The probability of recovery of acute lymphocytic leukemia patient can be increased by the early identification of its symptoms. Leukemia (cancer) is a malignant disease seen in people of any age groups either in children or adults but usually affects people in their 50s and 60s.

In the literature, there is a lot of work for leukemia recognition based on many approaches like gene expression analysis [1] and holographic microscope images [2].

Artificial intelligent methods are based on automated systems that can speed up identification and make it much easier, in addition, the amount of data analyzed are higher moreover increase the classification accuracy specially in telemedicine applications. Many prediction methods used for analysis and classification of leukemia like KNN algorithm [3], other prediction methods use endoscopic images technique [4] and image processing techniques [5].

This paper presents a fully automatic method for leukemia identification that classifies blood images to know if the blood cell is normal or leukemia (cancer) cell.

Leukemia images identification is very important for diagnosis and therapy of cancer patients.

The proposed leukemia identification method is based on transforming blood cell image into 1D signals and executing the same processes performed on speech signals. The speech

signal operation used the Mel-frequency cepstral coefficients (MFCCs) for feature extraction fused with discrete wavelet transform (DWT) and signals denoising.

MFCCs are applied in speech recognition methods and its values are not very powerful in existence of additive noise [6], and so we propose leukemia identification system as an application of this idea by transforming the leukemia image 2D object into 1D signals and executing the same processes performed on speech signals.

Lately, wavelet based features are employed in various types of research. The DWT has good representation for time and frequency and can be used for multi dimensional localized contribution in time and frequency dominion for the signal of interest. moreover, wavelet denoising is a technique that can be used for reduction of noise from the speech signal [7], [8]. In our algorithm, we combine wavelet transform of the image with features extracted from MFCCs to assist in achieving a better recognition rate.

Five various classification techniques are used as classifiers in the proposed algorithm for the leukemia recognition method. Classification techniques are Radial SVM, Neighborhood Component Analysis classifier (NCA), Naive Bayes classifier, NaiveBayes Kernel classifier, and Quadratic Linear Discriminant Analysis (LDQ) classifier.

The rest of the paper is ordered as follows: An overview of leukemia, problem statement, and a brief survey of the current research area in this field are explained in Section 2. The process of extracting features existing in leukemia image using MFCCs is discussed in Section 3. The proposed leukemia recognition system, discrete wavelet transform, wavelet denoising are summarized in Section 4, feature matching (classification) is discussed in Section 5, Section 6, explains the experimental results and discussion. Finally, Section 7 summarizes the concluding remarks and future work.

## II. RECOGNITION SYSTEM

The process of identification of the leukemia recognition system consists of four phases: feature extraction, a training phase followed by a testing phase, and classification [9], [10]. Only useful information of the object is kept in feature extraction process. One of the most familiar methods that are used for feature extraction is MFCC [11]. MFCC perform with frames of the data so it uses 1D speech or voice signals. Training and testing steps both of them contain feature extraction technique. In the training step, each signal is showed

using a set of training data. Features are only the characteristic information of the signal and unnecessary info is stripping away. While in the testing step, feature extraction is also used and the resulting info is compared to the models in the database of the leukemia images to allow the unknown image to be identified. Finally, the classification process is performed to locate the exact signal corresponding to the leukemia image or normal blood cell image, therefore the system model is built and feature matching for testing the effectiveness of this model is performed by implementing a set of testing data to be compared with the stored features in the database. In classification phase, each unknown image is shaped by using a set of data samples in the training step, where a set of feature vectors is produced and kept in a database by deleting all needless information in the training samples keeping only the distinctive information to construct image models. When some unknown leukemia sample arrives, a mapping is made by the pattern matching techniques to match the features of the unknown sample to identify leukemia class [9], [12].

Feature extraction is a very vital step for recognition of unknown images. The only helpful data is selected in feature extraction that describes the signal and undesirable data is excluded. MFCC is a famed and excellent method for feature extraction from a speech signal [13], [14] that can be also used for face, gesture, palm print, satellite and iris image identification [15]-[19].

Thus this method is one of the finest techniques for feature extraction, mainly for automatic speech and speaker recognition system.

### III. GENERATION OF COEFFICIENTS USING MFCC

MFCC technique selects features from a given image. In case of voice identification, we obtain features by the following steps: first, the one dimensional signal is separated into minor frames or segments to make its statistical specification fixed, then frequencies are suppressed by windowing at boundaries and increase its center frequencies, the signal is transformed to frequency domain by FFT, Mel-Scale determine the space and size between each filter, then log will normalize signal after making discrete cosine transform (DCT), finally we get MFCC factors which are the last step in feature extraction process which are the characteristic information of the image. Now matching is made between MFCC coefficients of the given sample and the dataset sample to recognize and validate the blood cell if it is normal or a leukemia cell. At the beginning, the 2D image should be converted to a 1D signal, and then fed to MFCC algorithm to extract features as done in a voice signal.

Leukemia identification mainly involves two phases: the first phase is to extract the features from the leukemia image sample and collect a dataset, this is known as training step and the second phase is to extract features from a testing sample and match them with the samples present in the database, this is known as a testing phase. Feature extraction and conversion are common steps in both training and testing phases of leukemia recognition system.

Feature extraction is the method of keeping image discriminative information while decreasing the amount of data

present in the input image sample. This method is important to identify leukemia image from normal blood cell image by producing enough information for good leukemia recognition. There are many feature extraction techniques can be used in signal recognition system like linear prediction coefficients (LPC), linear predictive cepstral coefficients (LPCC), perceptual linear predictive analysis (PLP), and Mel-Frequency spectrum coefficients (MFCC). MFCCs is the famous one and it is used in this research. MFCCs coefficients that have been used to represent the signal distribution, moreover its features come from cepstral analysis and warped to the Mel-scale which assures low frequency components over the high frequency components.

Voice recognition steps are the same as leukemia recognition, however first the leukemia 2D image should be converted to 1D signal since MFCC works on a 1D signal. The following Fig. 1 shows the steps from 2D image to MFCC coefficients.

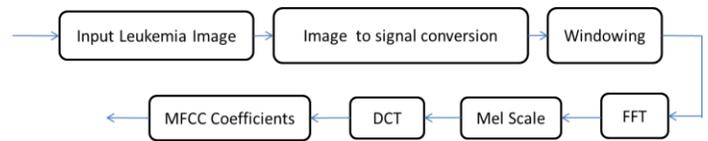


Fig. 1. MFCC feature extraction process.

#### A. Input leukemia Image

We apply MFCC technique for leukemia recognition, it is the same steps used in voice recognition, the difference is the conversion of blood cell image from the 2D image to 1D signal then applying MFCC technique, the rest of the steps are the same.

#### B. Image to signal conversion

The 1D image signal from the previous step is framed and windowed using Hamming window technique then applying Fast Fourier Transform, the resulting magnitude of the FFT spectrum is warped by a series of Mel-filter banks according to the Mel scale. The next step is taking the log of the spectrum, followed by applying a discrete cosine transform [11].

Mel is the measuring unit of the perceptual scale of perceived pitch or frequency of a tune, so the Mel-scale is a conversion between the real frequency scale in Hz and the perceived frequency scale in Mels. Mel means Melody to show that the scale is based on pitch comparisons. So, the conversion is virtually linear below 1 kHz and logarithmic above. This is the formula for converting actual frequency  $f$  hertz to the  $s$  mel-scale frequency (1):

$$s = 2595 \log_{10} \left( \frac{f}{700} + 1 \right) \quad (1)$$

#### C. Windowing

Usually a Hamming or Hanning window is used. In this procedure, every frame is multiplied by a tapering function, after windowing the signal the output is:

$$Y(s) = X(s) * W_n(s), 0 \leq s \leq N_s - 1$$

Where,  $Y(s)$  represents the output signal,  $X(s)$  is the input signal acquired from framing,  $N_s$  is the number of samples

within every frame and  $W_n(s)$  is hamming window symbolized as:

$$W_n(s) = 0.54 - 0.46 \cos(2\pi s / (N_s - 1)), 0 \leq s \leq N_s - 1 \quad (2)$$

#### D. Fast Fourier Transform (FFT)

Fourier Transformation is performed on the sliced signal. FFT is used to map a signal from time domain to frequency domain [20].  $N_s$  samples in each frame are converted to the frequency domain. FFT is a fast processing algorithm to apply and has easy computational speed. FFT transformation is made for each frame separately when the signal is divided into small frames.

#### E. Mel Scale

The previously calculated spectrums are converted on the Mel scale to know the estimate about the existing energy at each position in the spectrum. Mel scale with the triangular overlapping window is recognized as a triangular filter bank. This filter bank is an array of different band pass filters with a spacing of the preset stationary bandwidth along Mel frequency time.

Thus, the Mel scale controls the space of the given filter and calculates the width of it, when the frequency gets higher filters also get wider. The appropriate spread filters give us the energy present in the signal at each point. The conversion formula of frequency is (1).

We apply Log of base 10 to the output spectrum from Mel bank then applying DCT for standardization, this step is important for the DCT calculation to make small value large enough and large small enough.

#### F. Discrete Cosine Transform (DCT)

DCT is performed on the log Mel spectrum to convert it to the time domain. The result coefficients are named Mel frequency cepstrum coefficients (MFCCs).

$$C_n = \sum_{k=1}^k (\log S_k) \cos \left[ n \left( k - \frac{1}{2} \right) \frac{\pi}{k} \right], n = 0, 1, \dots, K - 1 \quad (3)$$

Where,  $n = 0, 1, \dots, N - 1$ ,  $K$  is the number of filters,  $C_n$  is the MFCC and  $N$  is the number coefficients here  $n = 13$  so, the total number of coefficients obtained from each frame is 13.

#### G. MFCC Coefficients

The output amplitudes of the spectrum are the MFCCs. As a summary of leukemia recognition steps: the image is mapped from 2D to a 1D image. Therefore The 1D signal image is split into small frames, then windowing is applied to suppress edges at high frequencies. The signal is transformed to the frequency domain by applying FFT technique. Furthermore, the size and space between filters are provided by Mel-scale, and then the signal is filtered by Mel-Bank. Accordingly, the output of Mel-Bank is logged then DCT is applied to get final Mel-Frequency Cepstral Coefficients.

### IV. THE PROPOSED LEUKEMIA RECOGNITION SYSTEM

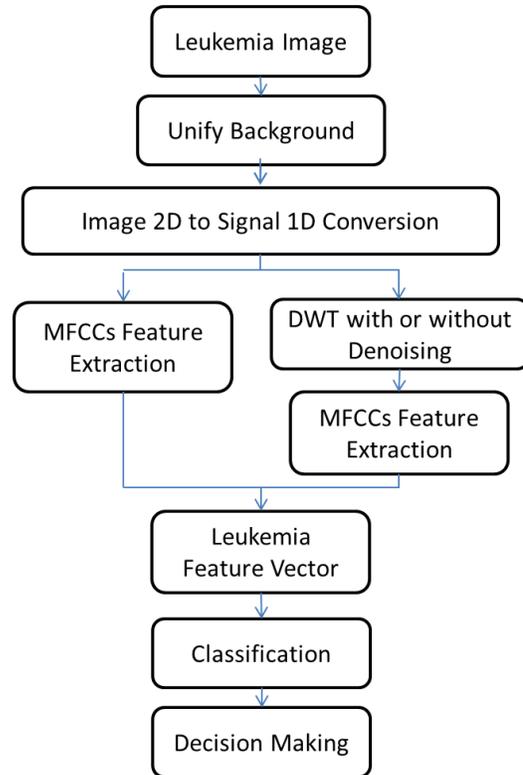


Fig. 2. Components of leukemia classification system.

The proposed leukemia recognition system consists of signal modeling, feature extraction, and feature matching (classification) as shown in Fig. 2.

For training and testing phase the image 2D is mapped to a 1D signal, MFCC features are obtained from 1D signal while discrete wavelet transformed get additional features which support the MFCC features from the original signal, furthermore, wavelet denoising is applied to signal and DWT signal to get additional features to assist MFCC features. In classification process, the unknown image features are used to predict the signal that corresponds to the class of this unknown image which is leukemia image or normal blood image.

#### A. Discrete Wavelet Transform (DWT)

It is a mathematical tool that hierarchically decomposing functions, furthermore any type of function like a curve, image, surface, or signal. Wavelet is a good technique that representing details of this function. DWT is a good method for the analysis of non-stationary signals. It represents the signal as a series of approximation at a different resolution where the low pass part corresponds to the signal while the high pass part corresponds to the details. It is the same as filtering the signal with a bank of bandpass filters whose impulse responses are all nearly given by scaled forms of a mother wavelet. The filters outputs are generally extremely reduced so the number of output samples of DWT equivalent to the number of the input samples, therefore no repetition arises in this transform.

The output features of this DWT vector are added to the features of the MFCCs generated from the original blood cell

signal to form a big feature vector that can be used for leukemia identification. These are more strong features in case of the existence of degradations.

### B. Wavelet Based Denoising

Wavelet based denoising is improving the noise robustness [21]. Wavelet denoising steps are the following: the first step is the decomposition by choosing the correct filter and applying the wavelet transform to the noisy signal to create the noisy wavelet coefficients till properly distinguish the occurrence of the perfect decomposition. The second is the vital step in wavelet denoising where we select suitable threshold border at each level and threshold technique like soft or hard thresholding to best eliminate the noises. Finally, the reconstruction step where the calculation is made for the thresholded wavelet coefficients to get the inverse wavelet transform to acquire a denoised signal.

## V. CLASSIFICATION METHODS

There are many classification methods that can be used to distinguish normal blood cell from leukemia. In this paper, we used five different classification techniques:

1) *Neighborhood Components Analysis*: This algorithm attempts to exploit a stochastic of nearest neighbors of the leave-one-out KNN score on the training set [22], its breviation is NCA.

2) *Support Vector Machines* [23]: Radial SVM are used where 5 fold validation method are used to set the values of  $C$  and  $\gamma$  within the range of (0.5-1.5). The implementation is made by the LIBSVM software.

3) *Bayes classifier* [24]: Using densities approximation of the class-condition according to the Gaussian density as a kernel function.

4) *Naive Bayes kernel classifier (NBK)*: This is mainly a naive Bayes classifier, where the one dimensional densities are approximated using a Parzen window density estimate, in place of Gaussian approximation.

5) *Discriminant Analysis (LDQ)*: Linear Discriminant Analysis can only learn linear decision boundaries, while Quadratic Discriminant Analysis can learn quadratic decision boundaries and is, therefore, more flexible.

## VI. EXPERIMENTAL RESULTS AND ANALYSIS

### A. Experimental Results

In this experiment, we need to preprocess images to unify the background of images before converting it to signal. The original colored image is converted to gray scale image (green channel), then green channel histogram is obtained. The next step is creating a binary image via thresholding to a specific value, and removing small objects from the binary image to get a clean binary image. The final step in preprocessing is masking the RGB image with a white background as shown in Fig. 3.

Total numbers of images are 210, classified as 107 normal blood cell images and 103 leukemia cell images. Normal blood cell images with its Mel-frequency cepstral coefficient are shown in Fig. 4 while leukemia cell images with its Mel-frequency cepstral coefficient are shown in Fig. 5.

In the previous Fig. 4 and 5, the X-axis is the number of frames (MFCCs) which comes from input signal while the Y axis is the feature vector values for each frame. After this step, the classification techniques are applied to calculate best accuracies.

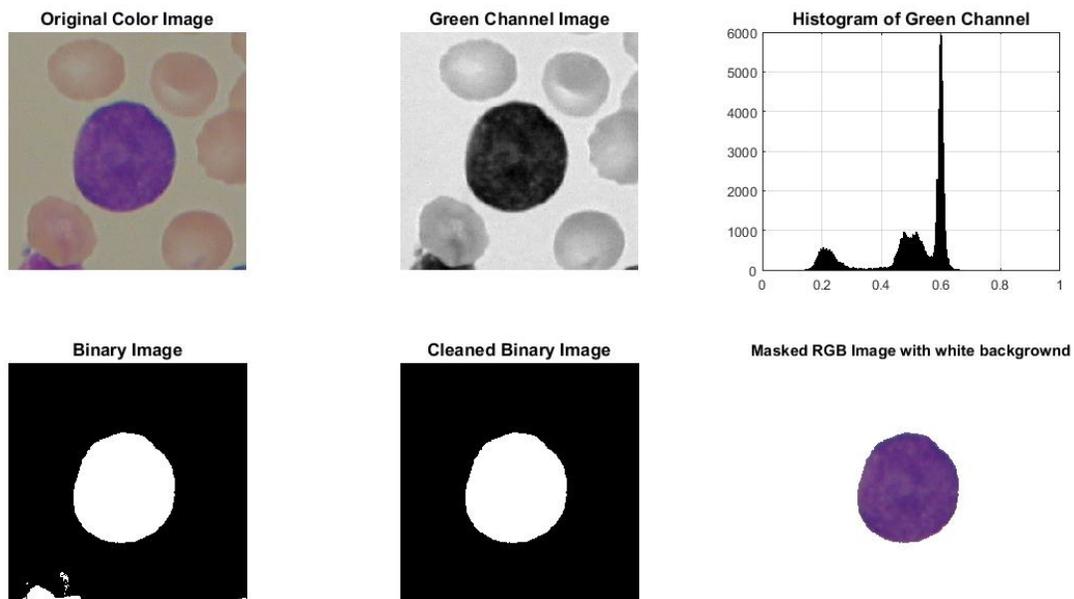


Fig. 3. Preprocessing of images.

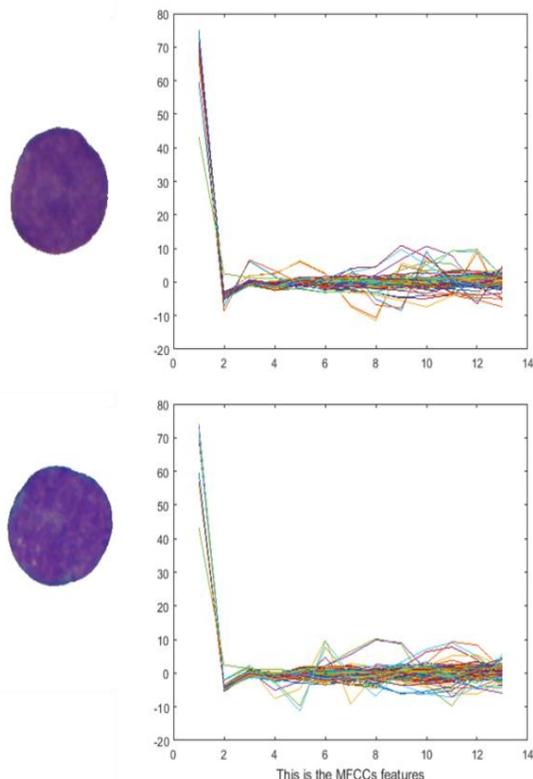


Fig. 4. MFCCs of normal blood cells.

**B. Results and Discussion**

In this paper, we used six techniques for extracting features. In the first technique, MFCCs are extracted from the blood cell signals only. In the second one, the features are extracted from the MFCCs of the DWT of the blood cell signals. In the third technique, denoising process is applied to signal and features are extracted from denoised MFCCs technique of the blood cell signals. In the fourth technique, the features are extracted from the MFCCs of the denoised DWT technique of the blood cell signals. In the fifth technique, denoising process is applied and features are extracted from both the denoised signals and DWT of the denoised signals and concatenate these features in a single feature vector. In the sixth technique, denoising process is applied to MFCCs signal only so the features are extracted from denoised MFCCs signals and DWT of the blood cell signals and concatenate these features in a single feature vector. A comparison between the six experiments is given in Fig. 6.

Fig. 6 illustrates that MFCC features extraction of the DWT of the blood cell signal (the second method) and features extracted from the MFCC of the denoised DWT of the blood cell signal (the fourth method) have the equivalent recognition rate and they both achieved the finest recognition rates in the existence of noise. This result shows the strength of the wavelet features to facilitate the recognition of the leukemia images and normal blood images with and without noise. We focus on these two methods (the second and the fourth) in next experiments and exclude other methods.

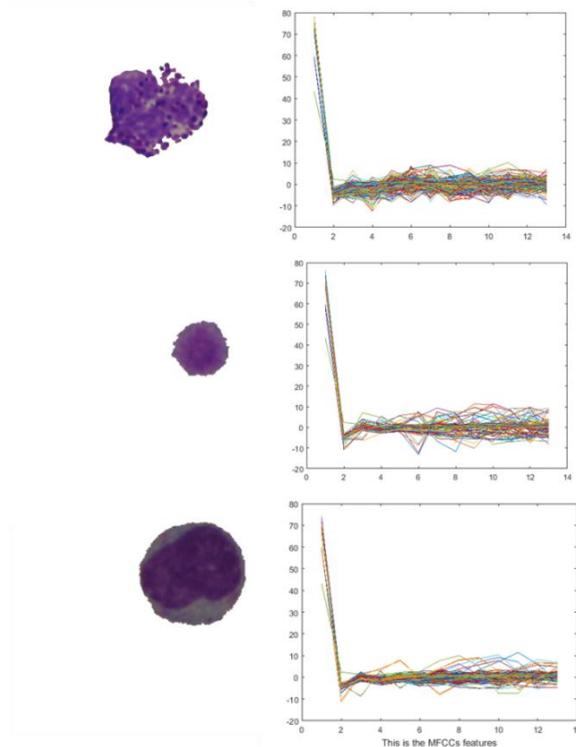


Fig. 5. MFCCs of different leukemia images.

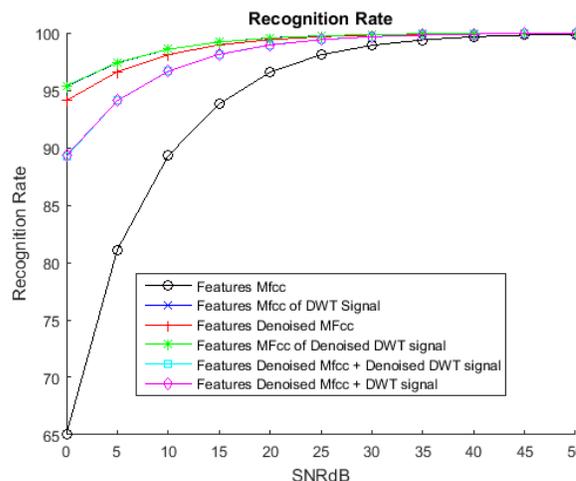


Fig. 6. Recognition rate to SNR for blood cell images.

Receiver Output Characteristic (ROC) curves are shown in Fig. 7 and 9 for the five classification techniques used for the 2<sup>nd</sup> method and 4<sup>th</sup> method, respectively. ROC curve represents the performance curve for classifier output and as we see the NBayes kernel has the highest (area under the curve) in both Fig. 7 and 9 for features from the MFCCs of the DWT signal (2<sup>nd</sup> method) and features from the MFCCs of the denoised DWT signal (4<sup>th</sup> method), respectively. The detailed values of ROC curves for the previous two methods are shown in Table 1.

Classification accuracies are shown in Fig. 8 and 10 for the five classification techniques for the 2<sup>nd</sup> method and 4<sup>th</sup> method, respectively. The NBayes kernel has the highest

identification accuracy of 92.85% of the features from the MFCCs of the DWT signal (2<sup>nd</sup> method).

The detailed values of classification accuracies for the previous two methods are shown in Table 1.

In summary, Fig. 11 shows a comparison of the classification accuracies between the five classification techniques used (NBayes kernel, NBayes, LDA Quadratic, SVM Radial, and NCA) for the 2<sup>nd</sup> method and 4<sup>th</sup> method. The highest accuracy classifier is Naïve Bayes kernel and the second highest accuracy is equally between NBayes and LDA Quadratic while the worst one is NCA.

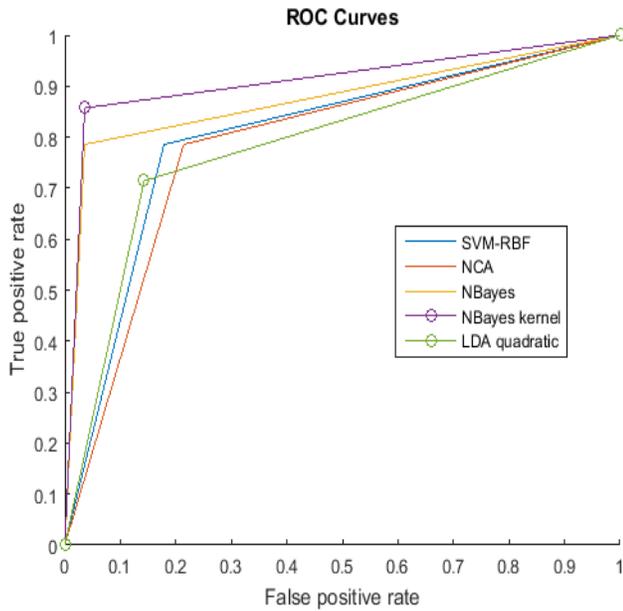


Fig. 7. Roc Curves of features from MFCC of the DWT Signal (2<sup>nd</sup> Method).

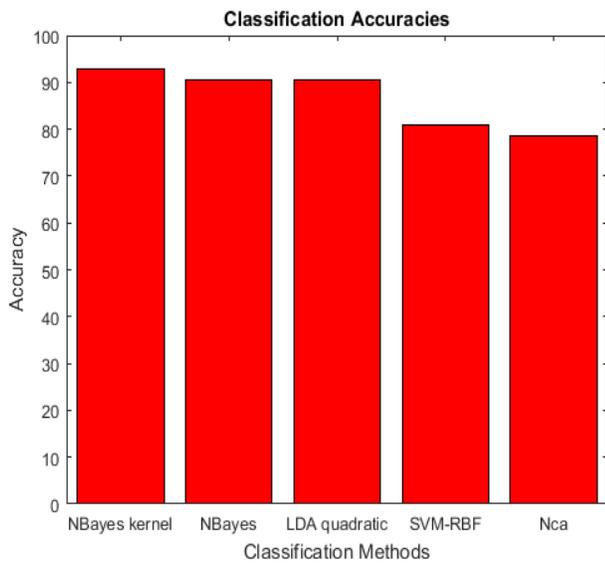


Fig. 8. Classification Accuracies of MFCC of the DWT Signal (2<sup>nd</sup> Method).

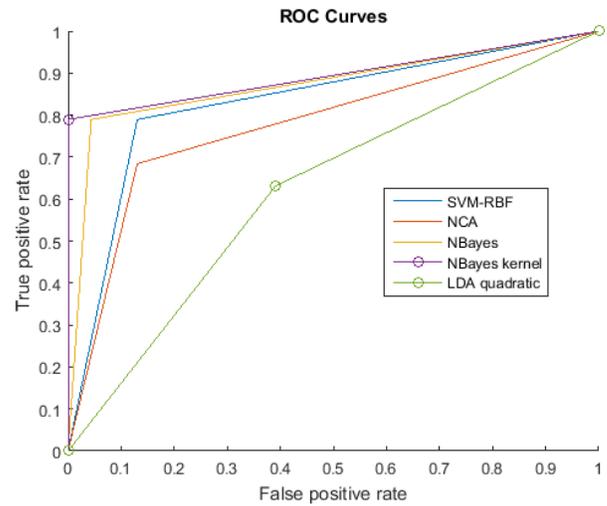


Fig. 9. Roc Curves of features from MFCC of the denoised DWT Signal (4<sup>th</sup> Method).

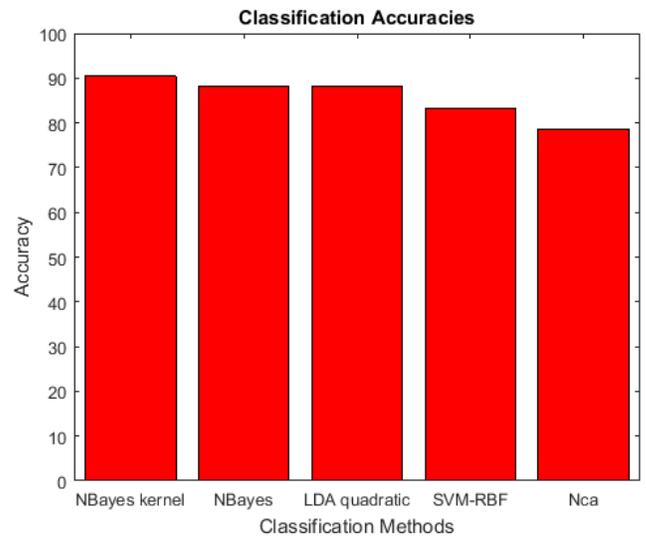


Fig. 10. Classification Accuracies of MFCC of the Denoised DWT Signal (4<sup>th</sup> Method).

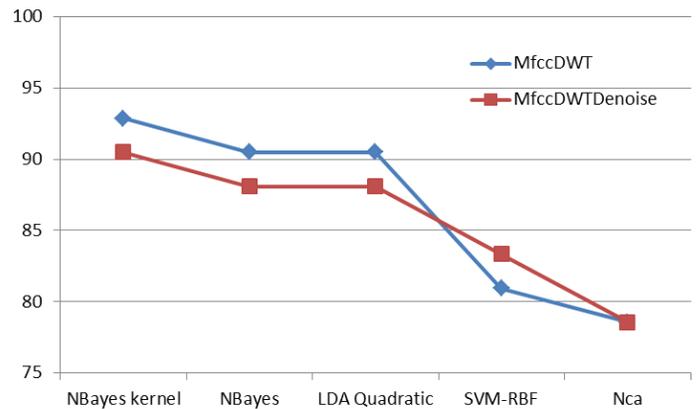


Fig. 11. Comparison of classification accuracies for the five classifiers.

TABLE I. RESULT OF CLASSIFICATION ACCURACY AND AREA UNDER THE CURVE

	Classification Accuracy		Area Under the Curve (ROC)	
	Mfcc of DWT	Mfcc of denoised DWT	Mfcc of DWT	Mfcc of denoised DWT
<b>NBayes kernel</b>	92.85714	90.47619	0.910714	0.894737
<b>NBayes</b>	90.47619	88.09524	0.875	0.872998
<b>LDA Quadratic</b>	90.47619	88.09524	0.803571	0.829519
<b>SVM Radial</b>	80.95238	83.33333	0.785714	0.620137
<b>NCA</b>	78.57143	78.57143	0.785714	0.776888

## VII. CONCLUSION AND FUTURE WORK

This paper has given a strong identification method for leukemia identification images based on MFCC and the wavelet transform techniques. Five different classification techniques are used to know the best classifier for identification.

The experiments illustrate that the best results are accomplished by extracting features from the MFCCs of the DWT signal with and without the presence of noise. The experimental results have shown that the recommended technique is beneficial for classification of blood cell images to leukemia or normal blood cell and it is a novel application for MFCC method because it is mainly used for speech or voice recognition.

The best classification technique of blood cell images to be normal or leukemia cells are Naïve Bayes kernel Classifier.

In future work, the experiment will be done on a larger dataset, and the MFCC will be combined by different classification techniques to increase the accuracy of the leukemia image recognition rate.

## REFERENCES

- [1] G. Ji Z. Yang, and W. You, "Pls-based gene selection and identification of tumor-specific genes," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. PP, no. 99, pp. 1-12, 2010.
- [2] M.R. Asadi, A. Vahedi, and H. Amindavar, "Leukemia cell recognition with zernike moments of holographic images," in Proceedings of the 7th Nordic Signal Processing Symposium (NORSIG 2006), pp. 214-217., June 2006.
- [3] Pratik M. Gumble, and Dr.S.V.Rode, "Analysis & Classification of Acute Lymphoblastic Leukemia using KNN Algorithm," International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN: 2321-8169, PP: 94 - 98, Volume 5 Issue 2, February 2017.
- [4] Shrikant D.Kale, and Dr.S.B.Kasturiwala, "Detection of abnormality in endoscopic images using endoscopic technique," IJRITCC, Vol - 4 , Issue - 1, January 2017.
- [5] Pratik M. Gumble, and Dr.S.V.Rode "Study & Analysis of Acute Lymphoblastic Leukemia Using Image Processing," IJRITCC, Vol- 5, Issue-2 January 2017
- [6] Vivek Tyagi, and Christian Wellekens, "On desensitizing the Mel-Cepstrum to spurious spectral components for Robust Speech Recognition," Proceedings in Acoustics, Speech, and Signal Processing, (ICASSP '05). IEEE International Conference, vol. 1, pp. 529-532., 2005.
- [7] Jong B. C., "Wavelet Transform Approach For Adaptive Filtering With Application To Fuzzy Neural Network Based Speech Recognition," PhD Dissertation, Wayne State University, 2001.
- [8] Sarikaya R., "Robust And Efficient Techniques For Speech Recognition in Noise," PhD Dissertation, Duke University, 2001.
- [9] Pullella D., "Speaker Identification Using Higher Order Spectra," Dissertation of Bachelor of Electrical and Electronic Engineering, University of Western Australia, 2006.
- [10] A. Shafik, S. M. Elhalafawy, S. M. Diab, B. M. Sallam, and F. E. Abd El-samie, "Wavelet Based Approach for Speaker Identification from Degraded Speech," IJCNIS, 2009.
- [11] K. Sri Rama Murty, and B. Yegnanarayana, "Combining Evidence from Residual Phase and MFCC Features for Speaker Recognition," IEEE Signal Processing Letters, VOL. 13, NO. 1, January 2006.
- [12] R. Chengalvarayan, and L. Deng, "Speech Trajectory Discrimination Using the Minimum Classification Error Learning," IEEE Transactions on Speech And Audio Processing, Vol. 6, No. 6, pp. 505-515, 1998.
- [13] Noda J.J., Travieso C.M., and Sánchez-Rodríguez D. , "Fusion of Linear and Mel Frequency Cepstral Coefficients for Automatic Classification of Reptiles.," Appl. Sci. 178, 7, 2017.
- [14] Pooja Prajapati and Miral Patel, "Feature Extraction of Isolated Gujarati Digits with Mel Frequency Cepstral Coefficients (MFCCs)," International Journal of Computer Applications 163(6):29-33, April 2017.
- [15] Biswas Sangeeta, "MFCC based Face Identification," Titech Japan, 2007.
- [16] T. M. Talal and A. El-Sayad, "Identification of Satellite Images Based on Mel Frequency Cepstral Coefficients," ICCES, 2009.
- [17] Fahmy M. M., "Palmpoint recognition based on Mel frequency Cepstral coefficients feature extraction," Ain Shams Engineering Journal, p.9, 2010.
- [18] Shibli Nisar, Mushtaq Ali Khan, and Muhammad Usman, "Iris Recognition using Mel-Frequency Cepstral Coefficient," International Journal of Engineering Research Volume No.3, Issue No.2, pp : 100-103, 2014.
- [19] Shikha Gupta, Jafreezal Jaafar, Wan Fatimah, wan Ahmad, and Arpit Bansal, "Feature Extraction Using MFCC," Signal & Image Processing : An International Journal (SIPIJ) Vol.4, No.4, August 2013.
- [20] Jorge Martinez, Hector Perez, Enrique Escamilla, and Masahisa Mabo Suzuki, "Speaker recognition using Mel Frequency Cepstral Coefficients (MFCC) and Vector Quantization (VQ) Techniques," Electrical Communications and Computers (CONIELECOMP), 2012 22nd International Conference, pp. 248 - 25, , Feb. 2012.
- [21] M. Farge, N. Kevlahan, V. Perrier, and E. Goirand, "Wavelets and Turbulence," Proceedings of the IEEE, Vol. 84, No. 4, pp. 639-669, 1996.
- [22] Goldberger J Roweis S, Hinton G, Salakhutdinov R "Neighbourhood components analysis.," Advances in Neural Information Processing Systems vol. 17, 2004.
- [23] Scholkopf B Smola A "Learning with kernels: Support vector machines, regularization, optimization, and beyond " MIT Press, Cambridge, MA, 2001.
- [24] Duda RO Hart PE, Stork DG, "Pattern Classification, 2nd Edition," Wiley InterScience, 2000.

# A New Strategy of Validities' Computation for Multimodel Approach: Experimental Validation

Abdennacer BEN MESSAOUD  
National Engineering School of  
Tunis, LR11ES20 LACS Laboratory  
Tunis El Manar University, Tunisia

Samia TALMOUDI BEN AOUN  
National Engineering School of  
Tunis, LR11ES20 LACS Laboratory  
Tunis El Manar University, Tunisia

Moufida LAHMARI KSOURI  
National Engineering School of  
Tunis, LR11ES20 LACS Laboratory  
Tunis El Manar University, Tunisia

**Abstract**—The evaluation of validities is a fundamental step in the design of the multimodel approach. Indeed, it is thanks to validities that we estimate the contribution of each base-model in the reproduction of the behavior of the global process in a given operating area. These coefficients are calculated most commonly by the approach of the residues formulated by the distance between the real output and the sub-models' outputs. In this paper, a strategy allowing to improve the performances of the residues' approach in terms of precision and robustness is proposed. This strategy is based on a quasi-hierarchical structuring. A simulation example and a validation on a semi-batch reactor showed the interest and the effectiveness of the proposed strategy.

**Keywords**—Validities; residues' approach; multimodel; quasi-hierarchical structuring; experimental validation

## I. INTRODUCTION

The multimodel approach has been of considerable interest for many years. The works of [1] and [2] define the idea of the multimodel approach as the apprehension of a nonlinear behavior of a system by a set of local models (linear or affine) characterizing the system operation in different operating zones. The motivation of this approach ensues from the fact that it is often difficult to conceive or to identify a model taking into account the complexity of the studied system. Using this definition, the multimodels can be understood as models defined around different operating points.

The multimodel approach offers an interesting alternative and a powerful tool to bypass the difficulties to identify, control and diagnose a nonlinear and complex system [3]-[8].

The multimodel modeling concept consists of simply to represent the dynamics of a nonlinear system by a family of relatively simple models properly characterizing the functioning of the system in its different operating areas. This models family constitutes the base of local models of the system.

In the literature, three different methods may be employed for the determination of the models' base [9]. The first one is only based on the measures of inputs/outputs of the system from which are estimated the different models' parameters [5],

[10], [11]. For the second and the third method, we supposed to have a nonlinear mathematical model, the base-models are obtained either by linearization around the different operating points [1], or by convex polytopic transformation [6], [12].

The global model output, so-called multimodel output, is obtained by a combination of the local models' outputs weighted by their validity indexes representing the relevance degree of each model estimated at each instant by a suitable decision process.

Several validities' calculation methods have been proposed in the literature [13]-[17]. The residues' approach is the most commonly used [3], [4], [12], [13], [18]-[20]. This approach is based on the calculation of the residues formulated by the distance between the real output and the sub-models' outputs. However, the performances of the multimodel approach, whose base-models' validities are calculated by the residues' approach, are considerably deteriorated in several cases of complex systems [13], [16], [17]. A new strategy, presented in this paper, allows to solve this problem and to improve the precision of the multimodel approach.

The paper is organized as follow: In Section 2, the validity concept is presented. The classical residues' approach is presented in Section 3. The new strategy of validities' computation by the residues' approach is detailed in Section 4. To illustrate the interest and efficiency of the new strategy, the simulation results are given in Section 5. Section 6 is reserved for validation on a semi-batch reactor.

## II. VALIDITY CONCEPT

The validity  $v_i$  represents the contribution of the local model  $M_i$  in the description of the behavior of the global system in a given operating area. It is estimated at each instant by a suitable decision process (Fig. 1). When it is equal to 0, the corresponding model is considered as inactive and consequently has no influence on the global multimodel system. If, on the contrary, this validity takes the value 1, the model represents perfectly the process at the considered instant. In the case where the relevance degree is between 0 and 1, the corresponding model  $M_i$  represents partially the system behavior [13].

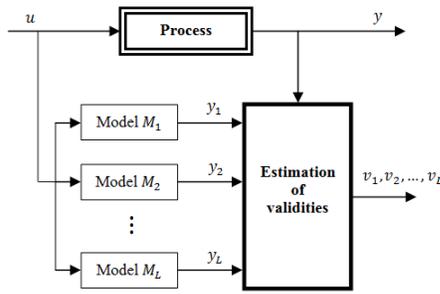


Fig. 1. Structure of the validities' estimation for the multimodel approach.

The base-models' validities satisfy the following convex sum property [21]:

$$\begin{cases} v_i \in [0,1] ; i = 1,2,\dots,L \\ \sum_{i=1}^L v_i = 1 \end{cases} \quad (1)$$

Where,  $v_i$  is the validity of the  $i^{\text{th}}$  model and  $L$  is the number of base-models.

Once the validities are estimated, the multimodel output is obtained by a combination of the local models' outputs weighted by their respective validities and given by the following formula:

$$y_{mm}(k) = \sum_{i=1}^L v_i(k) y_i(k) \quad (2)$$

Where,  $y_i$  is the output of local model  $M_i$ .

### III. RESIDUES' APPROACH

This method requires only knowledge of the base-models outputs and the global system response [14].

The validities' calculation is based on the residues which are based on the online calculation of the difference between the process output and those of the various models  $M_i$  of the base:

$$r_i = |y - y_i| ; i = 1, \dots, L \quad (3)$$

Where,  $y$  is the process output and  $y_i$  is the output of the model  $M_i$ .

The validities are deduced from the following equation:

$$v_i = 1 - \frac{r_i}{\sum_{j=1}^L r_j} ; i = 1, \dots, L \quad (4)$$

These validities are normalized by (5).

$$v_i^{simp} = \frac{v_i}{\sum_{j=1}^L v_j} ; i = 1, \dots, L \quad (5)$$

The validities' approach presented above is effective in cases where the operating areas present overlapping [10].

In some cases the validities values are so moved closer that we need to implement methods known as reinforcement methods in order to distinguish them. This reinforcement operation may be defined for example by (6).

$$v_i^{renf} = v_i \prod_{\substack{j=1 \\ j \neq i}}^L (1 - v_j) ; i = 1, \dots, L \quad (6)$$

The normalized reinforced validities, satisfying the convex sum property, are given by [3]:

$$v_i^{renf}_n = \frac{v_i^{renf}}{\sum_{j=1}^L v_j^{renf}} ; i = 1, \dots, L \quad (7)$$

## IV. THE NEW PROPOSED STRATEGY

### A. Problem Statement

In order to highlight the deterioration phenomenon of the quality of the approximation by the multimodel representation using the residues' approach to estimate the validities' indexes, the approximation problem of the following static nonlinear function is considered:

$$y(t) = 1 + \exp(-t^2) \sin(\pi t) ; t \in [-2; 2] \quad (8)$$

The system operating space can be decomposed into two operating areas. Each zone is then characterized by a sub-model (Fig. 2):

$$\begin{cases} \text{Model } M_1 : y_1(t) = 3.03t + 1 \\ \text{Model } M_2 : y_2(t) = -1.78t + 2.74 \end{cases} \quad (9)$$

The multimodel is then given by the following equation:

$$y_{mm}(t) = v_1(t) y_1(t) + v_2(t) y_2(t) \quad (10)$$

whose validities are estimated by the residues' approach.

The Fig. 3 illustrates a dynamic behavior of the nonlinear function badly approached by the multimodel in the range [0.1, 0.7]. This is due to the insufficiency of the decomposition of the operating space into two zones. Decomposing then the operating space into three zones (Fig. 4), a third model is added to the models' base:

$$\text{Model } M_3 : y_3(t) = -1.82t - 0.77 \quad (11)$$

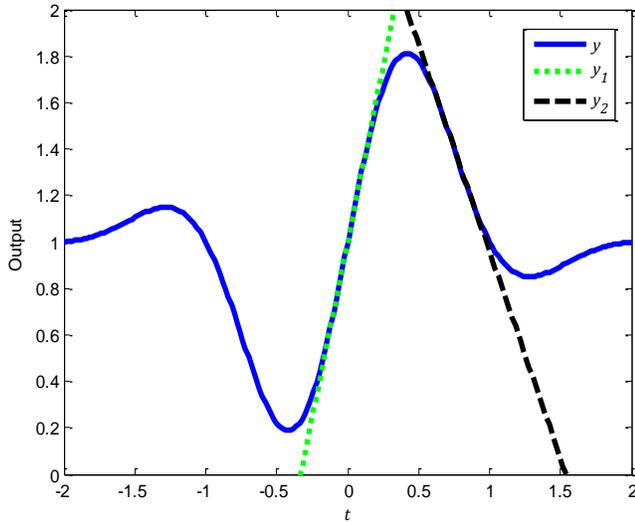


Fig. 2. Nonlinear system and local models with L=2.

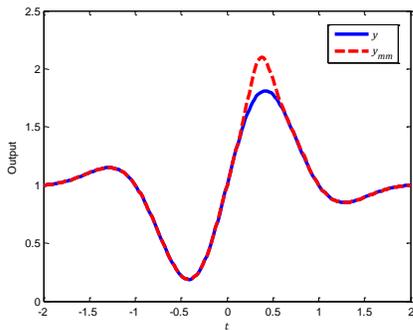


Fig. 3. Nonlinear system and multimodel approximation with L=2.

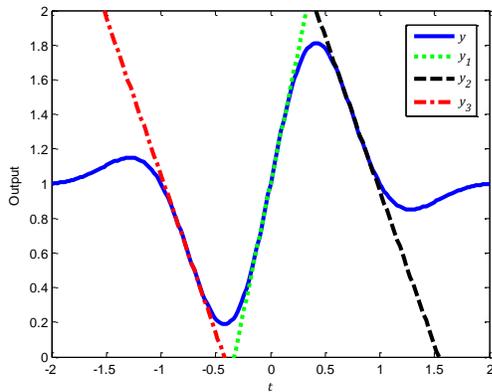


Fig. 4. Nonlinear system and local models with L=3.

Re-calculating the relevance degrees of the different sub-models by the residues' approach, the result of approximation by the multimodel approach is given in Fig. 5. This figure

shows a high deterioration of the approximation quality compared to the case when the operating space is decomposed into two areas.

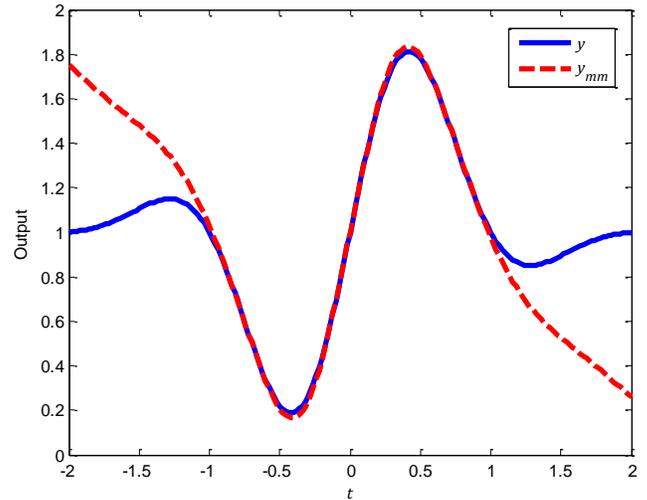


Fig. 5. Nonlinear system and multimodel approximation with L=3.

By examining the strategy on which is based the estimation of the validities values by residues' approach, it can be concluded that the weakness of this strategy is related to the normalization phase relative to the different residues. Indeed, this phase does not take into account the result obtained in the case of the decomposition of the operating space into two areas. To remedy this problem a new strategy will be proposed in the following.

### B. New Proposed Strategy

Let us assume that at instant  $k$ , the residues' calculation using (3) gives ascending values  $(r_1 \leq r_2 \leq \dots \leq r_L)$ . The new strategy is based on a quasi-hierarchical structuring as shown in Fig. 6 whose validities  $(v'_1, v'_2, \dots, v'_L)$  and  $(v_{pmm_1}, v_{pmm_2}, \dots, v_{pmm_{L-2}})$  are calculated by (5) or (7).

The validity of each base-model is given, therefore, by the following equation:

$$\begin{cases} v_1 = v'_1 \prod_{j=1}^{L-2} v_{pmm_j} \\ v_i = v'_i \prod_{j=i-1}^{L-2} v_{pmm_j} ; i = 2, \dots, (L-1) \\ v_L = v'_L \end{cases} \quad (12)$$

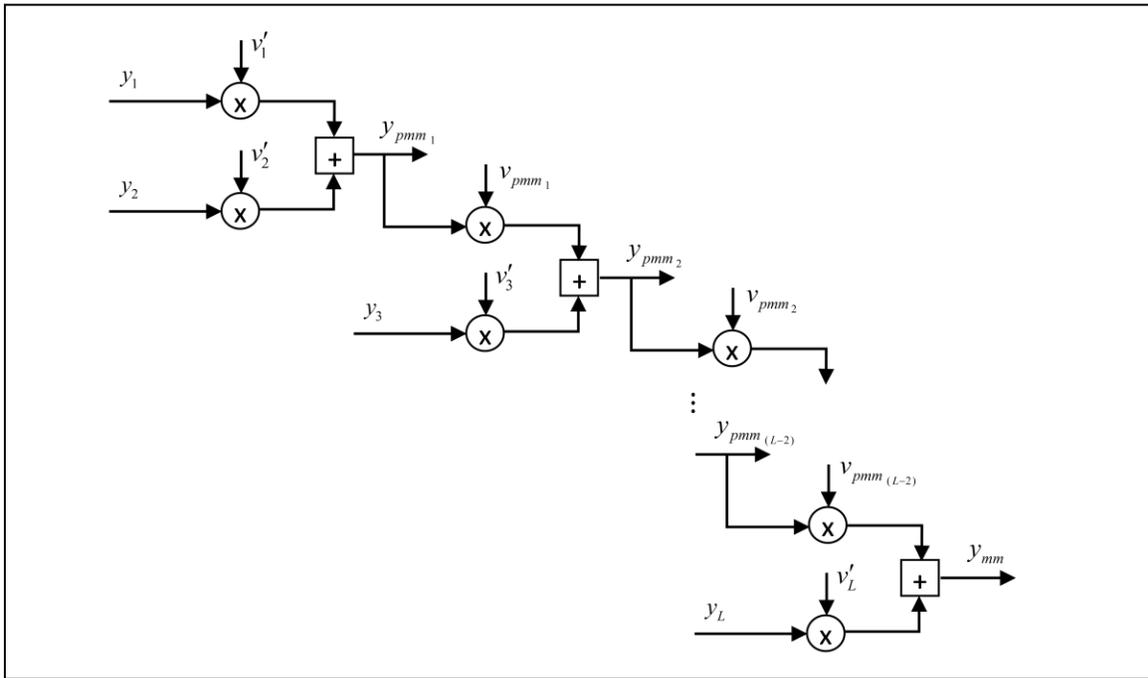


Fig. 6. New strategy: quasi-hierarchical structure.

In the general case and for each instant  $k$ , the base-models validities are calculated by Algorithm 1.

**Algorithm 1 : Validities Computation**

```

1 Begin
2 Create a table:  $Tab_y = [y_1 \ y_2 \ \dots \ y_L]$ 
3 Calculate the corresponding residues by equation (3);
4 Create a table:  $Tab_r = [r_1 \ r_2 \ \dots \ r_L]$ ;
5 Arrange  $Tab_r$  in ascending order into  $TabC_r$  and create a
table  $Tab_{ind}$  for the corresponding indices of  $Tab_r$ ;
6 for  $i = 1$  to 2
7  $r_i \leftarrow TabC_r[i]$  and Calculate  $v'_i$  by (5) (or by (7));
8 end
9  $y_{pmm_1} = v'_1 \cdot Tab_y[Tab_{ind}[1]] + v'_2 \cdot Tab_y[Tab_{ind}[2]]$ ;
10  $j \leftarrow 1$ ;
11 for  $i = 3$  to  $L$ 
12  $r'_i \leftarrow TabC_r[i]$  and  $r_{pmm_j} = |y - y_{pmm_j}|$ ;
13 Calculate  $v'_i$  and  $v_{pmm_j}$  by (5) (or by (7));
14 if  $i = L$ 
15 STOP !;
16 else
17  $y_{pmm_{j+1}} = v_{pmm_j} \cdot y_{pmm_j} + v'_i \cdot Tab_y[Tab_{ind}[i]]$ ;
18 end
19  $j \leftarrow j + 1$ ;
20 end
21 for  $i = 1$  to  $L$ 
22 for  $j = 1$  to  $L$ 
23 if  $Tab_{ind}[j] == i$ 

```

```

24 if  $j == 1$ 
25  $v_i = v'_j \prod_{k=1}^{L-2} v_{pmm_k}$  ;
26 else if  $j == L$ 
27  $v_i = v'_L$ ;
28 else
29  $v_i = v'_j \prod_{k=j-1}^{L-2} v_{pmm_k}$  ;
30 end
31 end
32 end
33 end
34 end
35 end

```

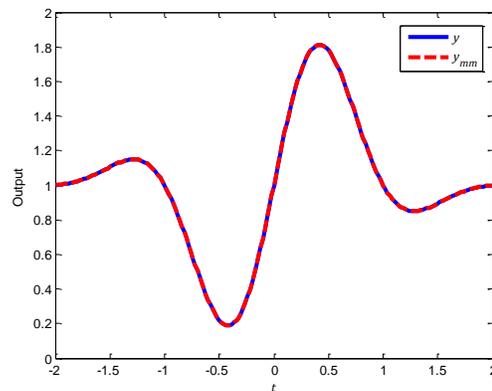


Fig. 7. Nonlinear system and multimodel approximation with  $L=3$  (new strategy).

Applying this new strategy to the static nonlinear function (8), a perfect adequacy between the nonlinear model data and those of the multimodel is shown in Fig. 7.

### V. SIMULATION EXAMPLE

In order to underline the interest of the new strategy of residues' approach for validities' computation, a simulation example was considered. The performances of the models are assessed using the *Mean Square Error (MSE)* and the *Variance-Accounted-For (VAF)* indicators given by the following equations [22]:

$$MSE = \frac{1}{N} \sum_{k=1}^N (y(k) - y_{mm}(k))^2 \quad (13)$$

$$VAF = \max \left\{ 1 - \frac{\text{var}(y(k) - y_{mm}(k))}{\text{var}(y(k))}, 0 \right\} \times 100\% \quad (14)$$

Where,  $y(k)$  and  $y_{mm}(k)$  are the system and the multimodel output, and  $\text{var}(\cdot)$  denotes the variance of a signal.

The considered example is a discrete system with time varying parameters, described by the following equation [17]:

$$y(k) = -a_1(k)y(k-1) - a_2(k)y(k-2) + b_1(k)u(k-1) + b_2(k)u(k-2) \quad (15)$$

The variation laws of different parameters of the process are given by Fig. 8.

By applying the multimodel approach, Talmoudi et al. [17] demonstrated that the models base is composed of three models whose transfer functions are given by:

$$H_1(z^{-1}) = \frac{0.18104z^{-1} + 0.07183z^{-2}}{1 - 1.1657z^{-1} + 0.2073z^{-2}} \quad (16)$$

$$H_2(z^{-1}) = \frac{0.10423z^{-1} + 0.1325z^{-2}}{1 - 1.2806z^{-1} + 0.3258z^{-2}} \quad (17)$$

$$H_3(z^{-1}) = \frac{0.018301z^{-1} + 0.20512z^{-2}}{1 - 1.3801z^{-1} + 0.42936z^{-2}} \quad (18)$$

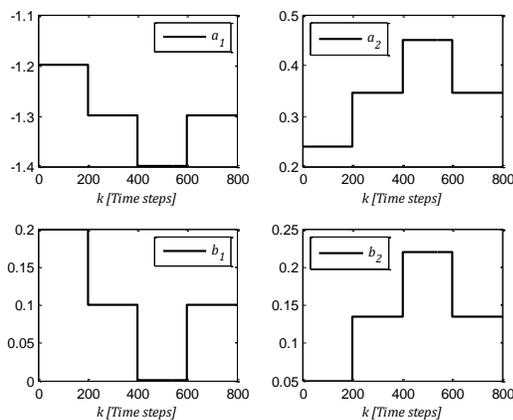


Fig. 8. Variation laws of different process parameters.

The multimodel output is obtained by the fusion of the different models outputs  $y_1(k)$ ,  $y_2(k)$  and  $y_3(k)$  weighted by their respective validity indexes  $v_1(k)$ ,  $v_2(k)$  and  $v_3(k)$ :

$$y_{mm}(k) = v_1(k)y_1(k) + v_2(k)y_2(k) + v_3(k)y_3(k) \quad (19)$$

Let us consider the following validation input sequence:

$$u(k) = 0.5 + (\exp(-0.05k)) \cos\left(\frac{k\pi}{30}\right) \quad (20)$$

The validities are calculated at first by the classical formulation of the residues' approach (reinforced validities), and secondly by the new strategy.

The simulation results are given in Fig. 9 and 10 where the relative error between the real and the multimodel outputs is calculated using the following equation:

$$e_{relative}(k) = \left| \frac{y(k) - y_{mm}(k)}{y(k)} \right| \times 100\% \quad (21)$$

These figures show that the new strategy of residues' approach for validities' computation clearly improves the precision of the multimodel output compared with the results obtained by the residues' approach in its classical formulation. This is also proved by the performance indicators (Table 1).

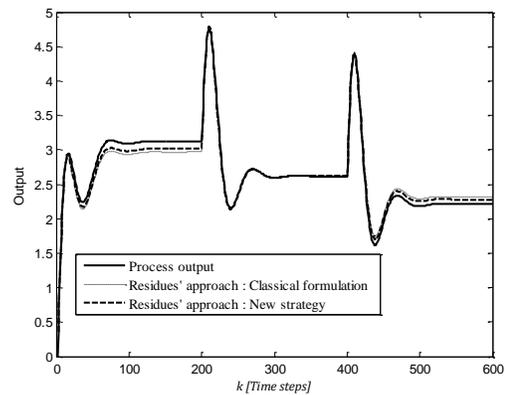


Fig. 9. Real and multimodel outputs.

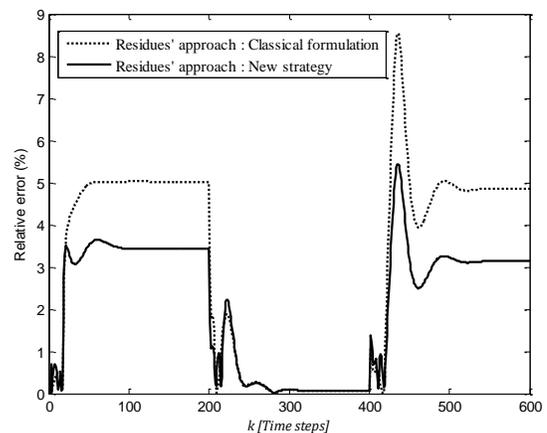


Fig. 10. Evolutions of the relative errors.

TABLE I. MSE AND VAF

	Classical formulation	New strategy
MSE	0.0104	0.0049
VAF	97.07%	98.65%

### A. Robustness Study

In order to evaluate the quality of the validities calculation method based on the new strategy while comparing it to the residues' approach in its classical formulation, a robustness study of these two methods was made. Three cases are examined:

#### 1) Robustness relative to output measurement error:

In this case, the variance of the measurement error was adjusted so that a noise-to-signal ratio at output (SNR) was 0.1%, 1%, 5% and 10% in power:

$$SNR(\%) = \frac{\text{var}(w(k))}{\text{var}(y_0(k))} \times 100 \quad (22)$$

Where,  $y_0(k)$  represents the part of the noise-free output signal and  $w(k)$  is the measurement error.

#### 2) Robustness relative to the base-models parameters:

The deviation of the base-models parameters is defined as follows:

$$\begin{cases} a_i = a_{i_0} + \Delta_{a_i} \\ b_i = b_{i_0} + \Delta_{b_i} \end{cases} \quad (23)$$

with:

$$\begin{cases} \Delta_{a_i} = \pm \frac{\delta a_i}{100} |a_{i_0}| \\ \Delta_{b_i} = \pm \frac{\delta b_i}{100} |b_{i_0}| \end{cases} \quad (24)$$

The simulation was made for the following variations:

$$\begin{cases} \delta a_i = 0.01; 0.1; 0.2 \text{ and } 0.5\% \\ \delta b_i = 0.01; 0.1; 0.2 \text{ and } 0.5\% \end{cases}$$

#### 3) Robustness relative to the base-models poles:

The deviation of the base-models poles is defined as follows:

$$p_i = p_{i_0} + \Delta_{p_i} \quad (25)$$

with:

$$\Delta_{p_i} = \pm \frac{\delta p_i}{100} |p_{i_0}| \quad (26)$$

The simulation was made for the following variations:

$$\delta p_i = 0.01; 0.1; 0.2 \text{ and } 0.5\%$$

By examining Tables 2 to 4, it can be noted that despite the influence of the validities calculation method based on the new strategy by the level of the measurement error as well as the deviations of the parameters and poles of the base-models, it generally gives better results than the classical residues method.

TABLE II. ROBUSTNESS RELATIVE TO OUTPUT MEASUREMENT ERROR

	SNR (%)	Classical formulation	New strategy
MSE	0.1	0.0109	0.0054
	1	0.0170	0.0105
	5	0.0320	0.0244
	10	0.0441	0.0372
VAF	0.1	96.93%	98.49%
	1	95.25%	97.07%
	5	91.32%	93.38%
	10	87.69%	89.62%

TABLE III. ROBUSTNESS RELATIVE TO THE BASE-MODELS PARAMETERS

	Deviation of parameters (%)	Classical formulation	New strategy
MSE	0.01	0.0105	0.0048
	0.1	0.0328	0.0239
	0.2	0.0511	0.0406
	0.5	0.0227	0.0183
VAF	0.01	96.99%	98.61%
	0.1	90.62%	93.26%
	0.2	90.30%	92.56%
	0.5	96.65%	94.83%

TABLE IV. ROBUSTNESS RELATIVE TO THE BASE-MODELS POLES

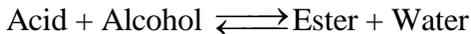
	Deviation of the poles (%)	Classical formulation	New strategy
MSE	0.01	0.0099	0.0046
	0.1	0.0030	0.0014
	0.2	0.0285	0.0212
	0.5	0.0660	0.0600
VAF	0.01	97.18%	98.72%
	0.1	99.18%	99.63%
	0.2	94.52%	96.29%
	0.5	81.14%	82.72%

## VI. VALIDATION ON A CHEMICAL REACTOR

The performances obtained by the new proposed validities calculation strategy incited us to apply it on a real model of chemical reactor [5], [11].

Fig. 11 shows the experimental device of the process.

The used reactor is a semi-batch reactor for the chemical esterification of the olive oil according to the following reaction:



The esterification reaction is carried out in a stirred tank surrounded by a jacket where a heat-transfer fluid which assures a thermal contribution to the reactor flows at a constant rate.

The heat-transfer fluid passes through a plate heat exchanger (E2) where it will be cooled, then through a resistor exchanger (E1) where it will be heated before arriving at the jacket.

Temperature sensors are used to measure the temperatures of the reactor ( $T_r$ ) and those of the heat transfer fluid at the inlet ( $T_{ede}$ ) and outlet ( $T_{sde}$ ) of the jacket.

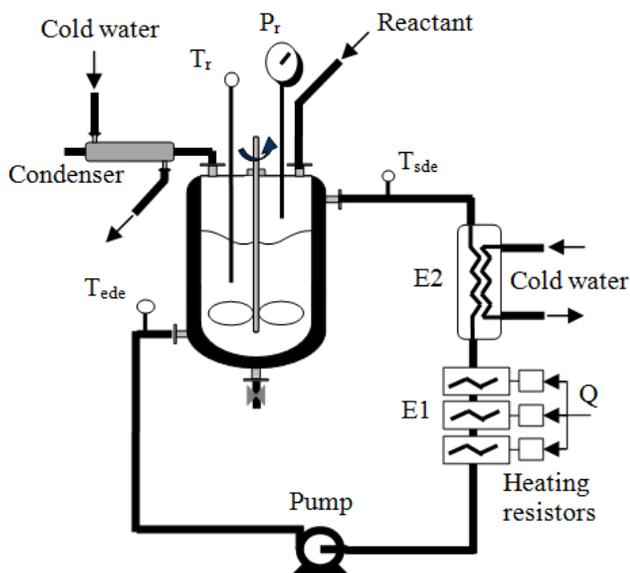


Fig. 11. Experimental device.

Three operational phases are distinguished in the production of ester:

- *Heating phase:* The reactive initially at an ambient temperature is heated to a temperature which corresponds to the reaction temperature.
- *Reaction phase:* During which the temperature of the reactional environment is maintained constant.
- *Cooling phase:* In order to retrieve the ester, the reactor is cooled back to the ambient temperature.

The process is considered as a mono-variable system where the control variable is the electric power  $Q$  supplied by the heating resistors while the output is the reactor temperature  $T_r$  [5].

Such a system is nonlinear and the use of multimodel approach is recommended [11].

Fig. 12 represents the set of identification data picked out of the reactor. The selected excitation signal  $Q$  is a Pseudo Random Binary Sequence (PRBS) applied to the reactor with a sampling time equal to 180 s.

The models' base is determined by applying the method based on the Kohonen networks [11]. This approach requires firstly determining the number of clusters. The next step consists in classifying the identification data set. And finally a step of structural and parametric estimation of the base-models is necessary.

Three second order systems are obtained:

$$H_1(z^{-1}) = \frac{5.95 \times 10^{-5} z^{-1} + 0.00185 z^{-2}}{1 - 1.238 z^{-1} + 0.2646 z^{-2}} \quad (27)$$

$$H_2(z^{-1}) = \frac{-3.255 \times 10^{-5} z^{-1} + 0.00113 z^{-2}}{1 - 0.7956 z^{-1} - 0.1906 z^{-2}} \quad (28)$$

$$H_3(z^{-1}) = \frac{9.959 \times 10^{-5} z^{-1} + 0.0018 z^{-2}}{1 - 1.135 z^{-1} + 0.1677 z^{-2}} \quad (29)$$

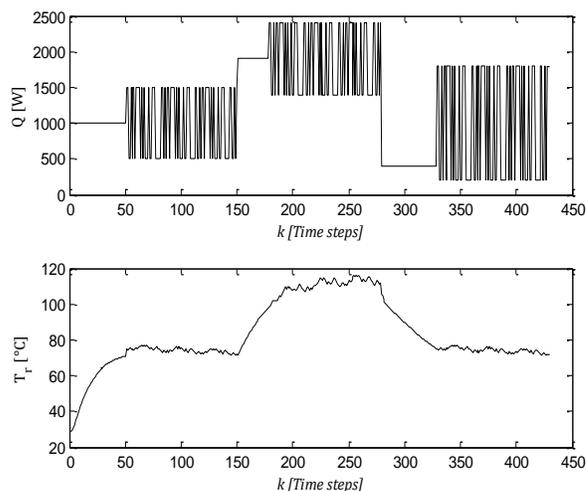


Fig. 12. Identification data set.

The result of the validation phase is given by Fig. 13 where the multimodel output is calculated by the fusion of the three base-models outputs weighted by their respective validity indexes determined at first by the residues' approach in its classical formulation (simple validities), and secondly by the new proposed strategy.

It can be seen that the new strategy of validities' computation by residues' approach offers a very satisfactory precision as compared to the residues' approach in its classical formulation. Indeed, the multimodel output, obtained by exploiting this new strategy, follows with a high precision the real output and describes perfectly the system behavior. However, the output obtained by the exploitation of the classical residues' approach follows the real output with a relatively important error. This is also proved on Fig. 14. On this figure, we drew the evolutions of the relative errors between the system output and the multimodel outputs

exploiting the residues' approach in its classical formulation and the new proposed strategy. It's clearly observed that the relative error is equal to zero by applying the new proposed strategy.

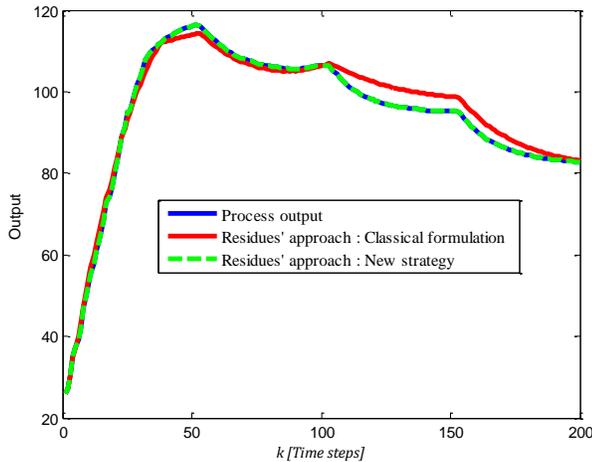


Fig. 13. Real and multimodel outputs of chemical reactor.

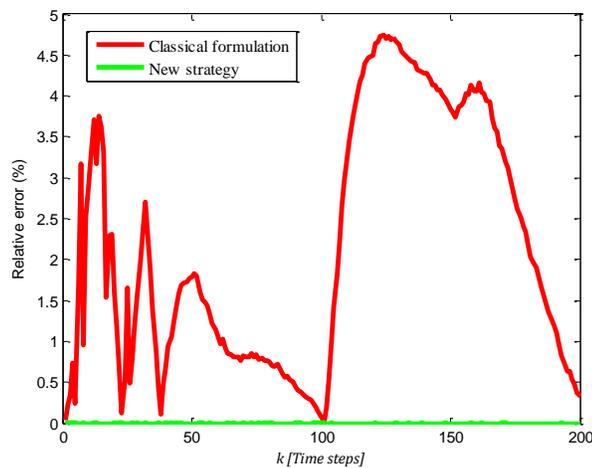


Fig. 14. Evolutions of the relative errors of chemical reactor.

TABLE V. MSE AND VAF OF CHEMICAL REACTOR

	Classical formulation	New strategy
<b>MSE</b>	6.1632	$8.116 \times 10^{-29}$
<b>VAF</b>	98.48%	100%

The performance indicator *MSE* (Mean Square Error) was calculated to evaluate the new proposed strategy of validities' computation by residues' approach compared to the classical formulation of the same approach. It is null in the case of applying the new proposed strategy (Table 5). The *variance-accounted-for VAF* asserts this result by its value that is equal to 100% (Table 5).

## VII. CONCLUSION

This paper treats one of the principal issues of the multimodel approach which is the base-models validities

estimation. In this study, a new strategy, allowing to improve the performances of the residues' approach for validities' computation, is proposed. This strategy is based on a quasi-hierarchical structuring. The different steps of validities computation were detailed. The numerical simulation results, described in this paper, prove the efficiency of the new proposed strategy as well as its impact on the improvement of the performances of the residues' approach in terms of precision and robustness. The use of the new strategy on a model of semi-batch chemical reactor showed that, in this case, the multimodel approach leads to a perfect modeling of the real process.

Operating modes management and resolving conflicting connections between them are meaningful and challenging issues, which will be studied in the future work.

## REFERENCES

- [1] T. A. Johansen and B. A. Foss, "Nonlinear Local Model Rrepresentation for Adaptative System", IEEE International Conference on intelligent control and instrumentation, Singapore, vol. 2, 1992, pp.677-682.
- [2] R. Murray-smith and T. A. Johansen, Multiple Model Approaches to Modelling and control, Taylor & Francis, London, 1997.
- [3] N. Elfelly, J. Y. Dieulot, M. Benrejeb, and P. Borne, "Multimodel Control Design using Unsupervised Classifiers", Studies in Informatics and Control, vol. 21, no. 1, 2012, pp. 101-108.
- [4] A. Karoui and M. Ksouri, "Diagnosis by the Multimodel Approach in the Frequency Domain", 4th International Conference on Logistics (LOGISTIQUA) Proceedings, Hammamet, Tunisia, 2011, pp. 261-267.
- [5] M. Ltaief, A. Messaoud, and R. Ben Abdennour, "Optimal Systematic Determination of Models' Base for Multimodel Representation: Real Time Application", International Journal of Automation and Computing, vol. 11, no. 6, 2014, pp. 644-652.
- [6] A. M. Nagy, G. Mourot, B. Marx, J. Ragot, and G. Schutz, "Systematic Multimodeling Methodology Applied to an Activated Sludge Reactor Model", American Chemical Society: Industrial & Engineering Chemistry Research ,vol. 49, no. 6, 2010, pp. 2790-2799.
- [7] C. Othman, I. Ben Cheikh, and D. Soudani, "On the Internal Multi-Model Control of Uncertain Discrete-Time Systems", International Journal of Advanced Computer Science and Applications (ijacsa), vol. 7, no. 9, 2016. <http://dx.doi.org/10.14569/IJACSA.2016.070912>.
- [8] H. Bouzaouache, "Output Feedback Controller Synthesis for Discrete-Time Nonlinear Systems", International Journal of Advanced Computer Science and Applications (ijacsa), vol. 8, no. 4, 2017. <http://dx.doi.org/10.14569/IJACSA.2017.080441>
- [9] M. Chadli and P. Borne, Multiple Models Approach in Automation: Takagi-Sugeno Fuzzy Systems, Wiley-ISTE, 2013.
- [10] N. Elfelly, J. Y. Dieulot, and P. Borne, "A Neural Approach of Multimodel Representation of Complex Processes", International Journal of Computers, Communications and Control, vol. 3, no. 2, 2008, pp. 149-160.
- [11] S. Talmoudi, K. Abderrahim, R. Ben Abdennour, and M. Ksouri, "Multimodel Approach using Neural Networks for Complex Systems Modeling and Identification", Nonlinear Dynamics and Systems Theory, vol. 8, no. 3, 2008, pp. 299-316.
- [12] A. Salem, A. S. Tlili, and N. B. Braiek, "On the Plotytopic and Multimodel State Observers of Inductor Motor", Journal of Automation & Systems Engineering, vol. 2, no. 4, 2008.
- [13] R. Ben mohamed, H. Ben Nasr, and F. M'sahli, "A Multimodel Approach for a Nonlinear System based on Neural Network Validity", International Journal of Intelligent Computing and Cybernetics, vol. 4, no. 3, 2011, pp. 331-352.
- [14] F. Delmotte, Analyse Multi-modèle, PhD Thesis, Université des sciences et techniques de Lille, France, 1997.
- [15] M. Ksouri, Contributions à la Commande Multimodèle des Processus Complexes, PhD Thesis, Université des sciences et techniques de Lille, France, 1999.

- [16] M. Ltaief, K. Abderrahim, R. Ben Abdennour, and M. Ksouri, "A Fuzzy Fusion Strategy for the Multimodel Approach Application", *WSEAS Transactions on Circuits and Systems*, vol. 2, no. 4, 2003, pp.686–691.
- [17] S. Talmoudi, K. Abderrahim, R. Ben Abdennour, and M. Ksouri, "A New Technique of Validities' Computation for Models' Base", *WSEAS Transactions on Circuits and Systems*, vol. 2, no. 4, 2003, pp. 680-685.
- [18] S. Bedoui, M. Ltaief, and K. Abderrahim, "New Method for the Systematic Determination of the Model's Base of Time Varying Delay System", *International Journal of Computer Applications*, vol. 46, no. 1, 2012, pp. 13-19.
- [19] I. Chihi, A. Abdelkrim, and M. Benrejeb, "Multi-model Approach to Characterize Human Handwriting Motion", *Biol Cybern*, vol. 110, no. 1, 2016, pp. 17-30.
- [20] A. Sassi and A. Abdelkrim, "New Stability Conditions for Nonlinear Systems described by Multiple Model Approach", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, no. 1, 2016, pp. 177-187.
- [21] R. Orjuela, B. Marx, J. Ragot, and D. Maquin, "State Estimation of Nonlinear Systems based on Heterogeneous Multiple Models: some Recent Theoretical Results", 7th Workshop on Advanced Control and Diagnosis, Zielona Gora, Poland, 2009.
- [22] R. Orjuela, B. Marx, J. Ragot, and D. Maquin, "Nonlinear System Identification using Heterogeneous Multiple Models", *International Journal of Applied Mathematics and Computer Science*, vol. 23, no. 1, 2013, pp. 103–115.

# Simulation and Analysis of Quality of Service (QoS) Parameters of Voice over IP (VoIP) Traffic through Heterogeneous Networks

Mahdi H. Miraz

School of Computer Studies  
AMA International University BAHRAIN  
Salmabad, Bahrain  
Centre for Ultra-realistic Imaging (CURI)  
Glyndŵr University, UK

Muzafar A. Ganie

Department of Computer Science &  
Software Engineering University of Hail,  
Hail, KSA  
KSA

Suhail A. Molvi

Department of Computer Science &  
Software Engineering, University of Hail,  
Hail, KSA

Maaruf Ali

Department of Science and Technology,  
University of Suffolk,  
Ipswich, Suffolk, UK

AbdelRahman H. Hussein

Department of Software Engineering,  
Al-Ahliyya Amman University,  
Amman, Jordan

**Abstract**—Identifying those causes and parameters that affect the Quality of Service (QoS) of Voice-over-Internet Protocol (VoIP) through heterogeneous networks such as WiFi, WiMAX and between them are carried out using the OPNET simulation tool. Optimization of the network for both intra- and inter-system traffic to mitigate the deterioration of the QoS are discussed. The average value of the jitter of the VoIP traffic traversing through the WiFi-WiMAX network was observed to be higher than that of utilizing WiFi alone at some points in time. It is routinely surmised to be less than that of transiting across the WiFi network only and obviously higher than passing through the increased bandwidth network of WiMAX. Moreover, both the values of the packet end-to-end delay and the Mean Opinion Score (MOS) were considerably higher than expected. The consequences of this optimization, leading to a solution, which can ameliorate the QoS over these networks are analyzed and offered as the conclusion of this ongoing research.

**Keywords**—Voice over Internet Protocol (VoIP); Quality of Service (QoS); Mean Opinion Score (MOS); simulation

## I. INTRODUCTION

Because of the ever increasing and global adoration of using the Internet, especially for Voice-over-IP (VoIP) calls on mobile devices, it is turning out to be progressively inexpedient to disregard the gravity of voice communications utilizing the Internet in our everyday lives. Due to the continuance of dissimilar types of protocols and networks (i.e. WiFi, WiMAX, 3G, 4G, LTE, CDMA, GSM, EDGE, GPRS etc.), in most cases the data has to traverse multiple assorted networks - there is an urgent need for this research. While VoIP traffic passes

through any such heterogeneous networks, the Quality of Services (QoS) suffers noticeable degradation. The solitary raison d'être of the research, presented in this paper, is to explore and investigate the level and magnitude of such degradation of the QoS of VoIP traffic traveling through these assorted networks. In pursuance of this aim, our objectives are of threefold: 1) to design, develop and configure appropriate sample networks using the OPNET modeler; 2) to run the simulation using various loads as well as to record the measured results of the QoS parameters; and finally 3) to articulate the research findings by analyzing the results procured through the simulations. The first two scenarios are made up of a number of VOIP clients transferring data through a couple of homogeneous networks i.e. WiMAX-to-WiMAX, WiFi-to-WiFi. The major QoS parameters of VoIP traffic such as the: Mean Opinion Score (MOS), Throughput, Availability, Crosstalk, Jitter, Distortion, Link Utilization Distribution, Attenuation, Loss and Echo, etc. are to be scanned and analyzed. The third set-up comprises of heterogeneous networks replacing the homogeneous ones. The VoIP traffic traverse a heterogeneous network made up of assorted protocols i.e. WiMAX-to-WiFi. The simulation will capture the same VoIP QoS parameters as in the first couple of scenarios. The results, thus obtained using the heterogeneous networks, will then be analyzed and compared with the previously attained results using the homogeneous set-ups.

The layout of this research comprises the arrangement of the following parts: The first section imparts a concise preamble to the research while the second section gives a detailed account of the background information as well as the

relevant technological/scientific terms referred to in this report. The third section comprises a "Literature Review" survey studying a broad selection of research projects and articles whereas the fourth section covers the research methodology together with the simulation scenarios of the networks as well as the necessary configuration/set-up to accomplish them. The fifth section analyzes and compares the results, followed by the concluding discussion together with the layout for potential future research directions and works.

## II. BACKGROUND TERMINOLOGY

### A. Voice over IP (VoIP)

Voice over Internet Protocol or more commonly known as "VoIP" [1] is simply defined as the digitized voice traffic intrinsically transmitted using a data network to make telephone calls. This differs from using a traditional analogue circuit switched public network, as now the data has been split into packets. These packets can take any route to reach the destination. Packetized data travel through a virtual circuit which differs from a circuit switched network in that the circuit does not need to be reserved for the entire duration of the call between the sender and the receiver with packet switching. Thus the channels may be utilized more by sharing with other users than compared to circuit switching. However, the data packet can arrive out of sequence, experience delay or even may never arrive as a consequence of traffic congestion and buffer overflows. These are some of the major disadvantages of sharing traffic across a virtual network that VoIP traffic has to contend with. On the other hand, the advantages offered include the multiple routing of the VoIP traffic ensuring a cheaper and often free of cost flow of traffic between the different intra-packet network components such as the routers and switches. Transmitting digital data in the format of packets signifies that all types of digitized data such as voice, video, fax, music and telephony have the opportunity to be carried together utilizing a shared common network at any given time.

The fact of being software packet based puts VoIP technology in a favourable or superior position. Thus, VoIP enjoys a distinct advantage and supremacy of budget scalability in comparison with the currently operational alternative telephony systems. This allows lines to be shared with other users and services thus helping to lower the overall costs over the circuit switched networks. However, being predominantly a network based on software - it is exposed to the possibility of being attacked or harmed by the progressively rising threat of cyber-attacks from crackers in terms of malware such as viruses and worms. In [2], the author discusses several security solutions to confront this potential problem.

Convergence has been accelerated with the deployment of 3G [3], WiMAX and considerably further recently by the deployment of LTE and 4G, particularly amongst internet, mobile and fixed services. Universal access to the internet regardless of the means of transportation is accelerating predominantly due to the widespread rollout of WiMAX, WiFi and femtocells in public spaces. The demand for greater

bandwidth to support multimedia broadband access is also increasing and being expected by the consumers. This was facilitated by the adoption of the IP Multimedia Subsystem (IMS) in the Rel. 5 version of UMTS (Universal Mobile Telecommunications System). The IMS is a packet based control overlay network used for transporting user data and signaling.

The Session Initiation Protocol (SIP), a development of the Internet Engineering Task Force (IETF) was embraced by the Third Generation Partnership Project (3GPP) for setting up IP-based multimedia sessions, this includes VoIP. The current IEEE 802.11 (WiFi) and 802.16 (WiMAX) networks completely support VoIP and many other real-time services [4].

### B. Session Initiation Protocol (SIP)

Making, maintaining and clearing a call requires control information and signaling to be exchanged between the network entities. This is actually a rather complicated process where internet mobility is involved across various types of devices with differing capabilities and network technologies. A protocol that has been chosen to manage these "sessions" is known appropriately as the "Session Initiation Protocol" or SIP [5]. SIP works alongside and in complement with the existing real-time protocols. The source and destination endpoints, known as the "user agents", discover each other and then negotiate the parameters for the efficient exchange of information by the use of SIP. The necessary user agents and intermediary nodes are handled by SIP by the creation of proxy servers. These proxy servers can then request and respond to 'invitation', 'registration' and other such SIP requests. SIP is a transport protocol independent of the type of session being established. SIP is designed to be agile, flexible and to handle various types of multimedia data exchange.

SIP being an application layer control protocol can take care of the entire multimedia call set-up to the termination process. It also includes the ability to handle multicast call set-up, including the removal of the participant. SIP is designed for mobility with features such as redirection and name mapping. A powerful feature of SIP is the ability to maintain an externally visible identifier, invariant of location [6]. For example, SIP supports these call set-up features: session set-up, session management, user availability, user location and user capabilities.

### C. QoS Parameters of VoIP Traffic

The data networks being flexible in its ability to handle multifarious types of data services over the Public Switched Telephone Network (PSTN) puts the Plain Old Telephone Service (POTS) at a financial disadvantage [4]. The QoS parameter of VoIP traffic varies, and can be quantified by a range of divergent metrics, such as the: jitter, end-to-end delay and Mean Opinion Score (MOS), as shown in Table 1.

The Mean Opinion Score (MOS) has been used to subjectively measure the voice quality in a telephone network. It is based on a perceptual scale of 1 to 5 as shown in Table 1.

TABLE. I. SCALING AND CLASSIFICATION OF MOS [7].

Score	Quality	Scale of Listening Effort
5	Excellent	No effort is required.
4	Good	No considerable effort is required.
3	Fair	Moderate effort is required.
2	Poor	Considerable effort is required.
1	Bad	Not understood even with considerable effort.

Jitter “is the variation in [the] arrival time of consecutive packets” [10]. Jitter is calculated over an interval of time [7]. It should be noted that the buffers can both under-fill and over-fill, triggering packet drops.

The packet end-to-end (E2E) delay “is measured by calculating the delay from the speaker to the receiver [including the] compression and decompression delays” [8].

The International Telecommunication Union – Telecommunication (ITU-T) gives the guidelines for the delay and jitter for the different types of call quality, as presented in Table 2 [8].

TABLE. II. ITU-T PRECEPT FOR VOICE QUALITY [8].

Network Parameter	Good	Acceptable	Poor
Delay (ms)	0-150	150-300	> 300
Jitter (ms)	0-20	20-50	> 50

#### D. WiFi™ (IEEE 802.11x)

The contention wireless networking technology, WiFi, evolved from its counterpart wired IEEE Ethernet 802.3, outlining perceptions for the technology of Local Area Network (LAN), to become the IEEE 802.11 Wireless LAN or WLAN. The physical and data link layers are defined, operating over the two different frequency bands of 2.4 GHz and 5 GHz. Two popular WiFi standards are the 802.11b (11 Mbit/s) and the 802.11g (54 Mbit/s) with an operating range of 80-100 m. The protocols being a contention based system, the speeds quoted are a theoretical maximum. The contention causes the comparatively low bitrates and thus affects the QoS, especially for real-time services like VoIP. This is not helped by the large headers of the WiFi and VoIP protocols themselves. Its uptake and popularity has been due to the inexpensive price of the router and most network equipment coming with its built-in, including the WiFi antenna. WiFi has now become widespread covering: domestic, industrial, public spaces including on public transportation [9].

#### E. WiMAX™ (Worldwide Interoperability for Microwave Access) Technology

WiMAX, when it was first introduced ten years ago was meant to provide a global wireless high speed mobile Internet access. However, LTE (Long Term Evolution) has largely superseded this application. WiMAX, however, is not dead and there are around 580 operators in the world providing backhaul and rural access to fast broadband internet access, often in the less developed regions of the world. Typical application scenarios of WiMAX are shown in Fig. 1. WiMAX was designed to provide the same experience as that of fixed internet services, such as QoS, Service Level Agreement (SLA), interoperability with off course mobility, wide coverage

and security [10]. It is ironic that WiMAX, once touted as the “4G of Wireless Technology” has now been superseded ahead of its time by LTE. WiMAX is still probably the first all IP mobile internet technology allowing true scalability to carry multimedia traffic [11]. WiMAX provides a coverage area of 50 km<sup>2</sup> with data rates of 75 Mbps [12].

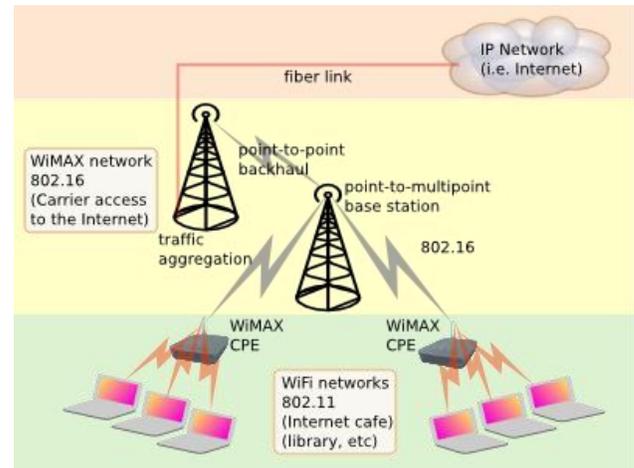


Fig. 1. Application Scenarios of WiMAX. (From: [http://www.accessmillennium.com/images/wifi\\_vs\\_wimax.png](http://www.accessmillennium.com/images/wifi_vs_wimax.png))

WiMAX comes in two types of technologies: the fixed IEEE 802.16/a/d version and the wireless IEEE 802.16-2005 (16e) amendment [13]. The latest version is known as WiMAX rel 2 or IEEE 802.16m. The latest version allows download bitrates up to 1 Gbit/s through channel aggregation for low mobility users.

### III. LITERATURE REVIEW

In another study Mahdi *et al.* [14], [15] investigated the same QoS parameters but for VoIP traffic travelling through UMTS and WiFi alone and together.

A previous simulation study of VoIP over both WiFi and WiMAX [9] has shown that VoIP activity does impact negatively on the overall throughput of both technologies. However, only in the WiFi network is packet loss and jitter experienced. The parameters commonly used to study the performance of the network, for example a study of WiMAX and UMTS using the OPNET network simulation software include: “MOS, end-to-end delay, jitter, and packet delay variation” [7].

It would appear that not all software implementations of VoIP clients are equal - as they vary in their effect on voice quality. This was revealed by a research experiment performed over the High Speed Packet Access (HSPA) [14] service.

To succeed in dealing with the severe problems of VoIP calls over WiFi while approaching the WiFi capacity limit and congestion, a new scheme, the Quality Assurance of Voice over WLANs (SQoSMA) [16] was proposed. SQoSMA took the approach of incorporating the data with the control and planes for detecting and mitigating congestion events. This was achieved by selecting the appropriate adaptive audio codec with the suitable bitrate and then implementing a call stopping method where needed to fix congestions.

An earlier similar scheme [17] was also explained with the use of edge VoIP gateway between the WLAN and the Internet Cloud. The task of the edge VoIP gateway was to determine the pertinent variable speech coding rate (64, 40, 32, 24 and 16 Kbit/s) to lessen the network congestion with a subsequent increase in the overall QoS of speech traffic.

A technique that reduces VoIP traffic's packetization delay (also known as transmission delay or store-and-forward delay) utilized a Transmission Control Protocol (TCP), Friendly Rate Control (TFRC) algorithm based 802.11e network which applied the EDCF (Enhanced Distributed Coordination Function)/HCF (Hybrid Coordination Function) scheme [17].

In [18], authors proposed using a routing and label based solution for transporting real-time VoIP traffic through WLAN which efficiently processed the procedures of call QoS, mobility and call admission. Their procedure utilized a 15 node wireless mesh network to implement distributive packet aggregation utilizing MAC waiting without unbounded packet delays. The fully optimized procedure resulted in a performance gain of 13 times for six hops.

Since human voice is assessed by humans and is therefore purely subjective, a metric to assess this for VoIP traffic is needed that takes into account human subjectivity — which is lacking in the purely objective SNR (signal-to-noise ratio) measure. A study [19] in this field was conducted to look at such metrics concentrating on the E-Model and the Perceptual Evaluation of Speech Quality (PESQ). The researchers studied the limitations of both measures and devised a new metric consolidating the advantages and benefits of them to devise the Advanced Model for Perceptual Evaluation of Speech Quality (AdmPESQ). AdmPESQ is particularly applicable for heterogeneous types of networks with differing delay parameters and packet losses.

The popularity of VoIP has been mushrooming since the last few years. VoIP is now routinely utilized by a wide range of diverse populations globally. While lowering the call price rates, VoIP facilitates almost all the advantages offered by the traditional Public Switched Telephone Network (PSTN). Furthermore, it incorporates several additional value added features. As a consequence of its widespread popularity and such advantages, many companies penetrated into the business of offering various VoIP services. The VoIP traffic, thus, has to pass across several different types of networks — often heterogeneous in nature. Degradation of Quality of Service (QoS) was thus experienced whilst the traffic traverses across such assorted networks. Materna [20], in his research paper “VoIP insecurity”, has enumerated four types of attacks that are relevant to VoIP, viz.:

- Eavesdropping;
- Service integrity;
- Service availability; and
- Spam over Internet Telephony (SPIT).

The successful availability without network outage is vital for the success on any well networked and connected

corporation. Thus protection against any forms of “Service Availability Attacks” is of paramount importance. Downtime in the telephony network will mean: lost revenues for the enterprise and the service providers, unplanned maintenance costs and lost productivity. The IP Telephony network must be protected against all known forms of attacks, which include: viruses, worms and especially the variations of “Denial of Service” (DOS). The effects of these may range from the degradation of the QoS to the total loss (also known as call drops) of the service. Degradation of the QoS is not just a minor nuisance but actually of major concern as customers often request the highest voice quality when they subscribe to an IP Telephony service.

The effect of such an attack on VoIP is actually more sensitive and harmful as it has a lower threshold and immunity than computer data networks. Computer data networks are protected more securely and are usually affected to a lesser degree than the VoIP network. Thus a generic worm may adversely affect the VoIP network precisely because of these reasons, in advance of the computer network. The worm at most, just slow down the computer data network. The worm may, however, totally bring the VoIP network down.

The aim of this research is to ascertain the degree to which the VoIP traffic's quality of service (QoS) deteriorates while traversing through heterogeneous networks. In order to achieve this aim, the authors of this paper, carefully designed, developed and simulated several network scenarios using the OPNET modeler. The results of the various VoIP QoS parameters, thus obtained through the simulation, were then analyzed, reported and published in the literature.

#### IV. RESEARCH METHOD

Due to financial constraints and equipment limitations, the simulation of a sample network, especially in academic research, is very important in the fields of computer networking and telecommunication. Not only does it help to get the perspective view of a network, it also provides guidance for the future. Jack Burbank [21] describes “Modeling and Simulation (M&S)” as an acute constituent in the “design, development and test and evaluation (T&E)” process. As reported by him, “It is almost always preferable to have [an] insight into how a particular system, individual device, or algorithm will [actually] behave and perform in the real world prior to its actual development and deployment” [21]. The advantages of M&S take account of the capability of exercising scenarios and case-studies which are not easily achievable through any empirical methods such as: network scalability testing; the capacity to adapt models to test the systems' sensitivity and to tune its performance [22]. In the case of two or more similar available technologies, it helps to compare and contrast in order to take deployment decisions. This project utilizes and takes advantages of the OPNET Modeler simulation software because it effectively incorporates a wide variety of protocols and technologies [23] while comprising a “development environment”. This smoothes the process of M&S of different types of networks and technologies including (but not limited to): VoIP, WiMAX, WiFi, 3G and LTE. Other networking technologies can be written in software or are available from third party sources.

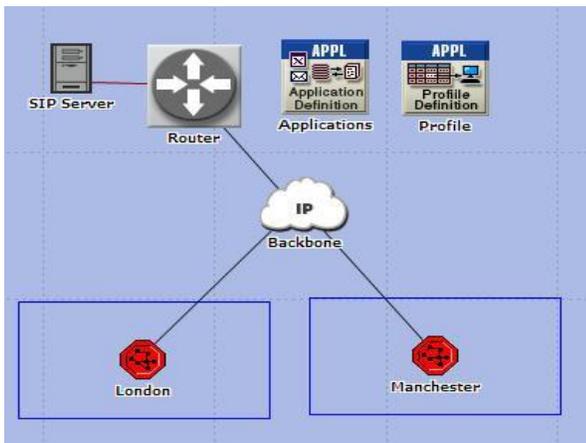


Fig. 2. WiFi network scenario.

In our first simulation scenario, a pair of WiFi subnets, namely London and Manchester, was designed and deployed. As shown in Fig. 2, both the subnets are configured with SIP server credentials connected via an IP cloud.

In our second simulation scenario, a pair of WiMAX subnets, namely Cambridge and Bradford was deployed instead of the WiFi ones. The last scenario replaces one of the WiMAX subnets (namely Bradford) from the second scenario by one of the WiFi subnets (namely Manchester) from the first scenario. Table 3 illustrates some details of the subnets deployed in this research project:

TABLE. III. LIST OF DEVICES USED CONFIGURING THE SUBNETS

Subnet Name	Scenario	Base Station Type	Work Station Type	Number of Work Stations
London	WiFi	WiFi	Mobile	4
Manchester	WiFi	WiFi	Mobile	4
Cambridge	WiMAX	WiMAX	WiMAX Workstation	4
Bradford	WiMAX	WiMAX	WiMAX Workstation	4
Manchester	WiMAX_WiFi	WiFi	Mobile	4
Cambridge	WiMAX_WiFi	WiMAX	WiMAX Workstation	4

It was obviously possible to add more workstations to the scenario, however, we were not interested in the network load, network complexity or routing. Rather, the aim of the research is to find the degradation of the QoS due to the heterogeneous source and destination. The workstations in both of the WiMAX and WiFi network models are configured to facilitate the execution of VoIP applications. The VoIP application, used in this project, is configured to operate as an ‘Interactive Voice’ service and produce one voice frame per packet. The application profile configuration has been set accordingly in order to make this VoIP application operate in a serial mode. A random generation approach was used to make ‘Calls’ to workstations. The ‘Calls’ were exponentially distributed while having an average duration of three minutes. Furthermore, the call inter-arrival periods are also exponentially distributed. In addition to the application profile and application configuration, the WiMAX network model contains a WiMAX

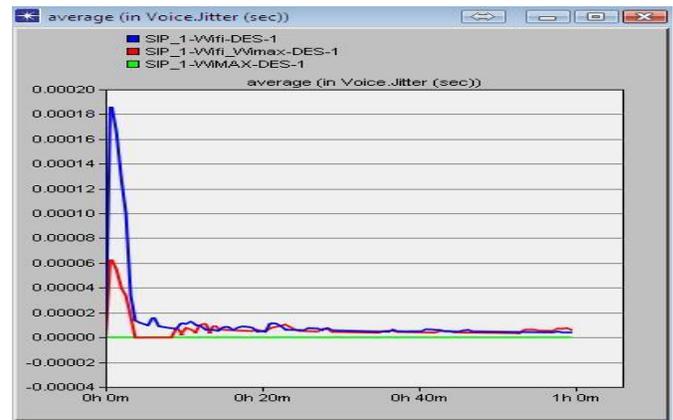
profile. In this profile, a service class of ‘Gold’ with UGS distribution for VoIP application has been created, which was deployed and classified on all the subscriber stations.

V. RESULTS AND DISCUSSION

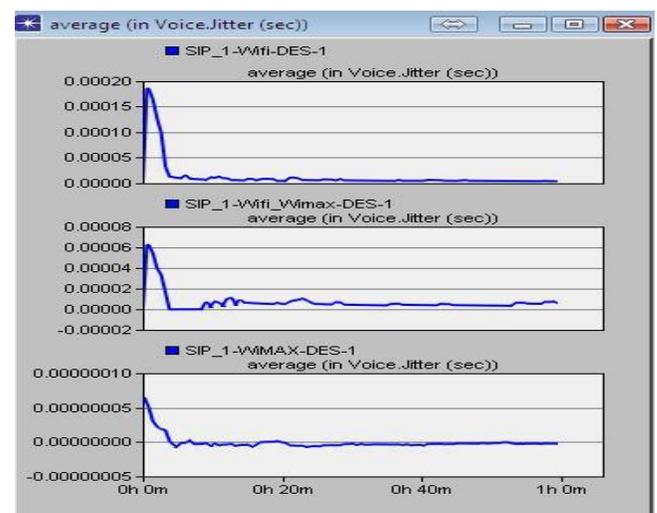
The average jitter graphs, as shown in Fig. 3(a) and 3(b), were obtained from simulating all three scenarios for one hour. They revealed that WiMAX always has better performance over WiFi.

WiFi also suffered from an extreme level of jitter during the initial five minutes, this was likely because of the nature of the convergence period. Although WiMAX, on the other hand, suffered from a similar hike, it was much lower than that observed for WiFi.

The most interesting result we have found is that the average jitter of WiFi-WiMAX scenario, at some points, exceeds that of WiFi. It should ideally always remain somewhere in-between WiFi and WiMAX. Because the simulation was run based on making random calls and no direct handover was associated, this result is very intriguing. However, further research is required to find out the reason(s) behind such a behavior of the WiFi-WiMAX scenario.



(a)



(b)

Fig. 3. (a) Average VoIP Jitter (Overlaid). (Top curve is WiFi, middle curve is WiFi-WiMAX, bottom line is WiMAX). (b) Average VoIP Jitter (Top: WiFi; middle: WiFi-WiMAX; bottom: WiMAX ).

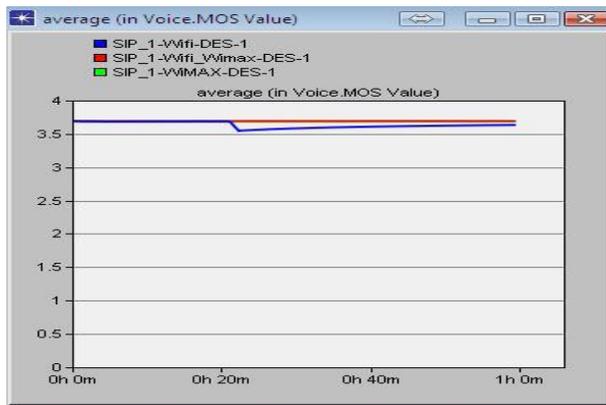


Fig. 4. Average MOS (Overlaid) of 3.7. (Top line is WiFi, bottom line WiFi-WiMAX).

In terms of the MOS, both WiMAX and WiFi-WiMAX observe similar levels of performance, as shown in Fig. 4. Although the call generation was exponentially distributed, the MOS performance of these two networks remains very steady over the whole simulation period.

On the other hand, although at the beginning of the simulation the WiFi network observes a similar level of MOS. However, as time passes, with the increased level of VoIP traffic due to the higher number of calls generated, the MOS decreases. As a result, taking into consideration the MOS, it can be deduced that both WiMAX and WiFi-WiMAX networks outperform the WiFi network. Moreover, although the MOS of the WiFi-WiMAX network scenario should theoretically be at some mid-point in-between the MOS graphs of WiFi and WiMAX, a much higher performance is observed.

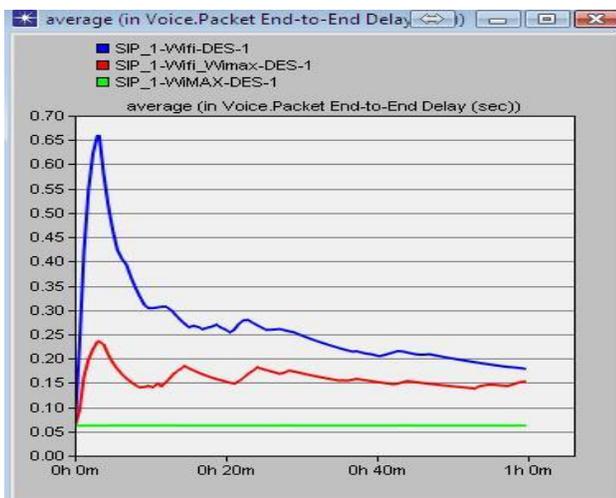


Fig. 5. Packet End-to-End Delay (Top curve is WiFi, middle curve is WiFi-WiMAX and bottom curve is WiMAX).

With regard to the packet end-to-end delay, WiMAX provides better services in comparison with either using just WiFi or WiFi-WiMAX, as illustrated in Fig. 5. In fact, WiMAX constantly remains in the “Good” range, as outlined in Table 2. Although WiFi observes a high level of packet end-to-end delay at the initial setup phase, it reaches and remains within the “Acceptable” band after the network has

converged. The WiFi-WiMAX network remains within the “Acceptable” band even during the convergence period.

## VI. CONCLUDING DISCUSSION

The paper presented the early findings related to VoIP traffic transmitted through WiFi, WiMAX and WiFi-WiMAX networks. Initially, two scenarios were designed where both generation and termination of the VoIP calls take place in an environment of homogenous networks such as WiFi and WiMAX. Another scenario was later added where calls were generated at the WiFi network and terminated at WiMAX networks and vice-versa.

One of the most thought-provoking findings of our research is regarding the average jitter value of the WiFi-WiMAX scenario of not being in-between WiFi and WiMAX. Our research shows that it does not always perform as expected; even, at some points in time, it exceeds that of WiFi.

The MOS of the WiFi-WiMAX network should ideally be somewhere near halfway of the WiFi and WiMAX MOS graphs. Our research has found that it exhibits a much higher performance than that. Similarly, the packet end-to-end delay of WiFi-WiMAX remains close to that of WiFi and is much higher than expected.

Since there are still a number of WiMAX providers, the study could be strengthened further if comparison between the simulation results against the corresponding results of a real deployment could actually be made. However, due to business and security reasons, companies tend not to reveal their data to the public. If any such data is received, we have plans to compare our results against them.

Future work will include other networks covering: GSM, GPRS, EDGE, UMTS (3G), CDMA, LTE and 4G. The analysis of such QoS parameters for Voice-over LTE (VoLTE) will be one of the particular future research directions. The effect of handover covering, soft, softer and hard on the network traffic will also be focused upon in the future works of this continuing research project.

Furthermore, the scope of this study will be broadened by including the investigation of the impact on other QoS parameters e.g. the packet drop rate, queuing delay and the throughput. To find out the reasons affecting the behavior of these parameters, they will be meticulously examined with the goal of attaining a better optimization and improved efficiency of the network.

## REFERENCES

- [1] Ralf Ebbinhaus, “VoIP lessons: Look who's calling,” IET Communications Engineer, vol. 1, no. 5, pp. 28 - 31, October-November 2013. [Online]. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1244763>
- [2] Floriana Gere, “Security in VoIP,” Journal of Knowledge Management, Economics and Information Technology, vol. 2, no. 2, pp. 1-10, April 2012. [Online]. <http://www.scientificpapers.org/information-technology/security-in-voip/>
- [3] Victor Y. H. Kueh, Rahim Tafazolli, and Barry G. Evans, “Performance analysis of session initiation protocol based call set-up over satellite-UMTS network,” Computer Communications, vol. 28, no. 12, pp. 1416-1427, July 2005. [Online]. <http://www.sciencedirect.com/science/article/pii/S0140366405000502>

- [4] Shaffatul Islam, Mamunur Rashid, and Mohammed Tarique, "Performance Analysis of WiMax/WiFi System under Different Codecs," *International Journal of Computer Applications*, vol. 18, no. 6, pp. 13-19, March 2011. [Online]. <http://www.ijcaonline.org/archives/volume18/number6/2290-2973>
- [5] J. Rosenberg et al., "SIP: Session Initiation Protocol," IETF RFC Editor, pp. 1-194, February 2002. [Online]. <http://ietfreport.isoc.org/all-ids/draft-ietf-sip-rfc2543bis-08.pdf>
- [6] Raj Pandya, "Emerging Mobile and Personal Communication Systems," *IEEE Communications Magazine*, vol. 33, no. 6, pp. 44-52, June 1995. [Online]. <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00387549>
- [7] Sheetal Jadhav, Haibo Zhang, and Zhiyi Huang, "Performance Evaluation of Quality of VoIP in WiMAX and UMTS," in *Proceedings of the 12th IEEE International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)*, 2011, Gwangju, South Korea, 20-22 October 2011, pp. 375 - 380. [Online]. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6118553>
- [8] James Yu and Imad Al-Ajarmeh, "Call Admission Control and Traffic Engineering of VoIP," in *Proceedings of the The Second IEEE International Conference on Digital Telecommunications (ICDT 2007)*, San Jose, California, USA, 1-5 July 2007, pp. 11-16. [Online]. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4270577>
- [9] M. Atif Qureshi et al., "Comparative study of VoIP over WiMAX and Wi-Fi," *International Journal of Computer Science Issues (IJCSI)*, vol. 8, no. 3, pp. 433-437, May 2011. [Online]. <http://ijcsi.org/papers/IJCSI-8-3-1-433-437.pdf>
- [10] F Ohrtman, *WiMAX Handbook: Building 802.16 Wireless Networks*. New York: McGraw-Hill, 2005.
- [11] (2009) *Delevering Broadband on the Go: WiMAX Technology*. [Online]. <http://www.intel.com/technology/wimax/>
- [12] Will Hrudehy and Ljiljana Trajković, "Mobile WiMAX MAC and PHY layer optimization for IPTV," *Mathematical and Computer Modelling*, vol. 53, no. 11-12, pp. 2119-2135, June 2011. [Online]. <http://www.sciencedirect.com/science/article/pii/S0895717710003821>
- [13] "Air Interface for Fixed and Mobile Broadband Wireless Access Systems," *IEEE 802.16e-2005*, 2005.
- [14] Mahdi H. Miraz, Suhail A. Molvi, Maaruf Ali, Muzafar A. Ganie, and AbdelRahman H. Hussein, "Analysis of QoS of VoIP Traffic through WiFi-UMTS Networks," in *Proceedings of the World Congress on Engineerings (WCE) 2014*, vol. 1, Imperial College, London, United Kingdom, 2-4 July, 2014, pp. 684-689.
- [15] Mahdi Hassan Miraz, Muzafar Aziz Ganie, Maaruf Ali, Suhail Ahmed Molvi, and AbdelRahman Hamza Hussein, "Performance Evaluation of VoIP QoS Parameters Using WiFi-UMTS Networks," in *Transactions on Engineering Technologies*, Gi-Chul Yang, Sio-Iong Ao, and Len Gelman, Eds. New York, USA: Springer-Verlag, 2015, ch. 38, pp. 547-561. [Online]. [http://link.springer.com/chapter/10.1007/978-94-017-9804-4\\_38](http://link.springer.com/chapter/10.1007/978-94-017-9804-4_38)
- [16] Badis Tebbani, Kamel Haddadou, and Guy Pujolle, "Quality assurance of voice over WLANs (VoWLANs) with differentiated services," *Institut Télécom and Springer-Verlag*, 2010.
- [17] Abdelbasset Trad, Qiang Ni, and Hossam Afifi, "Adaptive VoIP Transmission over Heterogeneous Wired/Wireless Networks," *INRIA, Planete Project 2004 Route des Lucioles, BP-93, 06902 Sophia-Antipolis, France*.
- [18] Dragos Niculescu, Samrat Ganguly, Kyungtae Kim, and Rauf Izmailov, "Performance of VoIP in a 802.11 Wireless Mesh Network," *NEC Laboratories America, Princeton, NJ*.
- [19] David Rodrigues, Eduardo Cerqueira, and Edmundo Monteiro, "QoE Assessment of VoIP in Next Generation Networks," *University of Coimbra, Department of Informatics Engineering, 3030-290 Coimbra, Portugal, Federal University of Para, Rua Augusto Corra, 01, 66075-110, Bel'em, Brazil*.
- [20] Bogdan Materna, "VoIP insecurity," *IET Communications Engineer*, vol. 4, no. 5, pp. 39-42, October/November 2006. [Online]. <http://ieeexplore.ieee.org/xpl/abstractAuthors.jsp?reload=true&arnumber=4049981>
- [21] Burbank, "Modeling and Simulation: A Practical Guide for Network Designers and Developers," *IEEE Communications Magazine*, vol. 47, no. 3, p. 118, March 2009.
- [22] A Maria, "Introduction to Modeling and Simulation," in *Winter Simulation Conference*, 1997.
- [23] OPNET Modeler. [Online]. [http://www.opnet.com/solutions/network\\_rd/modeler.ht](http://www.opnet.com/solutions/network_rd/modeler.ht)

# A Comparative Analysis of Quality Assurance of Mobile Applications using Automated Testing Tools

Haneen Anjum

Department of CS & SE,  
International Islamic University,  
Islamabad, Pakistan

Maham Khan

Department of CS & SE,  
International Islamic University,  
Islamabad, Pakistan

Zainab Shahid

Department of CS & SE,  
International Islamic University,  
Islamabad, Pakistan

Muhammad Imran Babar

Army Public College of  
Management & Sciences, Pakistan  
Corresponding Author

Saima Chaudhry

Department of CS & SE,  
International Islamic University,  
Islamabad, Pakistan

Furkh Zeshan

Department of CS, COMSATS,  
Lahore,  
Pakistan

Muhammad Jehanzeb

Army Public College of  
Management & Sciences,  
Pakistan

Summiyah Sultana

Department of CS & SE,  
International Islamic University,  
Islamabad, Pakistan

Shahid Nazir Bhatti

Department of SE, Bahria  
University, Islamabad,  
Pakistan

**Abstract**—Use of mobile applications are trending these days due to adoption of handheld mobile devices with operating systems such as Android, iOS and Windows. Delivering quality mobile apps is as important as in any other web or desktop application. Simplification and ease of quality assurance or evaluation in mobile devices is achieved by using automated testing tools. These tools have been evaluated for their features, platforms, code coverage, and efficiency. However, they have not been evaluated and compared to each other for different quality attributes they can enhance in the apps under test. This research study aims to evaluate different testing tools focusing on identifying quality factors they aid to achieve in the apps under test. Furthermore, it aims to measure overall trends of essential quality factors achieved using automated testing tools. The findings of this study are beneficial to the practitioners and researchers. The practitioners need to look up for specific tools which aid them to assure the desired quality factors in the apps under test. The researchers may base their studies on the findings of this study to propose solutions or revise existing tools in order to achieve maximum number of critical quality attributes in the app under test. This study revealed that the trend of automated testing is high on usability, correctness and robustness. Moreover, the trend is average on testability and performance. However, for assurance of extensibility, maintainability, scalability, and platform compatibility, only a few tools are available.

**Keywords**—Mobile application; quality assurance; automated testing; testing tools

## I. INTRODUCTION

Software testing enables the software testers to detect defects in the software and remove them to ultimately achieve improved software quality. Recently software testing became wide-spread and critical among software development companies. Software testing can be performed either manually

or automatically. Manual testing is to manually write the test cases and executing them without using any tool. In manual testing a tester performs the testing through carefully navigating through the different interfaces of the system under test, testing with different values of inputs, recording and comparing the observed results with the expected results of the tests.

Automated testing is done with the help of an automated testing tool. The automated testing tool provides a computer-controlled testing rather than manually. The testing tool executes the test cases to test the performance and functionality of the software under test. The aim of automated testing is to reduce the required human effort as in manual testing but it does not remove the need of manual testing at all [1]. Mobile platforms are being adopted worldwide because of a variety of software being offered to users in those handheld and portable devices. Testing is being used as a quality assurance technique for mobile apps too [2].

Several tools are proposed and implemented for this purpose. These tools have already been evaluated and compared for their unique features, supported platforms, code coverage, and efficiency. However, existing automated testing tools of mobile applications have not been evaluated and compared for different quality attributes they can enhance in apps under test. Therefore, two research objectives are formulated for this study that is: 1) to evaluate different testing tools of mobile apps focusing on identifying quality factors they aid to achieve in the apps under test; 2) to measure overall trends of essential quality factors achieved in the mobile apps under test using automated testing tools. In this paper, we have evaluated and compared automated testing tools for adding or enhancing valuable quality factors in mobile applications under test. The findings and result of this study are beneficial to the

practitioners as well as the researchers. The list of quality factors to be achieved varies among apps. The testing of different apps requires selection of different tools. Therefore, the practitioners may need to look up for tools which aid them to assure the desired quality factors in a particular App Under Test (AUT). The researchers who are interested in proposing the tools and techniques for testing of mobile apps may need to consider the quality factors highlighted in this study. Moreover, they can begin their own research study on the basis of these tools to propose merged, revised and enhanced solutions for achieving the maximum number of quality attributes in the AUT.

The rest of this paper is structured as follows: Section 2 gives a comprehensive knowledge about the background concepts of manual and automated software testing. Section 3 describes methodology that we used to achieve our research objectives. Section 4 presents description of a number of automated testing tools for mobile applications. Section 5 presents comparative study. Section 6 presents our findings and discussion. Finally, Section 7 concludes the paper.

## II. BACKGROUND

Success of any software product is determined by the quality of that software. This gives software quality assurance a great opportunity in software industry and customer satisfaction drives it. To develop a product of good quality and without any defects within the cost and time constraints have become critical. Implementing such products, with minimum or no bugs is a difficult task. This is the reason that the concept of software testing has got its existence [3]. In software industry, testing of software has become an extensive and vital phase of SDLC. It also provides final evaluation of other activities such as requirements specification, software design, and coding [4].

Software testing is an activity, which is performed to evaluate correctness and functionality of software for assuring fulfillment of user requirements and expected quality [5]. IEEE defines software testing as the process to evaluate the system or its components manually or by automated means to determine whether it fulfills the user requirements or to find the difference among actual result and expected result [6]. Hence, the software testing is to execute a software to identify defects or any missing features that were expected by the user requirements. Software testing results in improved quality and effectiveness of the software system, if it is executed appropriately. Detecting the defects in a software and removing those defects before the release of software leads to reduced maintenance cost.

All the activities of software testing can be conducted by two means: automated testing and manual testing. Manual testing is the fundamental software testing. It is conducted manually through moving about in the software application. A test plan or test cases are followed for manual testing. Test cases describe the complete test scenario in terms of actions to be performed during testing. On the other hand, in automated testing, the testing is conducted through some testing tool without the navigating through the different parts of the application manually.

Initially, manual testing was only performed. Because of human error, few defects may be ignored or unidentified through manual testing. So, through manual testing better quality of a software system cannot be ensured. To overcome this lack in manual testing, automated testing has evolved. The automated testing is helpful in quicker testing process. Recently automated testing got more attention and many testers prefer to use automated testing for the variety of software systems [7]. The basic element behind automated testing is the automated testing tool that is used to conduct the tests.

### A. Software Testing

Normally software testing is considered as an activity for detection of defects whereas there are different reasons behind conduction of software testing. Improved software quality is one of the major reasons. Software quality is improved by ensuring that the software product fulfils the user requirements and expectations. Smooth functioning of the software system can be ensured through testing. The software developing industries spend most of their time and cost on software testing during the SDLC [8]. If the testing is done early in the SDLC to prevent the occurrence of defects, it reduces the time and cost spent whereas, if the defects are detected in later stages, then the time to market and cost rises significantly. Therefore, performing testing throughout the SDLC is a better practice to detect the defects of the software. It is less expensive to remove the defects earlier, even before the release of the software [9].

Software testing aims to evaluate the capabilities of an application or the software and verify that it fulfils the quality principles such as reliability, portability, efficiency, security, usability, etc. Through testing all these principles should also be verified and ensured [10]. There are two main objectives of software testing. First, the detection of errors or defects. Second, preventing the number of occurrences of defects in the software system, that results in overall improved efficiency of the system.

### B. Manual Software Testing

Manual testing is the simplest level of testing in which the tests are executed as per test cases and by directly interacting with the software. In this testing, the tester prepares the test cases. Test cases, are the explanations of the features and the expected results of the software under test, and are written in simple natural language. The process of manual testing becomes too much time-taking as it requires all the activities to be performed manually. Though, manual testing is preferred in case of some complex systems where a few critical defects can only be discovered while testing manually. During manual testing the tester interacts with the system under test as the end user of that software would, and ensures the effectiveness of the system by navigating through the software [11]. Manual testing have the following drawbacks [12]:

- Time-taking
- Requires more testers
- Less accurate results
- Testing multiple features in parallel, not possible
- Lack of reusability of tests

- Lack of test completeness.

### C. Automated Software Testing

As the automated software testing got popular in software industries, the testing process become more effective. Automated software testing helps in easily executing various tests like performance testing and regression testing. The difficult testing activities got easier than before, as the automated testing evolved and improved, because it conducts the test for various datasets and the tests can be executed repeatedly without human involvement [1]. Automated software testing requires a little primary investment for the software but that doesn't have much economical effect as it results in reduced human efforts required for testing [13]. The automated software testing can be performed in various phases: preparation of test plan or developing the test cases, selecting the testing tool, creation of the test script and finally executing the test by using the automated testing tool and the script.

The main objective of automating software testing is to reduce the testing effort, time and cost. Testing automation results in improved efficiency, whereas reduction in human involvement in testing process. Automated testing supports the reusability of test scripts, using the testing tool, for different upgrades of the system under test [1]. Automated software testing simplifies the testing process and results in reduced maintenance cost of the software [10]. Automated testing has the following benefits [7]:

- Simplified regression testing
- Tests are repeatable and reusable
- Reduces time and cost

- Performance testing is possible due to simultaneous testing.

Automated testing has the following drawbacks [12]:

- It is more expensive
- All areas cannot be automated
- Manual testing cannot be fully discarded.

### D. Manual vs Automated Software Testing

Table 1 illustrates the differences between manual and automated software testing [1], [7], [12].

## III. METHODOLOGY

For mobile applications, nine essential software quality factors, as described in Table 2, are selected. These factors are the most significant quality attributes not only in software and web based applications, but also the mobile apps must conform to these quality requirements. Firstly, all industry-dominant and proposed mobile apps testing tools are identified from existing literature from 2010 to 2017. Secondly, each of these tools is studied in order to extract its features. Thirdly, for each tool, the quality factors it may aid to achieve in AUT are derived on the basis of its features and characteristics. All the derived and implied quality factors for each tool form a subset of the set of factors mentioned in Table 2. The tools are compared on the basis of their quality factors in Section 5. Moreover, for each tool, the derivation of the quality factors is also justified based on its features and characteristics. The summarized results of this comparative study are presented graphically in Section 6 to show an overall trend of quality factors achieved using automated testing.

TABLE. I. DIFFERENCES BETWEEN MANUAL AND AUTOMATED SOFTWARE TESTING

Manual Testing	Automated Testing
1. Time Consuming	Time Efficient
2. More human effort is required.	One-time human effort for creating the test scripts is enough.
3. Not accurate, due to room for human errors	More accuracy as less space for human error
4. Test cases cannot be reused	Supports reusability of test cases
5. More effective for functional testing and exploratory testing	Effective for regression testing, load testing & performance testing
6. Reduced short term cost (no automated testing tool is required) while increased long term cost (maintenance).	Increased short term cost (automated testing tool) while reduced long term cost (maintenance).

TABLE. II. SOFTWARE QUALITY FACTORS FOR COMPARATIVE ANALYSIS

Software Quality Factors	Description
Extensibility	Ability of software components to be added, modified and removed easily without badly effecting existing system. Flexibility is its category focused on ability of components to be added easily.
Maintainability	Maintainability is ability to make change for error corrections, supported by defined interfaces, documentations, comments in code.
Performance	Performance is related to acceptable response time.
Scalability	Ability to respond in an acceptable time in increased load or stress.
Robustness	Robustness is the ability of software to keep working and remain available in failure states by backup plans, data and hardware.
Usability	Usability is the ability of user to easily interact with the system using the user interface.
Platform compatibility	Software should run on several platforms like operating systems, browsers etc.
Testability	Testability refers to maximum and efficient code coverage by testing.
Correctness	Correctness is software should conform to with requirements or specifications.

#### IV. AUTOMATED TESTING TOOLS FOR MOBILE APPLICATIONS

In Software Development, Mobile Applications Development is a prominent area which is emerging rapidly. Therefore, testing also becomes significant in this area. Many tools are available for supporting different types and levels of testing in platforms like Android and iOS [14]. Following are some noteworthy tools that are being used in Software industry for their strong testing support for mobile apps. Robotium is one of the UI automation frameworks used for android systems. It is available free of cost in the market and can be used by enterprises and individuals as well. It assists the test case developers in writing functional, acceptance and system test scenarios, spanning a range of android activities. It is a Java based tool while JUnit test framework is a part of it as well. It is made to make it easy for test case developers to write robust and powerful automatic black box test cases. This tool cannot be used for Web or Flash apps [14].

Renorex is a testing tool and framework that supports the scriptless way of working and coding capabilities. This tool is mainly used for GUI supports in mobile and web apps. It offers a fast and intuitive way to write test cases as functions used in SUT. It gives some extra ability for creation of a robust regression testing. It supports cross browser testing too. The Renorex studio IDE delivers a feature ‘click and go Function’ in order to ensure the reusability of test actions and various UI element with the team of technical skill levels [12]. Appium is another cross-platform testing tool that allows test case developers to write test for multiple platforms such as iOS and android, using a single API. It enables code to be reused among iOS and android test suites. It is an open source tool used for web app and hybrid application of automating native mobile on both the iOS and android platforms, where the native apps can be written using android SDK or iOS [14].

MonkeyTalk is an open source tool used for functional testing. It is simple to use and powerful tool for testing mobile applications. This tool works with a range of real devices and emulators. It tests from a simple ‘smoke test’ to the sophisticated test suites such as data driven test suites. The tests are created for iOS and android if the parameterized tests are used. MonkeyTalk IDE is an eclipse based tool for recording, playing, editing and managing the functional test suites for iOS and Android applications that runs on emulators, simulators and devices [14]. UIAutomator is one of the testing

frameworks provided by Google’s Android. The tests run by this framework ensures an application to meet the functional requirements and it achieves a fine standard quality so that it can be successfully adopted by android users. It allows to run the tests reliable, fast, and repeatable manner [14].

Reran is a record and replay tool for smartphones that have Android operating system. It captures input event sent from the phone to the OS of a user session and after that allows the sequence of events to be sent into the phone programmatically at high level. Reran captures the low level events and replays them that are triggered on the phone, which allows it to capture and playback GUI events such as touchscreen gestures, and input sensors on device [15]. EvoDroid is used to test system of Android apps. It combines two techniques 1) to identify parts of the code open to be searched independently an android-specific program analysis; 2) an algorithm performs search step by step under the given info. Its main goal is to look for test cases that amplify code coverage [16]. MobiGUITAR models the state of the app’s GUI, which helps us more accurately model mobile apps’ state-sensitive behaviour. On the basis of state machine, it makes new test adequacy criteria. This test generation technique uses the models and criteria to generate test cases automatically. It delivers fully automatic testing that works on security policies of smartphone platforms [2].

Dynodroid automatically generates inputs to Android apps. It is capable of generating both UI inputs (e.g., touchscreen taps and gestures) and system inputs (e.g., simulating incoming SMS messages). It allows interleaving inputs from machine and human. Through a sequence of events it interacts with its environment. Dynodroid is an observe-select-execute cycle, it observes which events are important to current state, selects those events, and execute those events to make a new state in which it repeats this process [17]. FSMdroid is a guided approach to GUI testing of Android apps. Its basic idea is to 1) construct an initial stochastic model for the app under test; 2) iteratively mutate the stochastic model and derive tests. Compared with the traditional model-based testing approaches, it enhances the diversity of test sequences by 85%, but reduces the number of them by 54%. It first uses static analysis to identify UI events which can be missed during dynamic analysis [18]. Table 3 summarizes general information about above testing tools i.e. their support for testing types or levels, platform. According to Table 3, 90% of the tools support automated testing of Android apps. However, 20% of the tools support testing of the iOS apps.

TABLE III. AUTOMATED MOBILE APPLICATIONS TESTING TOOL

Testing Tool	Testing Type	Platform
Dynodroid	Event driven testing	Android
Evodroid	System testing	Android
FSM Droid	GUI testing	Android
MobiGUITAR	GUI testing	Android
Renorax	Compatibility testing	C#, Python, VB.net
Reran	GUI, system, stress, and security testing	Android
Robotium	GUI, system, functional, and acceptance testing	Android
Appium	GUI and functional testing	Android, IOS
MonkeyTalk	Compatibility and functional testing	Android, IOS
UIAutomator	Functional and GUI testing	Android

## V. COMPARATIVE ANALYSIS OF SOFTWARE TESTING TOOLS

The purpose of testing is to ensure that software meets its functional requirements and it is of desired or standard quality so that it is accepted and adopted by the user for its intended use [14]. Aforementioned tools are proficient in one or more from functional testing, system testing, code coverage and user interface testing, etc. of mobile applications. This section presents their comparative analysis on the basis of quality factors from Table 2 they test and thus enhance in mobile apps under test.

Dynodroid smartly plays the role of user of mobile app under test by generating input events automatically [17], thus giving an illusion of actual interaction of user with the application in expected environment. For each auto generated user event, this tool observes reaction of the application to further generate next possible event that could be performed by the user [17]. This proficiency makes Dynodroid fit for evaluating mobile applications for their usability. The test reports can help front-end developers to improve usability by reshaping the possible interaction with user while still fulfilling his needs. Furthermore, it allows tester's intervention at any stage for entering relevant and intelligent input in any sequence of events [17] to evaluate correctness of application. Results of these customized tests reveal the level of correctness achieved in application so that further conformance to requirements can be achieved. Studies have proved that this tool also finds bugs [17] that may crash the application, which are corrected by developers. Thus, this tool contributes to reliability and robustness of solution just tested. If the promised percentage of source code is covered under tests [17], then it shows that the testability of software is achieved. If there is less source code coverage, then the application has not attained the quality factor of testability.

Evodroid aims to perform system testing of mobile applications [16]. System testing exercises application for its overall behaviour to check correctness, so as to state that application fits for its intended user. It offers much higher code coverage [16] which can easily evaluate testability in an application. Despite of higher coverage of code being offered, if not a good percentage of code is being covered, then the application's design must be modified to reduce testing effort. So, in complicated solutions, other quality factors like correctness, robustness, maintainability, etc. can be evaluated after deployment also. Evodroid effectively provides features of deploying, maintaining, and enhancing mobile applications [16]. Thus, it adds to correctness, flexibility, and maintainability of apps by following its methods and tips of utilizing these features. FSMdroid focuses on Graphical User Interface (GUI) testing [18]. GUI is the interaction point between user and system. When GUI is tested for prompting input, displaying output and scenarios of erroneous inputs from user, it ultimately gives good evaluation of usability and accessibility of application's features under test. It also evaluates testability as it also offers high coverage of code [18]. It also reveals fatal bugs in code [18], which must be solved with proper handling of exceptional error scenarios in code. In this way, it contributes to robustness of application under test. It helps to make GUI models which consume minimum events

[18], thus improving performance of application by avoiding duplication and complex GUI events sequences.

MobiGUITAR helps to model state of an application's GUI to test behaviour at a particular sensitive state of GUI [2]. This feature lets testers monitor correctness of an application by mapping response or behaviour with GUI events and states. This tool is proficient in finding concurrency error [2] that may lead to severe concurrency issues, fatal errors, and crashes. It highlights other logical errors [2] too. All these errors are fixed for achieving robustness, fault tolerance and reliability in application being tested. It adds to testability also by its acceptable code coverage [2]. Renorax performs platform compatibility testing [12] which assures that the software is of good quality in terms of its diverse usage on a variety of famous platforms and configurations like operating systems, browsers, web programming languages, etc. It also adds features for supporting further addition and enhancement [12] thus adding flexibility factor for easy maintenance and updates of the software.

Reran tests applications which take user inputs from device sensors and sophisticated GUI operations [15] like zoom, tap, swipe, etc. Reran evaluates the usability of application with all complex application and system level events from rich controls of GUI and sensors. Such application should perform with greater accuracy and precision of time [15] due to sudden inputs from sensors. Reran evaluates the performance and efficiency of application by its strong testing support. Bugs indicated during debugging [15] and test results are corrected by developers which ultimately adds to correctness, performance and robustness of the application under test. It also performs stress testing [15] which evaluates the scalability of application for achieving optimum quality under stress or load conditions. It also catches security related bugs caught after invalid user inputs or malicious plugins [15] to give clues to developers to not leave any vulnerability and make the app and its data secure.

Robotium also supports a good evaluation of usability by performing tests on rich GUI controls of a touch screen mobile device [14]. Test results of function, system and acceptance testing [14] on Robotium allows developers to improve correctness and performance of applications to an optimum level. Appium focuses on testing interaction of user with the content of mobile web applications. Automated test cases are configurable with Safari and Chrome web browsers [14]. Test results are used to evaluate correctness and user experience with the mobile web application in terms of usability or accessibility of the features.

If an application is platform independent or cross-platform, it means it is applicable for a diverse use on different operating systems. It is a plus point to check quality factor of platform compatibility. MonkeyTalk serves this purpose to perform tests for mobile application's compatibility with iOS and Android by offering cross platform testing [14]. It creates test scripts to perform functionality tests against action of each user interface event or command [14]. UIAutomator ensures that mobile app under test is a quality app considering factors of correctness and usability by performing UI functional testing. It automates user test cases to reflect user experience and correctness of

behaviour against input entry and events in asynchronous GUIs like dialogs, alerts, etc. also. This tool is proficient in automating functional UI tests even on two or more devices. [14]

VI. FINDINGS AND DISCUSSIONS

According to Table 4, Evodroid and Renorax aid to achieve quality factors ‘extensibility’ and ‘maintainability’. ‘Performance’ of the AUT can be enhanced by using three tools i.e. FSM Droid, Reran, and Robotium. Among all the tools, only Reran aims to achieve ‘scalability’ of the AUT. ‘Robustness’ can be achieved by five tools i.e. Dynodroid, Evodroid, FSM Droid, MobiGUITAR and Reran. Many of the tools assure ‘usability’ of the AUT i.e. Dynodroid, FSM Droid, Reran, Robotium, Appium, and UIAutomator. The ‘platform compatibility’ testing is supported by only two tools, namely, Renorax and MonkeyTalk. ‘Testability’ of the app can be verified and enhanced using four tools, namely, Dynodroid,

Evodroid, FSM Droid, and MobiGUITAR. Most of the tools, namely, Dynodroid, Evodroid, MobiGUITAR, Reran, Robotium, Appium, MonkeyTalk, and UIAutomator assure the ‘correctness’ of the AUT.

Fig. 1 presents a graph showing the results of this comparative study. Ten dominant automated testing tools for mobile applications are considered for this study. Each tool focused one or more quality factors to achieve or enhance quality of apps under test. Moreover, Fig. 1 shows an overall trend of quality factors achieved by using automated testing. The quality factor of ‘correctness’ will be achieved using almost every automated testing tool. ‘Usability’ is also a major aspect of mobile apps which can be evaluated and achieved by using approximately 60% of the available software testing tools. Approximately 50% of these tools focus on achieving desired or optimum level of ‘robustness’ of mobile apps. A close to average percentage of testing tools attain quality factors of ‘testability’ and ‘performance’ in the app under test.

TABLE IV. QUALITY FACTORS ACHIEVED BY AUTOMATED TESTING OF MOBILE APPS

Software Testing Tools	Software Quality Factors								
	Extensibility	Maintainability	Performance	Scalability	Robustness	Usability	Platform compatibility	Testability	Correctness
Dynodroid					✓	✓		✓	✓
Evodroid	✓	✓			✓			✓	✓
FSM Droid			✓		✓	✓		✓	
MobiGUITAR					✓			✓	✓
Renorax	✓	✓					✓		
Reran			✓	✓	✓	✓			✓
Robotium			✓			✓			✓
Appium						✓			✓
MonkeyTalk							✓		✓
UIAutomator						✓			✓

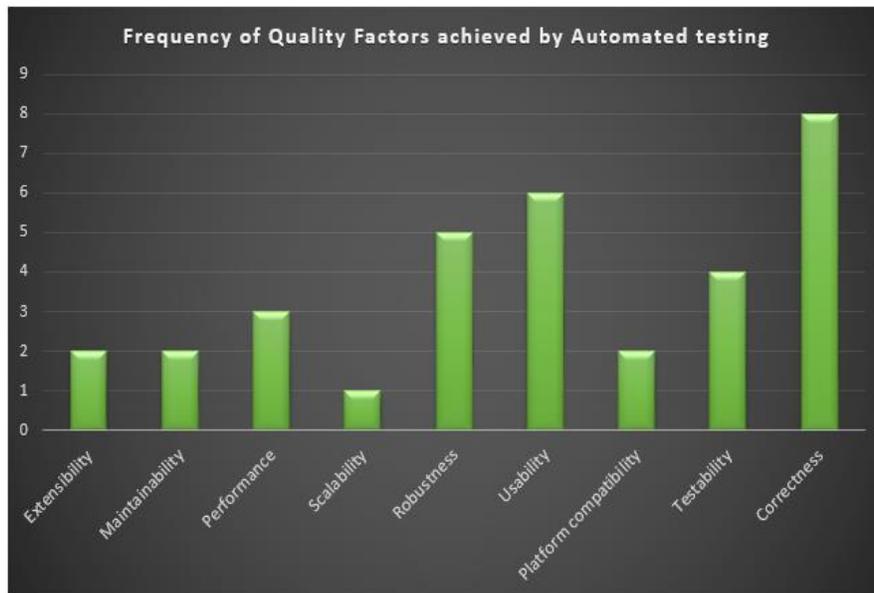


Fig. 1. Frequency of software quality factors achieved using automated testing tools for mobile applications.

A lesser percentage of tools are observed for other important quality attributes like extensibility, maintainability, scalability, and platform compatibility. Reran and Evodroid are better than other tools because they tend to achieve five out of nine quality factors. The rest of the eight tools help to achieve less than five quality attributes in the AUT. Therefore, it is recommendable that for an AUT, more than one tool should be used to assure all the critical quality factors.

There is no automated testing tool or solution for mobile apps which tests for all possible quality factors that are mentioned in Table 2. Most testing tools cover only usability, correctness and robustness, which are desired by almost every mobile app. To support incremental development with testing, and post deployment maintainability and flexibility, only a few tools serve this purpose. Therefore, trend of automated testing is high on usability, correctness and robustness, average on testability and performance, and lesser on extensibility, maintainability, scalability, and platform compatibility.

## VII. CONCLUSION AND FUTURE WORK

There is no mobile app testing tool which tests for all possible quality factors. Most testing tools cover only usability, correctness and robustness, which are desired by almost every mobile app. To support incremental development with testing, and post deployment maintainability and flexibility, only a few tools serve this purpose. Trend of automated testing is high on usability, correctness and robustness, average on of testability and performance, and lesser on extensibility, maintainability, scalability, and platform compatibility. In automated testing of mobile applications, further research can be done to propose automated mobile apps testing tool that aims to achieve all quality factors mentioned in Table 1. A similar analysis can be made by considering testing tools for other mobile operating systems as well like windows. A comparative analysis can also be done on quality of apps of different mobile operating systems based on automated testing tools of each platform.

Several tools are proposed and implemented for testing of mobile apps. In this research study, these tools are evaluated focusing on identifying the quality factors they aid to achieve in the apps under test. Moreover, overall trends of essential quality factors achieved using automated testing tools are measured. This study revealed that the automated testing provides best support for assurance of usability, correctness and robustness. An average number of tools aid to assure testability and performance. However, for assurance of extensibility, maintainability, scalability, and platform compatibility, only a few tools are available. In automated testing of mobile applications, further research can be done to propose automated mobile apps testing tool which aims to achieve all quality factors mentioned in Table 2. A similar analysis can be made by considering testing tools for other mobile operating systems too, e.g., windows. A comparative analysis can also be done on quality of apps of different mobile operating systems based on automated testing tools of each platform. Moreover, on the basis of the tools identified from this study, revised and enhanced solutions can be proposed for achieving the maximum number of quality attributes in the AUT.

## ACKNOWLEDGEMENT

Special thanks to International Islamic University, Islamabad Pakistan and Army Public College of Management & Sciences, Rawalpindi, Pakistan for providing support in order to complete this research.

## REFERENCES

- [1] P. Rathi and V. Mehra, "Analysis of Automation and Manual Testing Using Software Testing Tool," 2015.
- [2] D. Amalfitano, et al., "MobiGUITAR: Automated model-based testing of mobile apps," IEEE Software, vol. 32, pp. 53-59, 2015.
- [3] X. Wang and G. He, "The research of data-driven testing based on QTP," in 2014 9th International Conference on Computer Science & Education, 2014.
- [4] M. Monier and M. M. El-mahdy, "Evaluation of automated web testing tools," International Journal of Computer Applications Technology and Research, vol. 4, 2015.
- [5] K. M. Mustafa, et al., "Classification of software testing tools based on the software testing methods," in Proceedings of the International Conference on Computer and Electrical Engineering (ICCEE'09), 2009, pp. 229-233.
- [6] G. Saini and K. Rai, "Software Testing Techniques for Test Cases Generation," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, 2013.
- [7] N. Islam, "A Comparative Study of Automated Software Testing Tools," 2016.
- [8] D. Shikha and K. Bahl, "Software Testing Tools & Techniques for Web Applications,"
- [9] S. Jagannatha, et al., "Comparative Study on Automation Testing using Selenium Testing Framework and QTP," 2014.
- [10] S. Sharma and M. VISHAWJYOTI, "STUDY AND ANALYSIS OF AUTOMATION TESTING TECHNIQUES," Journal of Global Research in Computer Science, vol. 3, pp. 36-43, 2013.
- [11] H. Kaur and G. Gupta, "Comparative Study of Automated Testing Tools: Selenium, Quick Test Professional and Testcomplete," Harpreet Kaur et al Int. Journal of Engineering Research and Applications ISSN, pp. 2248-9622, 2013.
- [12] A. Jain, et al., "A Comparison of RANOREX and QTP Automated Testing Tools and their impact on Software Testing," IJEMS, vol. 1, pp. 8-12, 2014.
- [13] T. Xie, "Improving effectiveness of automated software testing in the absence of specifications," in 2006 22nd IEEE International Conference on Software Maintenance, 2006, pp. 355-359.
- [14] S. Gunasekaran and V. Bargavi, "Survey on Automation Testing Tools for Mobile Applications," International Journal of Advanced Engineering Research and Science (IJAERS), vol. 2, pp. 36-41, 2015.
- [15] L. Gomez, et al., "Reran: Timing-and touch-sensitive record and replay for android," in 2013 35th International Conference on Software Engineering (ICSE), 2013, pp. 72-81.
- [16] R. Mahmood, et al., "Evodroid: Segmented evolutionary testing of android apps", in Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering, 2014, pp. 599-609.
- [17] A. Machiry, et al., "Dynodroid: An input generation system for android apps," in Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering, 2013, pp. 224-234.
- [18] T. Su, "FSMdroid: guided GUI testing of android apps," in Proceedings of the 38th International Conference on Software Engineering Companion, 2016, pp. 689-691.

# Detection of Cardiac Disease using Data Mining Classification Techniques

Abdul Aziz

Faculty of Computer Science & IT  
Superior University, Lahore, Pakistan

Aziz Ur Rehman

Faculty of Computer Science & IT  
Superior University Lahore, Pakistan

**Abstract**—Cardiac Disease (CD) is one of the major causes of death. An important task is to identify the Cardiac disease very minutely and precisely. Generally medical diagnostic errors are dangerous and costly. Worldwide they are leading to deaths. Data mining techniques are very important to minimize the diagnostic errors as well as to improve the patient's safety. Data mining techniques are very effective in designing a medical support system and enrich ability to determine the unseen patterns and associations in clinical data. In this paper, the application of classification technique, decision tree for the detection of heart disease have been introduced. Classification tree uses many factors including age, blood sugar and blood pressure; it can detect the probability of patients fallen in CD by using fewer diagnostic tests which save time and money.

**Keywords**—Cardiac disease; classification technique; decision tree; knowledge discovery

## I. INTRODUCTION

In rapidly growing world, as the time moves individuals need to carry on with a very deluxe life, consequently they work like a gadget with a specific end goal to get a lot of cash and carry on with a casual life. But they overlook to take care of themselves. Their whole way of life is changing as their foods are changing. In this sort of life style they get tension and have blood pressure and sugar issues. It moves towards a major threat, namely, heart disease, an utmost vital organ having a deep influence in all body parts of an individual. Heart is the only organ in human body that really works hard [1]. Cardiac disease is the major cause of mortality in the worldwide over few decades [2]. Various factors exists which are supportive in detection of coronary illness, for example hypertension, smoking, high cholesterol, family history, obesity, blood sugar etc. [3].

In most cases, identification of disease is usually done by the medical specialist abilities and on the basis of current test result. To diagnose an illness is so crucial task that needs high skills and much experience [4]. The main focus is in detection of cardiac disease by using data mining techniques. The enthusiasm to my study is the approximation provided by W.H.O. According to W.H.O by year 2030 just about 23.6 million individuals will kick the bucket due to coronary illness [5].

Thus to reduce this risk the detection of coronary illness should be performed. Current year's medical sector is producing large amount of data related to diagnosis disease, patients, hospital resource and medical devices, etc. [6]. This

data is the main source for effective analysis of data and from this, extract key information that motivates the healthcare community. The detection of heart disease using Data Mining techniques provides us better result. To extract and discover unseen patterns related to heart disease form the existing coronary illness database classification tree which plays an important role. Data Mining focuses machine learning, statistical analysis and databank technology [7]. It assists the medical practitioner and analyst to mark intelligent medical decision which outmoded support system cannot. Some risky elements for CD are: obesity, family history, level of cholesterol, inactiveness and hypertension [8].

According to a survey about 50 percent of victims have no indications till heart attack arises. Analysis of many factors are done to investigate the heart disease, generally physicians make conclusion by assessing current result of the patients tests but it depends upon doctor experience and abilities.

## II. LITERATURE REVIEW

In this section, current literature to diagnose the heart disease by using various data mining methods and tools have been discussed. Among of them a few researches that supported my work are discussed here. M Kumari diagnoses the heart disease by applying a data mining classifier that is Decision Tree. The research scholar analyzes the presentation of this algorithm on various factors that is accuracy rate, sensitivity and error rate. He concludes that the accuracy of Decision Tree is 79 percent [9].

Research demonstrates that by using data mining techniques in health industry then this industry would be in healthier position to fulfill their long term as well as short term goals [10]. By using biomedical mining algorithms heart disease is predicted, the author used classification technique that constructed on supervised machine learning procedures. The author use the decision tree that has error rate 0.2775 and having accuracy of 79.05% [11].

Decision tree algorithms have been applied for classification in various application areas that is production, medicine manufacturing and monetary analysis [12]. N. Kausar and S Planiappan perform a comparison between decision tree and naïve Bayes algorithms. They used UCI data set for risk prediction and stated that decision tree gave high accuracy then naïve Bayes that is 96.4% accuracy [13]. Dr. P Alli, Jenzi and Paryanka offered a new system that depends on mining algorithms to detect the cardiovascular disease. They gathered various patterns to estimate the CD. It happened to them that

decision tree was very cool to fathom and had a better accuracy rate for detecting heart disease [14]. Meenu and Kawaljeet show in his research that Bhatla applied three classifiers such as Decision Tree, Naïve Bayes and neural network for likelihood of CD. Their examination shows that neural network have extraordinary correctness in neural network and after it Decision Tree outer performed over other data mining algorithms [15].

K. Kaurand Lalit shows in their research that he performs many experiments with KNN, Naïve Bayes and Decision Tree. Among all over them DT (Decision Tree) have very high performance of accuracy. Afterwards pre-processing the data correctness of Naïve Bayes and Decision Tree have been enhanced, they use Tanagra tool to classify their data [16]. The study was carried out via C4.5, Decision Tree for identification of stroke disease. The uppermost percentage of this algorithm was 75%, 65% and 75% for Myocardial Infarction disease [17]. Most of the above work was done with WEKA tool.

### III. RESEARCH METHODOLOGY

#### A. Decision Tree

Decision Tree is pondering the most famous technique for diagnosis the cardiac disease. To build a decision tree by using accessible data which can pact with the glitches related to different research areas is very important [18]. Corresponding to the flow chart in which each non leaf node shows a test on a specific attribute and each branch shows the result of that test and each leaf node need a class tag. Root node is the upper most node of the decision tree [19].

The utmost usage of decision tree is in processes research analysis for computing conditional probability. Few advantages of Decision Tree are easily understandable, perform well in huge dataset, simply interpret and robust as well as it knobs both categorical and numerical data [20]. In this the structures that convey supreme information are carefully chosen for classification while other features are put off, by this means computational efficacy is enhanced [21].

#### B. Data Source

These experiments are being carried out for the detection of heart disease using Decision Tree algorithm. The data set is taken form University of California Irvin (UCI) Cleveland Data set and there are total 52 instances from which only 8 attributes are taken for experiments work such as age, chest pain, blood pressure, blood sugar that achieved, angina electro cardiogram. SPSS Clementine 12.0 has been used for [22] calculation and analysis of Data due to its efficiency in finding patterns, analysis and having ability of good prediction.

#### C. Data Set

Selection of data sets is very important because all the experiments and results are based on the data sets. It has been tried to choose the latest, accurate and clean data set so that best results could be obtained. An extra care has been taken in this regards. A total of 210 instances are taken from patient database of Cheema Heart Complex and Cleveland hospital database.

TABLE. I. DATASET AND ITS DESCRIPTION

Attribute	Description
Age	30 to 50 = 1 , 51 to above = 2
Chest pain	1: typical angina 2: atypical angina 3: non-anginal pain 4: asymptomatic
Blood Pressure	Normal (80 to 120) =1 High (above 120) =2
Blood Sugar	False = 0, True = 1
ECG	0: normal 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV) 2: showing probable or definite left ventricular hypertrophy
Heart Rate	Normal (60 to 100) =1 High (above 100) = 2
Angina	Yes = 1, No = 0
<b>Class Attribute</b>	
Disease	0: <50% diameter narrowing (No Disease) 1: > 50% diameter narrowing (Heart Disease)

From 210 only 8 attributes are selected for experiments. The data set contained 117 patients without heart disease and 92 patients with heart disease. We identify a diagnosis class having value 0 with no heart disease and value 1 with heart disease. Table 1 shows the selected attributes and their description.

#### D. Proposed Model

A new model is planned which gives finest result and perfections over previous models. In this section full framework has been discussed as shown in Fig. 1.

The first step in model is the selection of data that is the data source. After sourcing field option is used and a type field is selected that allows field metadata to be determined and controlled. And then the modeling phase occurs, the algorithm C5.0 is selected to constructs a predictive Decision Tree or rule set depends on your own choice and nature of data. After executing the predicted model, performance analysis is performed as shown in Fig. 2. Here performance of algorithm can be evaluated.

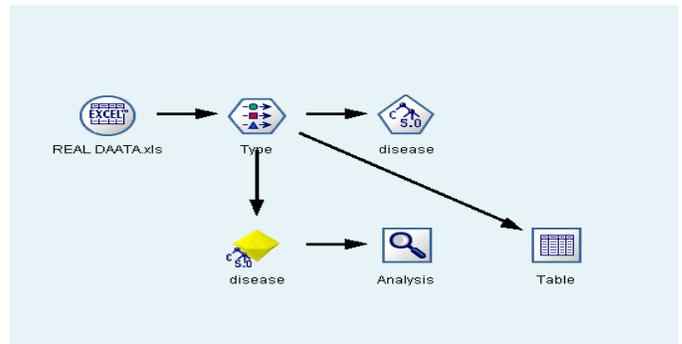


Fig. 1. Classification model.

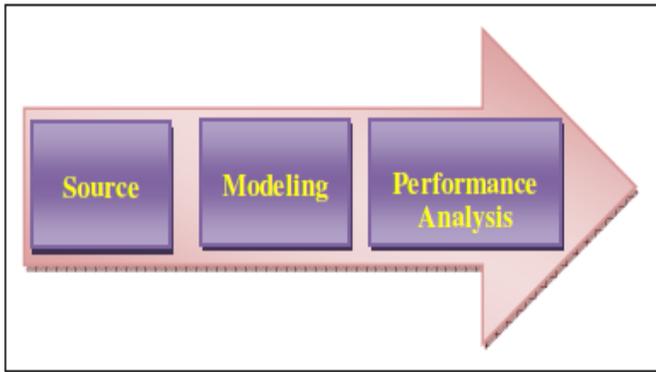


Fig. 2. Block diagram of proposed model.

E. Experiments and Results

Experiments were made using SPSS Clementine tool. Data Set of 210 Patients with 8 medical attributes is used. All attributes are in discrete form and resolved the discrepancies among them. Decision Tree performs best having a good estimation probability of 79.9% by using 8 attributes. That is shown in Fig. 3. There are total 8 attributes from which 7 are numeric and the last taking two values 0 and 1 (0 mean negative and 1 mean positive disease) is my class attribute.

The tree diagram shows all the results of predicted disease. Fig. 4 illustrates that a class attribute (disease) has two child nodes (exercise angina and chest pain) and maximum 2 depth tree. In class attribute Disease (1 presence of heart disease & 0 shows the absence of heart disease), initially 92 persons are found to be infected by heart disease (out of a dataset of 209 patients records), rest of them took further tests/observations like patients having angina during exercise.

For the sake of this, add exercise angina attribute which have two nodes with value 1 and 0 (presence or absence). In node 1, there is probability of 83% having the heart disease while rest 17% are not the victims of heart disease. In node 2 out of 137 people, 23% individuals i.e. 32 people would be infected by heart disease based on the test of chest pain, those who have chest pain type (typical angina and asymptomatic angina: typical angina and non-anginal pain) are 52 percent and 8 percent respectively. The remaining 105 people are not infected by the heart disease (see Fig. 4 and 5).

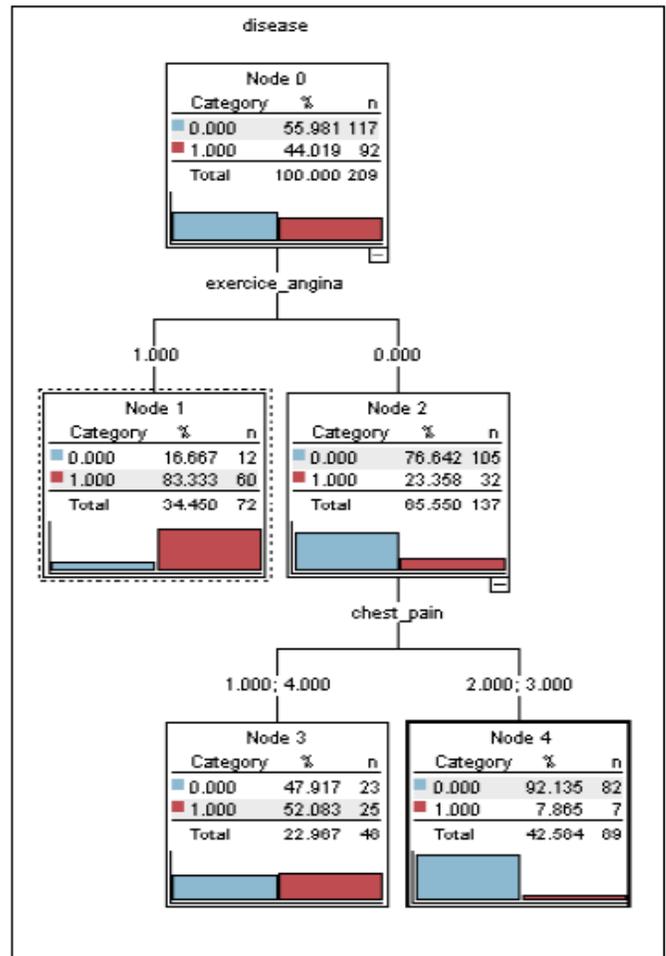


Fig. 4. Model of decision tree.

From the clinical point of view it is a common practice to carry out all the tests whenever a patient attends the clinic with chest pain. Usually, it takes much time to reach at the conclusion whether the patient has a heart disease or simply he suffers from muscular pain. Moreover, in addition to long decision time it is also very costly to patients. With the help of classification tree, numbers of diagnostic tests are reduced which also helps to reduce cost significantly.

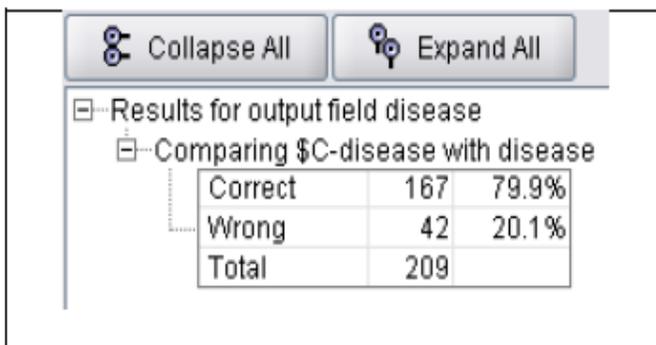


Fig. 3. Results and accuracy.

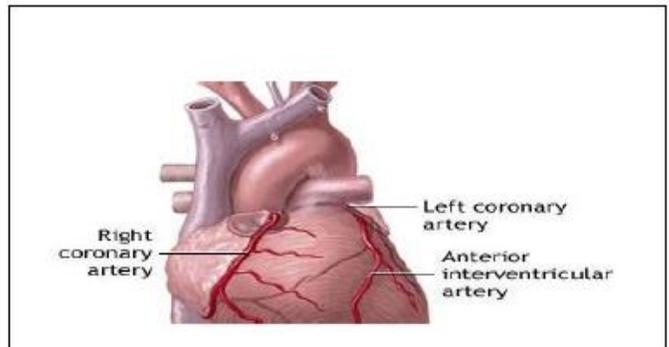


Fig. 5. Heart disease that leads to stroke.

#### IV. CONCLUSION

The most widely used technique of Data Mining in healthcare sector is the classification. The extensive classification method used for the prediction of heart disease is the decision tree that is used in this research. Sometimes poor observations lead towards death. All practitioners are not so expert to diagnose the heart disease with minimal number of tests. The main purpose of this research is to diagnose the heart patients more precisely and more accurately with minimum number of tests (reduction of attributes). This research plays a vital role in the cost reduction of treatment, diagnose disease and additional enhancement of the medical studies. The purposed research work can further be boosted and expended for the prediction of various types of heart diseases.

#### ACKNOWLEDGMENT

The authors thank and acknowledge the support and assistance of Cheema Heart Complex (CHC), Gujranwala, Pakistan. In addition to providing the useful cardiac data they also provided the technical guidance for the success of this research project. We also express gratitude for the commitment, expertise, and actions of members of the CHC. The findings and conclusions in this research paper are those of the authors and do not necessarily represent the official position of the CHC for Disease Control and Prevention.

#### REFERENCES

- [1] Kuldip Singh, Singh G. Alterations in Some Oxidative Stress Markers in Diabetic Nephropathy. *Journal of Cardiovascular Disease Research*. 2017; 8(1):24-7.
- [2] Nagre SW. Mobile Left Atrial Mass – Clot or Left Atrial Myxoma. *Journal of Cardiovascular Disease Research*. 2017; 8(1):31-4.
- [3] Krishnamurthy VT, Venkatesh SA. Negative Pressure Pulmonary Oedema after Sedation in a Patient Undergoing Pacemaker Implantation. *Journal of Cardiovascular Disease Research*. 2017; 8(1):28-30.
- [4] Paryad E, Balasi LR, Kazemnejad E, Booraki S. Predictors of Illness Perception in Patients Undergoing Coronary Artery Bypass Surgery. *Journal of Cardiovascular Disease Research*. 2017; 8(1):16-8.
- [5] Berenson GS, Srinivasan SR, Hunter SM, Nicklas TA, Freedman DS, et al. (1989) Risk factors in early life as predictors of adult heart disease: The Bogalusa heart study. *Am J Med Sci* 298:141-151.
- [6] Dhuper S, Buddhe S, Patel S (2013) Managing cardiovascular risk in overweight children and adolescents. *PaediatrDrugs* 15:181-190.
- [7] Berenson GS, Srinivasan SR, Bao W, Newman WP III, Tracy RE, et al. (1998) Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The Bogalusa Heart Study. *N Engl J Med* 338:1650-1656.
- [8] Newman WP, Freedman DS, Voors AW, Gard PD, Srinivasan SR, et al. (1986) Relation of serum lipoprotein levels and systolic blood pressure to early atherosclerosis. The Bogalusa Heart Study. *New Eng J Med* 314:138-144.
- [9] National High Blood Pressure Education Program Working Group on High Blood Pressure in Children and Adolescents (2004). The fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents; *Paediatrics* 114:555-76.
- [10] Dinarević S, Mulaosmanović V (2005) Primary prevention of Hypertension in Sarajevo Children: Role of Adiposity. 29th UMEMPS Congress Union of Middle Eastern and Mediterranean Paediatric Societies, pp. 154-156.
- [11] Dinarević S, Mesihović H, Simeunović S, Zulić I (1994) Dyslipoproteinaemia in Children with Heart Disease. *Intercontinental Cardiol* 3:126-9.
- [12] Kapur G, Ahmed M, Pan C, Mitsnefes M, Chiang M, et al. (2010) Secondary hypertension in overweight and stage 1 hypertensive children: A Midwest Paediatric Nephrology Consortium report. *J Clin Hypertens* 12:34-39.
- [13] Dinarević S, Hasanbegović S (2010) Problem of obesity in children and youth in Canton Sarajevo. *Pediatr Res* 68:1091.
- [14] WHO Europe (2007) The challenge of obesity in the WHO European Region and the strategies for response. Copenhagen: WHO Regional Office for Europe.
- [15] Daniels SR, Arnett DK, Eckel RH, Gidding SS, Hayman LL, et al. (2005) Overweight in children and adolescents: Pathophysiology, consequences, prevention and treatment. *Circulation* 111:1999-2012.
- [16] Berenson GS, Wittingly WA, Tracy RE, Newman WP, Srinivasan SR, et al. (1992) Atherosclerosis of the aorta and coronary arteries and cardiovascular risk factors in persons aged 6 to 30 years and studied at necropsy (The Bogalusa Heart Study). *Am J Cardiol* 70:851-858.
- [17] McNiece KL, Gupta-Malhotra M, Samuels J, Bell C, Garcia K, et al. (2007) National High Blood Pressure Education Program Working Group: Left ventricular hypertrophy in hypertensive adolescents: Analysis of risk by 2004 National High Blood Pressure Education Program Working Group staging criteria. *Hypertension* 50:392-5.
- [18] Torrance B, McGuire KA, Lewanczuk R, McGavock J (2007) Overweight, physical activity and high blood pressure in children: a review of the literature. *Vasc Health Risk Manag* 3:139-149.
- [19] Jiber H, Blitti MC, Bouarhroum A. Acute type B Aortic Dissection Complicated by Acute Limb Ischemia: Case Report. *Journal of Cardiovascular Disease Research*. 2016;7(2):97-9.
- [20] Muhammad Subhi Al- Batah "Testing the probability of Heart Disease using Classification and Regression Tree" *Annual Research & Review in Biology* 4 (11): 1713-1725,(2014).
- [21] Setty HS, Hebbal VP, Channabasappa YM, Jadhav S, Ravindranath KS, Patil SS, et al. Assessment of RV function following Percutaneous Transvenous Mitral Commissurotomy (PTMC) for rheumatic mitral stenosis. *Journal of Cardiovascular Disease Research*. 2016; 7(2):58-63.
- [22] Patnaik L, Pattanaik S, Sahu T, Panda BK. Awareness of symptoms and risk factors of Myocardial Infarction among adults seeking health care from a rural hospital of India. *Journal of Cardiovascular Disease Research*. 2016;7(2):83-5.

# Image Encryption Technique based on the Entropy Value of a Random Block

Dr. Mohammed A. F. Al-Husainy  
Department of Computer Science  
Faculty of Information Technology,  
Middle East University  
Amman, Jordan

Dr. Diaa Mohammed Uliyan  
Department of Computer Science  
Faculty of Information Technology,  
Middle East University  
Amman, Jordan

**Abstract**—The use of digital images in most fields of information technology systems makes these images usually contain confidential information. When these images transmitted via the Internet especially in the Cloud, it becomes necessary to protect these images in a way that ensure putting the confidential information that are contained far away from the attackers. A proposed image encryption technique has been presented in this work. This technique used a secret key that is extracted from the image content itself. Therefore, there is no need to find a secret channel to exchange any key where, sender and receiver authenticate each other with regards to a shared secret key extracted from the image. The technique constructs its secret key that is used to encrypt the image, based on the entropy values of a set of randomly selected blocks from the image itself. Vairous experiments have been conducted to evaluate the strength and performance of the technique. The experimental results shows that the proposed technique can be used effectively in the field of image security to protect and authenticate images.

**Keywords**—Image security; image encryption; secret key; image authentication

## I. INTRODUCTION

With the rapid expansion of modern network technology such as Cloud computing, many of current applications such as Facebook employed cloud storage services to store multimedia data (e.g., Images and videos). Due to the fact, that images may contain private information that may be related to personal interests or financial affairs, the deliberate disclosure of confidential content becomes a critical issue for people and organizations [1]-[4].

Images need to be accessed and shared over the cloud securely. Image encryption is an efficient mechanism to contribute security for these images. Encrypting image is defined as protecting visual image through the Internet from hacker attacks [5]. In recent years, various image encryption algorithms have been proposed using cryptographic techniques by modifying their pixel values or locations [6]-[8]. Cryptographic techniques are categorized into symmetric and asymmetric encryption. Symmetric key encryption algorithm [9] uses one key for encrypting and decrypting image respectively. Obviously, it requires keeping the key secret. If the hacker knows the key, image decryption can be done easily. In contrast, asymmetric key encryption algorithm [10] employed two keys: public and private key. Recent studies mentioned that asymmetric key encryption is slower than

symmetric key encryption algorithms [11]. Furthermore, asymmetric encryption algorithm has higher computational complexity which, are most of the time prohibitive for images and mathematical correlation between public and private keys may help attackers to hack the image [12]. This might be solved by using secret keys for image encryption or longer sized keys which are difficult to violate by attackers.

Due to digital images have intrinsic characteristics such as redundancy of data, less sensitive, a correlation between pixels and data capacity, it is difficult to handle these issues by using asymmetric cryptographic techniques [13]. They are not suitable for image encryption, while symmetric key image cryptographic algorithms appears to be a promising direction which takes profit from these characteristics to encrypt images [14].

The proposed method assumes that image is encrypted at rest with some secret key which, is not available to the attacker. To achieve this issue, the proposed method is developed based on entropy values of the image itself as secret key. The secret key will differ from one image to another. The fact that the attacker may have the historical secret key no longer matters because all the old keys are meaningless. The proposed method has the ability to resolve security problems caused by large data capacity and high correlation among pixels for color image encryption.

The rest of this paper is organized as follows: Related Works will be covered in Section 2. Section 3 describes the Proposed algorithm in detail. Section 4 presents Experimental results and performance analysis. In Section 5, Conclusions are drawn.

## II. RELATED WORKS

The main goal of image encryption techniques is to convert source images into limited formats such as texture based or noise based format encrypted images. The pixel values of encrypted images have been changed to prduce a noise image, which arise the information leakage of image content and hide the visual meaning of these images over the cloud. From the security viewpoint, texture based or noise based format pixel features in the encrypted images would efficiently decrease the risk of an encrypted image being attacked and altered. This interesting issue motivates us to present a novel image encryption technique based on entropy features as secret keys

for transforming image into a nearly uniform distributed pixel values in the image to achieve privacy and confidentiality [15].

In recent years, image encryption techniques have been proposed to provide privacy preserving for digital images stored in cloud storage. Image encryption has become the hot topic of exhaustive research as its potential to transmit images more securely. Image encryption techniques can be categorized into: 1) Frequency based image encryption and 2) Spatial based image encryption. Hence, using secret key, the frequency based image encryption algorithms are developed to transform image content in the frequency domain such as the Discrete Cosine Transform (DCT) [16], Discrete fractional Fourier Transform (DFFFT) [17], [18], Quantum Fourier Transform (QFT) [19], Fresnel transform [20], Hartley Transform [21] and Gyrator Transform [22].

The spatial based image encryption techniques are based on two common operations: Substitution and Permutation, where substitution is used to change pixel values and a permutation process is used to shuffle pixel positions in the image. The permutation and substitution processes can be used in spatial based image encryption algorithms like Data Encryption Standard (DES) [15], Advanced Encryption Standard (AES) [23], Rivest, Shamir and Adleman (RSA) [24], P-Fibonacci transform [25], wave transmission [26], elliptic curve ElGamal [27], gray code [28], random grids [29], Latin squares [30] and chaotic mapping [31].

In the first category, the digital image is divided into blocks and transformed into the frequency domain to extract features. These features are disordered to make the original image is invisible. For instance, Phalavan et al. [16] proposed an image encryption method based on DCT coefficients, where the image is divided into  $8 \times 8$  blocks and then extract high frequency DCT coefficients from these blocks. The main advantage of high frequencies DCT is representing more details in the image content. The secret key for their technique is generated based on cellular automata. Finally, the image blocks were encrypted using a secret key with XOR operation to disorder the values in each block.

Guo et al. [17] used Double Random Phased Encryption technique (DRPE) to encrypt the image where, binary image is used as secret key. Their method is robust of noise addition and the errors in the secret keys, which employed through decryption steps. Similarly, Lima et al. [18] divided grayscale image into blocks with size  $8 \times 8$  pixels. The image blocks were encrypted with a secret key based on Galois Field Fractional Fourier Transform (GFFFT). The size of secret key is 140-bit, which makes their method has a large key space resist a brute force attack.

In [19], quantum gray level image representation and image encryption is proposed based on QFT. The correlation between adjacent pixels in the encrypted image and its original image is computed. Their method gives high level of security, where random relation exists in the encrypted image.

Singh et al. [20] proposed a scheme that transform an image into complex image subjected to Fresnel transform to extract frequency coefficients. The Devil's vortex toroidal lens

phase mask is applied in the frequency domain to produce an encrypted image. The mean square error is computed for their method to show the robustness of encryption algorithm against Gaussian and speckle noise.

Lin et al. [21] proposed to use image scrambling in frequency domain based on Hartley transform. The input image is converted into Arnold cat map space plane and then, it is divided the image into  $3 \times 3$  blocks. Later, blocks are encoded by H matrix of Heatley transform.

Liu et al. [22] proposed to use an iterative image encryption structure, in which Henon mapping is applied for the input image. Then, mapped image is transformed through gyrator transform to encode the image.

In the second category, Yun-Peng et al. [15] proposed a conventional encryption technique by combining DES algorithm with chaotic sequence to encode the image. The key size for encryption method is 264 bits, which is much larger than the traditional DES algorithm to resist against brute force attacks.

In [23], an image encryption algorithm with a framework of combining diffusion and permutation is proposed. The input image is divided into blocks with size  $8 \times 8$  pixels. Each block is encoded based on pseudo random numbers which extracted from spatiotemporal chaos. Finally, the permutation of each block is computed using AES. Their method achieved a high speed by avoiding some time consuming operations.

Zhao et al. [24] used RSA encryption algorithm to scramble grayscale image. This algorithm is limited and not suitable for practical images due to large number of permutation rounds. It may not be the most desirable algorithms for digital image encryption, especially for real time systems.

Zhou et al. [25] introduced a new method for encrypting images by combining bit plane decomposition and image permutation. They used Fibonacci P-code transformation to scramble image. Their method is robust to various common attacks like noise, data loss, brute force and plaintext attacks.

Chen et al. [31] developed an encryption method based on Henon chaotic function and Logistic map for encrypting image and the secret key respectively.

Both the frequency based and spatial based image encryption techniques have the ability to increase the level of security to protect images. They are evaluated based on four factors: Security, Speed, Key space analysis and Correlation.

### III. PROPOSED MODEL

One most concern of the proposed method is how to generate a secret key from the image properties itself. The image is encrypted and shared through some secret means between sender and receiver. So, the attacker struggles to know the secret key unless he has the same image and the algorithm used for the decryption of the image. The strength of the proposed method comes from the key where it is not dependent on algorithm steps being secret. This leads that, it is difficult to obtain the secret key value out of the possible key space. Therefore, a set of main objectives has been established to be

achieved within the proposed technique :

- A. No need to exchange any secret key via a secure channel between sender and receiver.
- B. A secret key used in encryption operation is extracted from the image itself.
- C. The size of the secret key varied based on the nature of the image.
- D. Trying to use a secret key that contains as more as possible of random values.
- E. Apply substitution and transposition operations on the image within two different levels of implementation (on block of bytes and on a single byte).

The main stages that the proposed image encryption technique involves are depicted in Fig. 1.

To give readers a clear understanding of the implementation details of each stage of the proposed technique, we give below some necessary definitions and terminologies:

1) **Source Image (I)** is a bitmap color image of size (Width×Height×Palette). Where: Width is the width of the image, Height is the height of the image and Palette equal 3 which represents the three colors (R: Red, G: Green and B: Blue).

2) **Encrypted Image (E)** is a bitmap color image of size (Width×Height×Palette). Where: Width is the width of the image, Height is the height of the image and Palette equal 3 which represents the three colors (R: Red, G: Green and B: Blue). The image (E) is produced from the technique after encrypting the source image I.

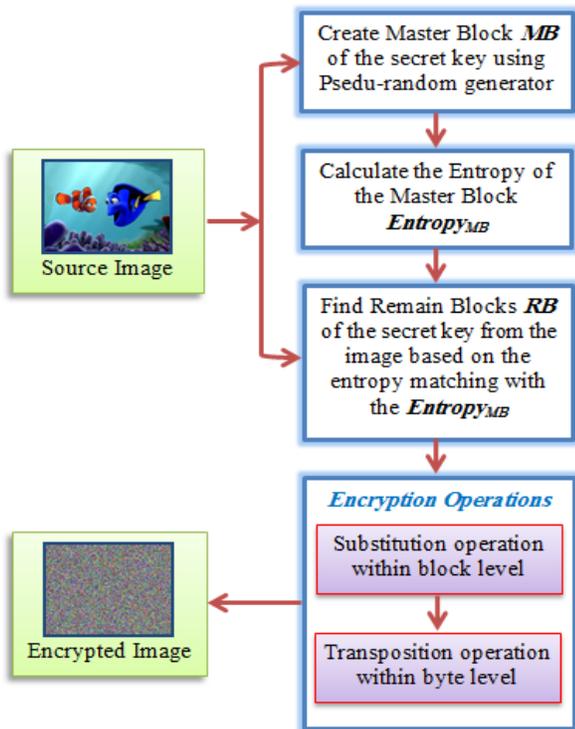


Fig. 1. Stages of the proposed image encryption technique.

3) **Secret Key (K)** Consists of a set of 2D blocks, it is divided into two parts: Master Block MB and the Remaining Blocks RB. All blocks are of the same size. The size of each block (i.e., number of rows and columns) is based on the image size and is calculated using the equations (1) and (2).

$$\text{NumberOfRows (R)} = \text{MaxDigitOf(Height)} \tag{1}$$

$$\text{NumberOfColumns (C)} = \text{MaxDigitOf(Width)} \tag{2}$$

4) **Entropy** is simply the average (expected) amount of the information from the data [32]. Information entropy is an important feature of randomness. Here, the entropy value is calculated by the equation (3).

$$\text{Entropy} = -\sum_{i=1}^n p_i \log_2(p_i) \tag{3}$$

Where  $n$  = number of different data values,  $p_i$  is probability of occurring the data value  $i$ .

The main stages that follow in the encryption phase of the proposed image encryption technique are given below:

**Stage 1:** From the source image size (i.e., its Width and Height), calculate the dimensions of the secret key blocks (R and C) using the equations (1) and (2).

**Stage 2:** Build the Master Block MB of the secret key K by using a pseudo random generator with a seed value (R×C). Random values of bytes are filled in MB. The value of each byte is between (0 ... 255). This part of the secret key K (i.e., MB) will be used later to construct the second part of the secret key K (i.e., RB). An example of the MB is shown in Fig. 2.

**Stage 3:** Calculate the entropy value of the MB Entropy<sub>MB</sub> using the equation (3).

**Stage 4:** Represent the bytes of the source image I bytes as a set of blocks of dimensions (R×C). The number of blocks in the source image I is calculated using the equation (4).

$$\text{NoOfBlocks} = (\text{Width} \times \text{Height} \times \text{Palette}) / (\text{R} \times \text{C}) \tag{4}$$

**Stage 5:** Search in the source image blocks that are created in Stage 4 to find all the blocks that have entropy values equal to the entropy value Entropy<sub>MB</sub>. These blocks are represented the remaining blocks RB of the secret key as shown in Fig.3. The number of remaining blocks RB found depends mainly on the image data. And these blocks are excluded from the implementation of the encryption operations that are applied in the next stages. The second part of the secret key K becomes ready to be used in the implementation of the encryption operations on the source image blocks.

C	0	1	2	3
R				
0	14	206	120	37
1	211	30	75	149
2	20	0	55	234

Fig. 2. Content of master block.

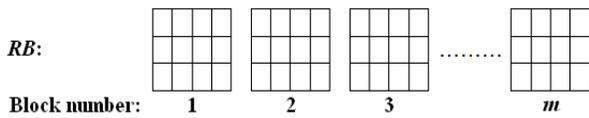


Fig. 3. Remaining blocks in the image.

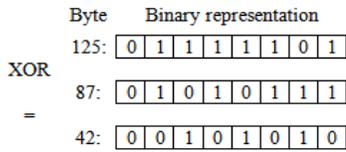


Fig. 4. XOR operation between blocks.

**Stage 6:** Perform the substitution operation of the encryption technique by doing the XOR logic operation between each block of the image with the one of the blocks in the RB set of blocks sequentially. When reach the block index *m* in the RB set of blocks, the next block index is 1. Example of XOR operation between two byte values is shown in Fig. 4 above.

**Stage 7:** Perform the transposition operation of the encryption technique by doing the randomly change the locations of all the bytes of the image including the bytes of the RM blocks. By using a pseudo random generator with a seed value (R×C), a new random location for each byte in the image be found.

**Stage 8:** Construct the encrypted image E from the collection of bytes that are produced in the Stage 7. As a result, the encrypted image is produced.

The same scenario has been implemented in the decryption phase to recover the original image from the encrypted one E, except the receiver should perform the transposition operation first.

In the proposed image encryption technique, we can note that there is no need to exchange any key between sender and receiver. The key size used (i.e., number of RM blocks used) and the values of bytes in each block vary based on the image data itself. This makes the technique easy to use by the users and more difficult against the attackers. Furthermore, the proposed method leads to the uniformization of the histogram of the encrypted images, which makes it more secure against statistical hacks as shown in Section 4.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

In order to evaluate the performance of the proposed technique, the necessary programming codes using C# language are written. Different images used in the experiments to test the technique. Some of these images are listed in Fig. 5.



Fig. 5. Some of the images use in the experiments.

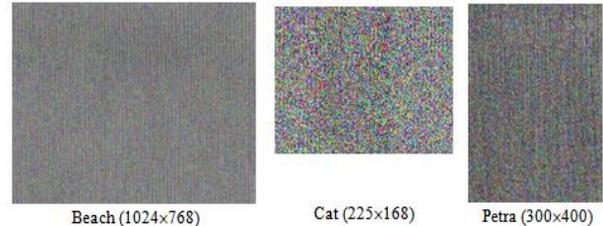


Fig. 6. The encrypted image from the images in Fig. 5.

The encrypted images that are produced from the source images in figure 1 are depicted in Fig. 6.

To make a performance comparison between the proposed image encryption technique and the well-known encryption techniques such as Data Encryption Standard (DES) [15] and Advanced Encryption Standard (AES) [23]. A set of measures (visually and numerically) has been used: Image histogram, Peak Signal to Noise Ratio (PSNR), encryption time, the key size, the complexity of the key and the sensitivity of the key.

1) **Image histogram:** good image encryption technique is the one that is able to achieve a high distortion in the distribution of color values of the encrypted image compared with the distribution of color values of the source image. Fig. 7 shows the histogram of the source image (Beach) and its encrypted image using the proposed technique and the DES and AES techniques.

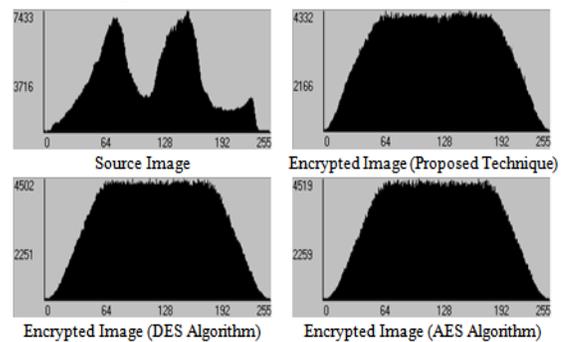


Fig. 7. Histogram of the source image and its encrypted images (Beach image in Fig. 5).

2) **Peak Signal to Noise Ratio (PSNR):** a numerical measure that is used to calculate the ratio of noise that is occurring in the encrypted image and it is caused by the implementation of the encryption technique on the source image. Good image encryption technique is the one that is producing high ratio of noise in the encrypted image with a low PSNR value [33], [34]. The PSNR is calculated using the equation (5) and (6), where MAX represents the maximum byte value in the image and I and E are the source image and the encrypted image respectively. Table 1 records the PSNR of the encrypted image in the proposed technique and DES and AES techniques.

$$NMAE = \frac{\sum_{k=0}^{(Width \times Height \times Palette) - 1} |I(k) - E(k)|}{Width \times Height \times Palette} \times 100 \quad (5)$$

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{NMAE} \right) \quad (6)$$

3) **Encryption time:** the time needed to complete the encryption phase is one of the common factors that is used to evaluate the performance of the encryption technique. A good image encryption technique is the one that is conducting its encryption operations in a short time. Table 2 summarizes the encryption time of the proposed technique and DES and AES techniques.

TABLE. I. PSNR OF THE ENCRYPTED IMAGE

Image	Algorithm	PSNR (db)
Beach	Proposed	8.42
	DES	6.56
	AES	6.56
Cat	Proposed	8.25
	DES	7.38
	AES	7.37
Petra	Proposed	8.71
	DES	7.81
	AES	7.82

TABLE. II. ENCRYPTION TIME

Image	Algorithm	Time (sec)
Beach	Proposed	3.15
	DES	3.70
	AES	3.37
Cat	Proposed	0.17
	DES	0.16
	AES	0.18
Petra	Proposed	0.52
	DES	0.59
	AES	0.53

4) **Key size:** the encryption technique becomes strong whenever it can be used a proportionally large key. The key used in the proposed technique is actually large because it consists of a set of blocks (i.e., RB) and each block consists of a number of bytes. This makes the bits that represent the secret key used by the proposed technique is large. To illustrate this, we assume that the block dimension is (R=4 and C=4). This means that each block of the RB has 16 bytes. And if there are 5 blocks in RB, this means that the number of bits that represent the key is 640, where it calculated using the equation (7). While the DES algorithm uses a key contains 64 bits and the AES algorithm uses a key contains 128 to 256 bits.

$$\text{NumberOfBits} = \sum_{i=1}^m [(R \times C) \times 8] \quad (7)$$

5) **Complexity of key:** is the ratio of randomness and the composite use of the key in the implementation of the operations of the encryption technique. The proposed technique uses a key that is extracted from the image itself and it contains really random bytes. In addition, it uses key in two levels of implementation of the operations (block and byte level).

6) **Sensitivity of key:** it means that if we change only one bit in the key used, the technique produces, from the encrypted image, an image that is a completely different from the source image. This forces the attacker to know all the bits of the key to be able to recover the source image from the encrypted image. To prove that, in the proposed technique, we changed one bit in the key used to encrypt one of the images in Fig. 5 and try to decrypt the encrypted image. The produced image is completely different from the source image as shown in Fig. 8.

7) **Correlation Analysis:** is the value that depicts the relationship between the adjacent pixels values in the encrypted image. Whenever the correlation value is small this means that the encryption technique achieved high randomness between the adjacent pixels in the encrypted image. The calculated correlation values for the proposed technique and DES and AES techniques are listed in Table 3.

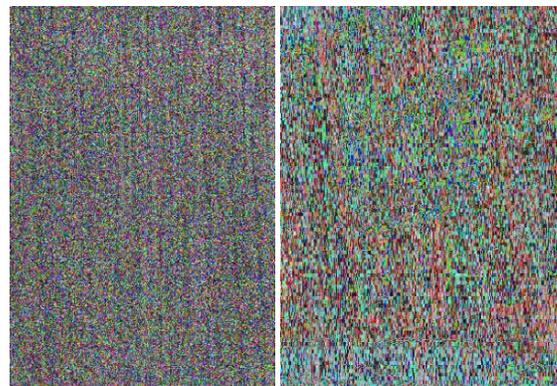


Fig. 8. The source image resulting from the encrypted image using the wrong key (Petra image in Fig. 5).

TABLE. III. CORRELATION BETWEEN ADJACENT  
PIXELS VALUES

Image	Algorithm	Correlation Value
Beach	Proposed	0.123
	DES	0.107
	AES	0.106
Cat	Proposed	0.120
	DES	0.089
	AES	0.089
Petra	Proposed	0.090
	DES	0.094
	AES	0.093

### V. CONCLUSION

This paper introduced a novel image encryption scheme based on Entropy values of selected blocks and perform XOR permutation at the block level with the transposition operation at the byte level. The Entropy is computed from the selected block in the image whose size is based on the image size. The rest of blocks are selected which they have the same entropy to be as candidates used for generating large enough secret key space to encrypt the image later. High level of security is achieved by using a random secret key. This leads that different image should have a different random key for encryption. Experimental results show that the scrambled image has approximately uniform histogram pattern and can be considered as a nearly random image. The security analyses also demonstrate that the proposed method is sensitive to the nature of source image and the encryption key. Therefore, the proposed method has high security and can resist against most common attacks. However, we have found from the experimental results that the secret key values may be different from one image to another which adds more ambiguity at the side of attackers about the key itself. The future work is to investigate the common malicious attacks that can be applied to images such as copy move forgery and image splicing in the field of image forensics.

### ACKNOWLEDGMENT

The authors are grateful to the Middle East University, Amman, Jordan for the financial support granted to cover the publication fee of this research article.

### REFERENCES

[1] Y.-Q. Zhang and X.-Y. Wang, "A symmetric image encryption algorithm based on mixed linear–nonlinear coupled map lattice," *Information Sciences*, vol. 273, pp. 329-351, 2014.

[2] D. M. Uliyan, H. A. Jalab, and A. W. A. Wahab, "Copy move image forgery detection using Hessian and center symmetric local binary pattern," in *Open Systems (ICOS)*, 2015 IEEE Confernece on, 2015, pp. 7-11.

[3] D. M. Uliyan, H. A. Jalab, A. W. Abdul Wahab, and S. Sadeghi, "Image Region Duplication Forgery Detection Based on Angular Radial Partitioning and Harris Key-Points," *Symmetry*, vol. 8, p. 62, 2016.

[4] D. M. Uliyan, H. A. Jalab, A. W. A. Wahab, P. Shivakumara, and S. Sadeghi, "A novel forged blurred region detection system for image forensic applications," *Expert Systems with Applications*, vol. 64, pp. 1-10, 2016.

[5] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 2594-2608, 2016.

[6] K. Loukhaoukha, J.-Y. Chouinard, and A. Berdai, "A secure image encryption algorithm based on Rubik's cube principle," *Journal of Electrical and Computer Engineering*, vol. 2012, p. 7, 2012.

[7] S. Liu, C. Guo, and J. T. Sheridan, "A review of optical image encryption techniques," *Optics & Laser Technology*, vol. 57, pp. 327-342, 2014.

[8] Z. Hua, Y. Zhou, C.-M. Pun, and C. P. Chen, "2D Sine Logistic modulation map for image encryption," *Information Sciences*, vol. 297, pp. 80-94, 2015.

[9] S. Sowmya and S. Sathyanarayana, "Symmetric Key Image Encryption Scheme with Key Sequences Derived from Random Sequence of Cyclic Elliptic Curve Points over GF (p)," in *Contemporary Computing and Informatics (IC3I)*, 2014 International Conference on, 2014, pp. 1345-1350.

[10] W. Liu, Z. Xie, Z. Liu, Y. Zhang, and S. Liu, "Multiple-image encryption based on optical asymmetric key cryptosystem," *Optics Communications*, vol. 335, pp. 205-211, 2015.

[11] J. J. Amador and R. W. Green, "Symmetric-key block cipher for image and text cryptography," *International Journal of Imaging Systems and Technology*, vol. 15, pp. 178-188, 2005.

[12] H. Rifa-Pous and J. Herrera-Joancomartí, "Computational and energy costs of cryptographic algorithms on handheld devices," *Future internet*, vol. 3, pp. 31-48, 2011.

[13] Y. Wang, K.-W. Wong, X. Liao, and G. Chen, "A new chaos-based fast image encryption algorithm", *Applied soft computing*, vol. 11, pp. 514-522, 2011.

[14] M. Saikia and B. Baruah, "Chaotic Map Based Image Encryption in Spatial Domain: A Brief Survey," in *Proceedings of the First International Conference on Intelligent Computing and Communication*, 2017, pp. 569-579.

[15] Z. Yun-Peng, L. Wei, C. Shui-ping, Z. Zheng-jun, N. Xuan, and D. Weidi, "Digital image encryption algorithm based on chaos and improved DES," in *Systems, Man and Cybernetics*, 2009. SMC 2009. IEEE International Conference on, 2009, pp. 474-479.

[16] A. Pahlavan Tafti and S. Janosepah, "Digital images encryption in frequency domain based on DCT and one dimensional cellular automata," *Informatics Engineering and Information Science*, pp. 421-427, 2011.

[17] C. Guo, S. Liu, and J. T. Sheridan, "Optical double image encryption employing a pseudo image technique in the Fourier domain," *Optics Communications*, vol. 321, pp. 61-72, 2014.

[18] J. B. Lima and L. Novaes, "Image encryption based on the fractional Fourier transform over finite fields," *Signal Processing*, vol. 94, pp. 521-530, 2014.

[19] Y.-G. Yang, X. Jia, S.-J. Sun, and Q.-X. Pan, "Quantum cryptographic algorithm for color images using quantum Fourier transform and double random-phase encoding," *Information Sciences*, vol. 277, pp. 445-457, 2014.

[20] H. Singh, A. Yadav, S. Vashisth, and K. Singh, "Optical image encryption using devil's vortex toroidal lens in the Fresnel transform domain," *International Journal of Optics*, vol. 2015, 2015.

[21] K. T. Lin, "Image Encryption Using Arnold Transform Technique and Hartley Transform Domain," in *Intelligent Information Hiding and Multimedia Signal Processing*, 2013 Ninth International Conference on, 2013, pp. 84-87.

[22] Z. Liu, L. Xu, C. Lin, J. Dai, and S. Liu, "Image encryption scheme by using iterative random phase encoding in gyrotator transform domains," *Optics and Lasers in Engineering*, vol. 49, pp. 542-546, 2011.

[23] S. H. Kamali, R. Shakerian, M. Hedayati, and M. Rahmani, "A new modified version of advanced encryption standard based algorithm for image encryption," in *Electronics and Information Engineering (ICEIE)*, 2010 International Conference On, 2010, pp. V1-141-V1-145.

[24] G. Zhao, X. Yang, B. Zhou, and W. Wei, "RSA-based digital image encryption algorithm in wireless sensor networks," in *Signal Processing*

- Systems (ICSPS), 2010 2nd International Conference on, 2010, pp. V2-640-V2-643.
- [25] Y. Zhou, K. Panetta, S. Aghaian, and C. P. Chen, "Image encryption using P-Fibonacci transform and decomposition", *Optics Communications*, vol. 285, pp. 594-608, 2012.
- [26] B. Norouzi and S. Mirzakuchaki, "A fast color image encryption algorithm based on hyper-chaotic systems," *Nonlinear Dynamics*, vol. 78, pp. 995-1015, 2014.
- [27] L. Li, A. A. A. El-Latif, and X. Niu, "Elliptic curve ElGamal based homomorphic image encryption scheme for sharing secret images," *Signal Processing*, vol. 92, pp. 1069-1078, 2012.
- [28] J.-x. Chen, Z.-l. Zhu, C. Fu, H. Yu, and L.-b. Zhang, "An efficient image encryption scheme using gray code based permutation approach," *Optics and Lasers in Engineering*, vol. 67, pp. 191-204, 2015.
- [29] S. J. Shyu, "Image encryption by random grids," *Pattern Recognition*, vol. 40, pp. 1014-1031, 2007.
- [30] Y. Wu, Y. Zhou, J. P. Noonan, and S. Aghaian, "Design of image cipher using latin squares," *Information Sciences*, vol. 264, pp. 317-339, 2014.
- [31] C.-K. Chen, C.-L. Lin, C.-T. Chiang, and S.-L. Lin, "Personalized information encryption using ECG signals with chaotic functions," *Information Sciences*, vol. 193, pp. 125-140, 2012.
- [32] H. Zhang, J. E. Fritts, and S. A. Goldman, "An entropy-based objective evaluation method for image segmentation," *Storage and Retrieval Methods and Applications for Multimedia*, vol. 5307, pp. 38-49, 2004.
- [33] M. A. F. Al-Husainy, "A novel image encryption algorithm based on the extracted map of overlapping paths from the secret key," *RAIRO-Theoretical Informatics and Applications*, vol. 50, pp. 241-249, 2016.
- [34] M. A. F. Al-Husainy, "Message segmentation to enhance the security of LSB image steganography," *International Journal of Advanced Computer Science and Applications*, vol. 3, pp. 57-62, 2012.

# Enhancing Lean Software Development by using DevOps Practices

Ahmed Bahaa Farid  
Information Systems Dept,  
Faculty of Computers and  
Information, Helwan University,  
Cairo, Egypt

Yehia Mostafa Helmy  
Business Information Systems Dept,  
Faculty of Commerce & Business  
Administration, Helwan University,  
Cairo, Egypt

Mahmoud Mohamed Bahloul  
Business Information Systems Dept,  
Faculty of Commerce & Business  
Administration, Helwan University,  
Cairo, Egypt

**Abstract**—Competition between companies has made a great pressure to produce new features continuously as fast as possible, subsequently successful software companies needs to learn more about customers and get new features out to them more rapidly. Lean software development cannot integrate between development and operation teams. DevOps enables this merge between them and creates operational parts as one part of the development process and made it up to date during the development phase, so reduced errors during the deployment. The purpose of this paper is to investigate how one can use devOps practices to improve the performance of lean software development production process and introduce a new framework that merge lean and devOps process. The research has been evaluated on a sample of 2 departments in Faculty of Commerce at Helwan University. The results of this work have led to reduce the response delivery time for customers and rapid feedback provides accurate expectations for customer needs that lead to lower levels of deployment pains and lower change fail rates.

**Keywords**—Lean software development; DevOps; development & IT operations; continuous delivery; monitoring; continuous integration

## I. INTRODUCTION

Changing business needs always required providing products faster to market due to competition among software companies which puts an increasing pressure to produce new features extremely fast. Projects in software development field always faced risks or problems as bugs, failure, past deadline and poor quality etc. In recent years, software companies need way in which you manipulate problems such as long development life cycles and rapidly changing requirements from customer [1].

Lean methodology is designed to minimize the wastes of resources that do not add any customer value to products [2].

Lean depends mainly on continuous improvement and to achieve this any defect or problem that may occur in the delivery process must be detected to get feedback continuously [3].

DevOps is set of practices and principles that is trying to improve life-cycle as a whole through integration between development and operations teams to reduce the release cycles and increase number of software deliveries [4].

The paper is organized as follows: Section 2 presents the background and related research of Lean Software Development and DevOps. Section 3 describes the research problem. Sections 4 and 5 describe the research goals, approach, and assumptions. Section 6 shows the theoretical mapping of the two approaches. Section 7 shows overall results of the theoretical mapping. Section 8 describes the background of the analyzed study. Section 9 summarizes the assessment results from the study. In Section 10, the results of this study are discussed. The last section concludes the paper with the key findings, research limitations and future work.

## II. OVERVIEW

This section consists of three parts. The first part presents an overview of Lean Software Development. The second part provides an overview of DevOps. The final part gives Related Work.

### A. Lean Software Development

Lean software development provides a set of principles to minimize wastes and maximize the customer value in software processes. Mary and Tom Poppendieck [5] have formulated a set of principles for the application of Lean thinking into software development.

There are seven main principles in Lean development process as the following: Eliminate waste, Amplify Learning, Decide as Late as Possible, Deliver as Fast as Possible, Empower the Team, Build Integrity In and See the Whole [5].

Eliminating waste is the first principle that explains the Waste as any unnecessary activities that add cost or time without adding value to the customers [6, 3]. There are many wastes that transferred by Poppendieck and Poppendieck [5] from manufacturing to software development are: partially completed work, extra Features, extra processes, task switching, Handoffs (Motion), delays (waiting) and defects.

The way lean works is by creating more value for customers with fewer resources through remove wastes from activities and eliminating whenever possible those steps that do not create value to enhance quality products [7].

Take the right time to adding the real value that satisfied the customer through remove anything that doesn't either add customer value directly or add knowledge about how to deliver that value more effectively [8].

## B. DevOps

Market needs are changing continuously. Therefore, there is always a need to adapt continuously to market needs and deliver quickly. DevOps appeared as a result of integration of development and operators team members to increase speed of new software releases and reduce time to respond to customer needs and changes [9].

Dubois [10] found that the developers had no knowledge of what was happening to the application on the deployment infrastructure, and that the operations team also did not care of the planning and priorities of projects.

Each of these teams has a different goal, development teams are interested for deliver new features and operations teams are interested for stability [11].

DevOps is approach that emerging to bridge this gap between these two teams and to achieve collaboration between them. The deployment process needs to be highly automated to enable continuous delivery of software so the DevOps provides a huge variety of practices to implement holistic deployment automation [12].

DevOps will improve productivity through accelerated customer feedback cycles and reduced overhead and rework in addition to provide a competitive advantage to a business through three dynamic capabilities. First a holistic collaborative work involving multiple stakeholders from business and software functions whereby speeding continuous planning and innovation of ideas [13]. Second continuously deploying of software builds through automating software delivery processes and eliminating wastes and this is known as continuous delivery [14]. Third Identify problems as early in the process and notify development teams as quickly as possible that means providing a feedback loop for continuous learning from customers by monitoring and optimizing the software driven innovation [15].

## C. Related Work

There have been a number of publications focusing on the relationship between Lean Software Development and Agile, understanding of the combined use of agile and lean approaches in software development and investigate how agile and lean approaches have been combined in software development [16,17]. On the other hand, Shahid Mujtaba identify waste-related problems in a software product customization process by using value stream maps (VSM) [18] but they did not provide empirical evaluation of value stream maps in the software engineering. Pilar Rodríguez [19] presented some of challenges when applying Lean Software Development as achieving flow, transparency and creating a learning culture but unfortunately did not elaborate on ways to overcome these challenges. Henrik Jonsson [20] provided a framework for lean software development but he did not provide empirical evidences.

Finally, Pilar Rodríguez et al. [21] identified some bottlenecks in Lean Software Development as lack of collaboration between the hardware and software teams and short feedback loops from teams that led to not easy to involve business management to prioritize the backlog and defining

the feature content but did not provide ways to overcome these challenges so this study is concentrated on determine challenges of lean software development to enhance lean process by using DevOps.

## III. PROPOSED DEFINITION

Business, competitive advantage, market and customer needs are forcing organizations to develop and deploy applications, products, and services at a fast rate. When talking about lean there are lack of coordination between different elements, tasks or features led to barriers in achieving process flow. Problems occurring in the integration between features or there is not enough time to apply these features and monitor them by operational teams. So it is needed to new approach to enable early integration between development and operation teams to enable merge between them and create operational parts as one part of the development process and made it up to date during the development phase, so reduced errors during the deployment.

## IV. RESEARCH APPROACH

The goal of this research was to study how using DevOps practices to enhance lean software development process through identify reasons of lean wastes and DevOps role to overcome this reasons and provide framework that allows integration between them. The approach was developed in Three-phase model. The first phase was to determine reasons of lean wastes and the role of DevOps in overcoming them. The second phase was to provide framework that allow merging between two approaches with implement an empirical study that applied to a sample university to make sure that merging is possible. The Third phase has been conducted in order to measure the effect of this merge.

## V. RESEARCH ASSUMPTIONS

- Using lean software development process and DevOps practices together will be useful for the organizations to achieve better team productivity and predictability of problems.
- Rapid feedback provides accurate expectations for customer needs that lead to lower levels of deployment pains and lower change fail rates.

## VI. STUDY PHASE 1: THE CHALLENGES OF LEAN SOFTWARE DEVELOPMENT AND THE ROLE OF DEVOPS TO OVERCOMING THEM

Any organization needs to find and fix issues early before they are available in production phase. For this reason, you need to team members work together as DevOps team, so the following section will explain the causes of the lean software development wastes and the role of DevOps in improving and addressing the following wastes:

LW1: Delays:

Reasons:

- Any worker in the system might be not understanding of the client's needs in a particular item partially, or

have not all of the knowledge that you need to complete your task, so delays work to be achieved.

- Approval on the work to transfer it from phase to another phase for example the completion of development work to beginning of deployment.
- Lack of trust between team members causes delays to the work.
- Delayed feedback lead to start with feature may be not needed by the customer.

DevOps practices applied to solve this reasons:

- Be sure to make the knowledge necessary available to execute the project and delivered on time by using a clear vision of the elements that will be implemented.
- Make sure the client's needs, which wants always through feedback iterations.

LW2: Extra Features:

Reasons:

- Misunderstanding of the customer expectations requirements.
- More features more testing that means quality assurance team will be busy with a lot of tests and this also will affect the developers.
- Developers add items that they believe will improve the product in the end, and it will add a better perspective of the product. This may have on increase features as a result of suggest other new features.
- Customers add items that they believe it will benefit the project in the end.

DevOps practices applied to solve this reasons:

- All extra items, features or codes must be tracked, compiled, integrated, tested and maintained so we need continuous improvement to delivery software processes.
- We need to more careful about what we produce and the real needs of the customer therefore it must take more interest about the product backlog. Adding item's functionality when it is necessary to meet a need of the customer. As a result, it helps to discover features that not added value to customer and will not developed.
- Continuous feedback as a DevOps practices enables rapid and continuous expectation of customer needs and give the ideal system, therefore eliminated extra features through tracking product easily and remove features that are not needed.

LW3: Re-Learning (Extra Processing):

Reasons:

- Passing knowledge from person to other person that required re-explain work to provide value and re-sharing it.

- Distributed tasks that lead to exchange between tasks, and thus loss of time to restore focus on the task to be performed.
- Captured poor knowledge lead to rediscovery of that same knowledge.

DevOps practices applied to solve this reasons:

- Sharing knowledge and meeting between team members and depend on experts that can provide the benefit of your project. Everyone in the team can face problem in his work so we can reduce rediscover something from another developer.

LW4: Motion:

Reasons:

- Need for more information because of a lack of understanding about task.
- Distributed work between team members lead to indirect communication between them.

DevOps practices applied to solve this reasons:

- Use visual boards to this can be helpful to reduce hand-off time.
- It's good to have cross-functional teams to create a single project team.
- Shorten feedback loops reduce the number of hand-offs.

LW5- Partially work done:

Reasons:

- Poor analysis for customer needs lead to problems in establishing proper requirements.
- Lack of active participation between environments as development and operations which leads to disruption work in production environment.
- The discovery of errors or detected later may lead to incomplete feature.

DevOps practices applied to solve this reasons:

- DevOps enables to see the whole system end to end view of system from inception to deployment the system to the customer through tracking work done. If feature "done" means "developers declare it to be done and deliver it into a production environment".
- Product backlog must be declared before started in execute features and determine the person who works on specific task and the time to do it so we need coordination between the team and the product owner.

LW6: Task Switching:

Reasons:

- There are tasks will deliver value to a different customer and every one of customers want realize value as

possible as and which results in switching between tasks to satisfy all customers.

- Delays may lead to task switching. If developer doesn't have the knowledge that needed to complete the task this will lead to switch to another task.

DevOps practices applied to solve this reasons:

- Make sure that you have a detailed knowledge well before the start of the task until does not happen disabled during execute the task.
- Determine the priority of the stories during planning phase. This prevents task switching.

LW7: Defects:

Reasons:

- Lack of proper the automation testing.
- Lack of understanding of the items clearly according to the standards that have been determined in advance.

DevOps practices applied to solve this reasons:

- Do some practices that enable team members to communicate with each other so easily give appropriate comments on the work before entering into began defect.
- In order to be successful project there must be a strong automated test which is a very important element to discovery defects early.
- Make sure you have a complete understanding about the item being implemented.

In short, DevOps can overcome the causes of lean software development wastes through using DevOps practices. This result can be displayed through table I.

### VII. STUDY PHASE 2: EMPIRICAL STUDY AND INTEGRATED FRAMEWORK FOR LEAN SOFTWARE DEVELOPMENT AND DEVOPS

Once theoretical study were established through determine the causes of Lean Software Development wastes and DevOps role in addressing them, using framework to enhance Lean Process as shown in Fig. 1.

TABLE I. MATRIX OF LEAN SOFTWARE DEVELOPMENT WASTES AND DEVOPS PRACTICES

Lean Wastes/DevOps Practices	Continuous Planning	Continuous Feedback	Continuous Delivery	Continuous Integration	Continuous Testing	Continuous Monitoring
Waiting	√	√				
Extra Features	√	√	√		√	√
Extra Processing	√					
Motion	√	√		√		
Partially done work	√		√			
Task switching	√	√				
Defects	√				√	√

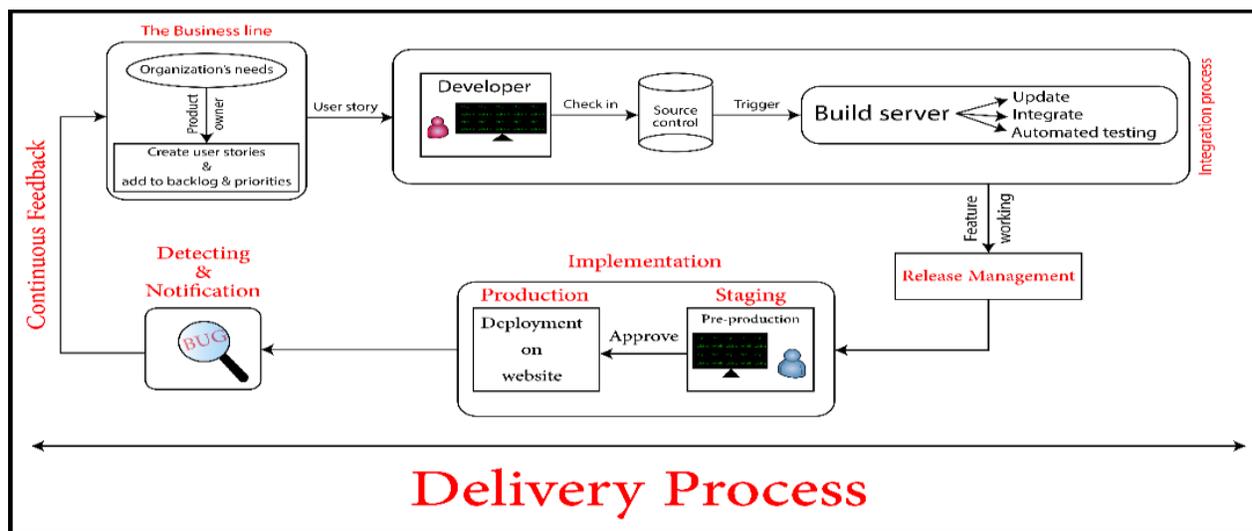


Fig. 1. Lean and DevOps framework.

### A. Study description

Applying this framework through empirical study using Evaluation website to provide ability to students to participate in evaluation of the doctors who taught and the subjects studied in the previous term in Faculty of Commerce Helwan University, also used Visual Studio Team Services used to manage and track work items cross the team to address the entire software development lifecycle and Application Insights to help in diagnose issues and to understand what users actually do with your app. There are two programs in Faculty of Commerce: the first BIS program which has been used for lean development software and the second FMI program which has been used for DevOps process. The framework can be illustrated by the following steps:

#### 1) The Business Line:

Any successful project or software the first step will be to plan and communicate continuously between the organization

and the team members responsible for this work. Continuous planning of DevOps practices allows doing that by always having a product backlog and prioritizing each item. As shown in Fig. 2. The project is divided into a set of User Stories and each user story is divided into a group of Tasks.

Product owners guide the development through creating a clear and inspiring product vision together with the team, ensure that customer value is transparent for everyone and focus on the highest priority goal at any one time.

#### 2) Integration Process:

Developers can now integrate code which allows checking at any time whether the product meets what the customer really values continuously. Continuous integration (CI) allows team members integrate their work frequently and verified by an automated build and testing to detect integration errors as quickly as possible as shown in Fig. 3.

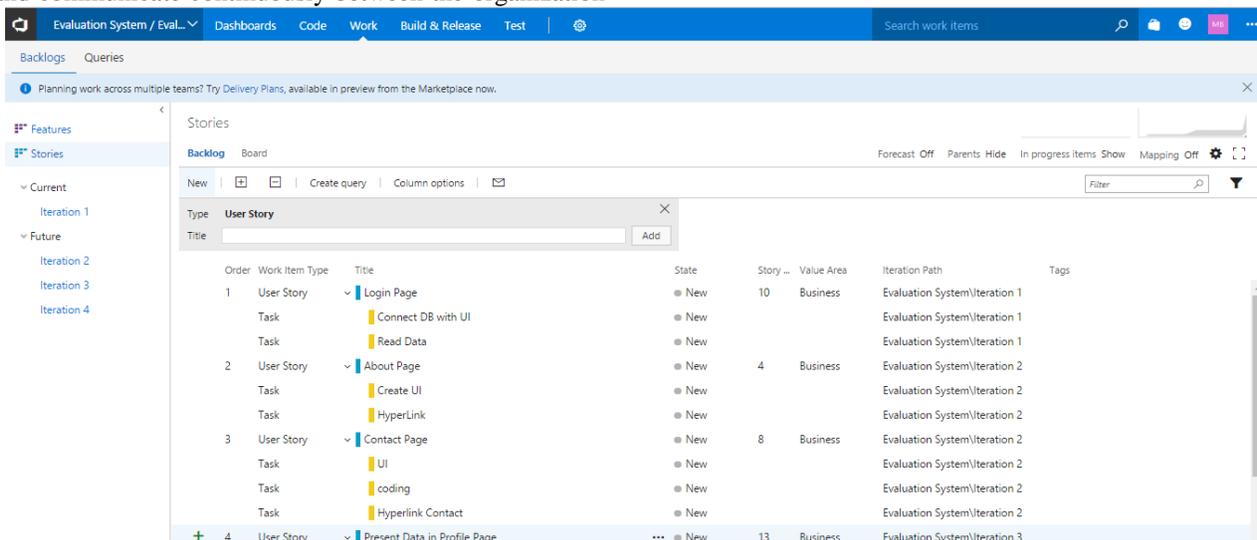


Fig. 2. User stories and Tasks I.

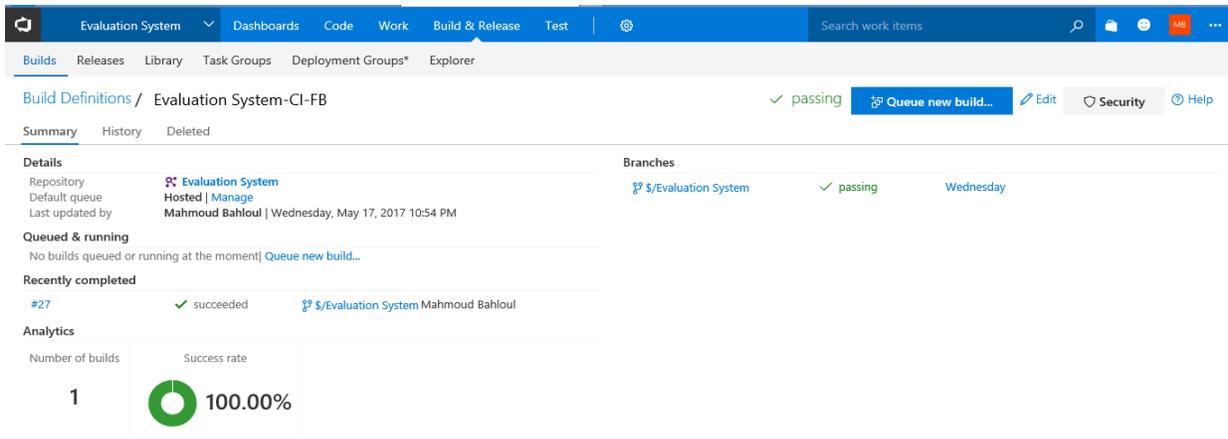


Fig. 3. Build process.

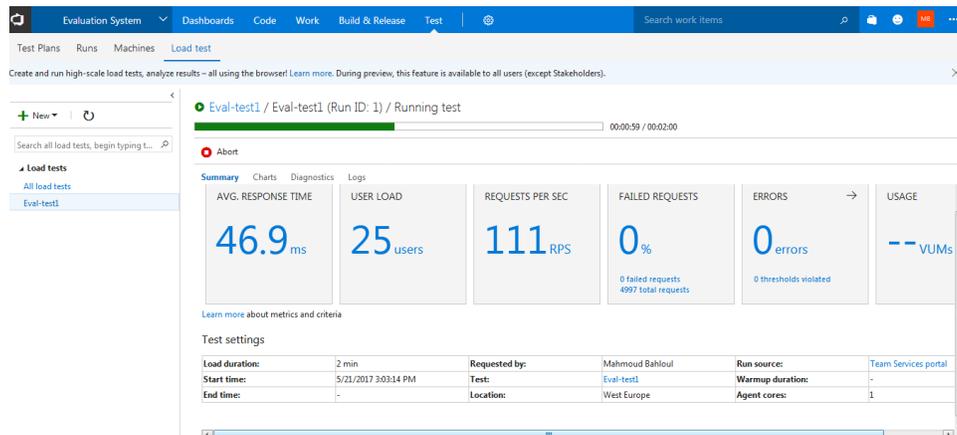


Fig. 4. Quality assurance environment.

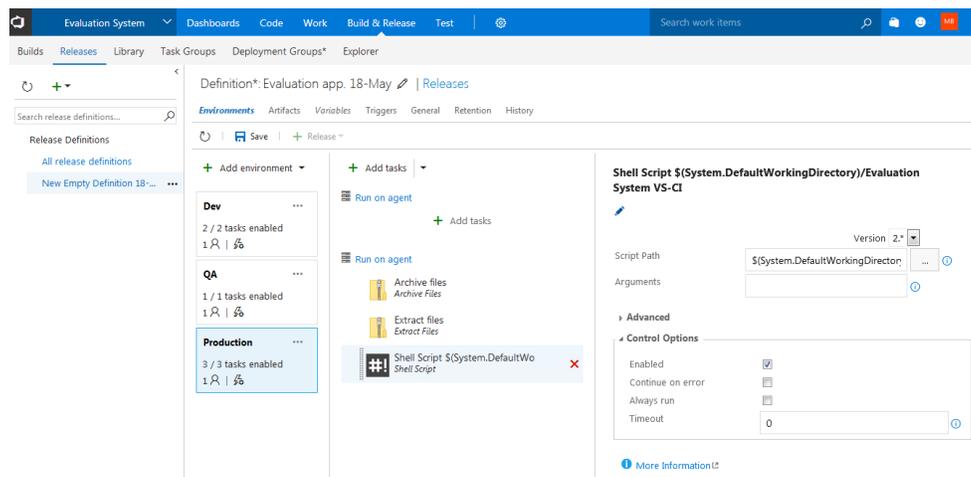


Fig. 5. Production environment.

The version control used to control changes in the source code and other software elements. The build server first check out the project from source control and executes targets from the build automation tool. It creates an integration build when any changes in the version control repository are made. Also as part of continuous integration process, we can do automated testing of the whole application along with code analysis to detect any issues or bugs and to analysis the code to obtain a quality build from the CI server.

When finish the developer from writing code and all unit tests is true, Quality Assurance (QA) team can only ensure that once a product is released it meets and matches all the quality criteria's and to build and ensure that the several testing and validation processes are improved continuously through automated testing which provides tests quickly as shown in Fig. 4.

### 3) Implementation:

The Release management is process responsible for planning, scheduling, controlling the build, testing and deploying release to increase numbers of successful releases through avoid unexpected outcomes. The quality assurance team gives the change its seal of approval. The change moves on to the staging server, where final acceptance testing

commences. The staging or a pre-production environment is providing for the end-users to test the application. When the end-user accepts about this feature, the release to the production environment is performed as shown in Fig. 5.

### 4) Delivery Process:

The main focus of DevOps is all about delivering *value* very quickly to users and customers. Communication between team members can definitely make it easier to automate deployments. If you need the team who write the deployment scripts to collaborate with the people who manage the environments and run the scripts, you need to continuous delivery.

Emphasis on collaboration and feedback is very important to get successful working software, so DevOps focus on everyone involved in a project must communicate with each other constantly and to continuous delivery new values and releases for customers to be able to achieve a very short time to market as shown in Fig. 6.

### 5) Detecting, Notification & continuous Feedback:

Using DevOps and lean software development all customer expectations can be answered quickly. It helps us to get connected to the users and to act rapidly against them and

understand how they use their system. As shown in Fig. 7 it explains how trace data can detect errors at the same time.

This direct tracking of data enables me to know why and where the error occurred and how often specific events happen

in my web app as student evaluated specific course. It has become very easy identify the features that are used constantly or not used. This continuous follow-up allow to know the most visited pages of the users as well as the pages do not care about as shown in Fig. 8.

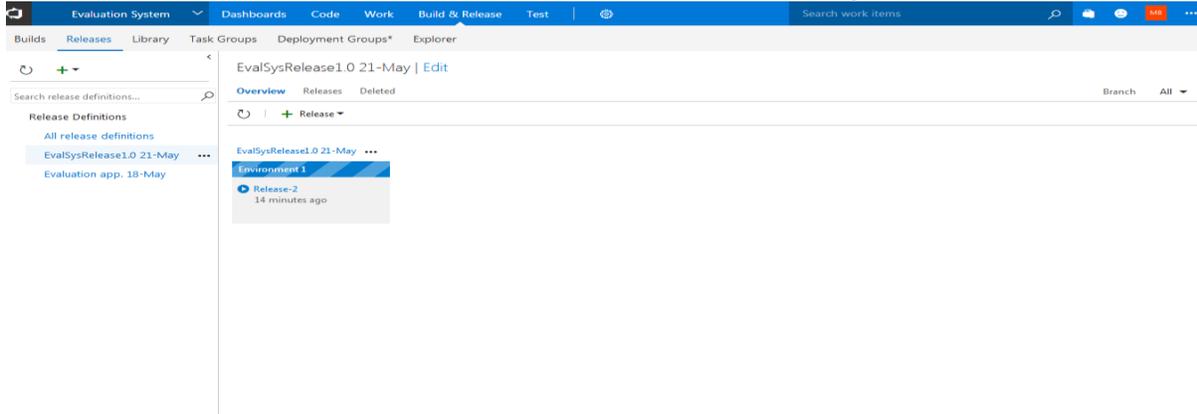


Fig. 6. Release delivery.

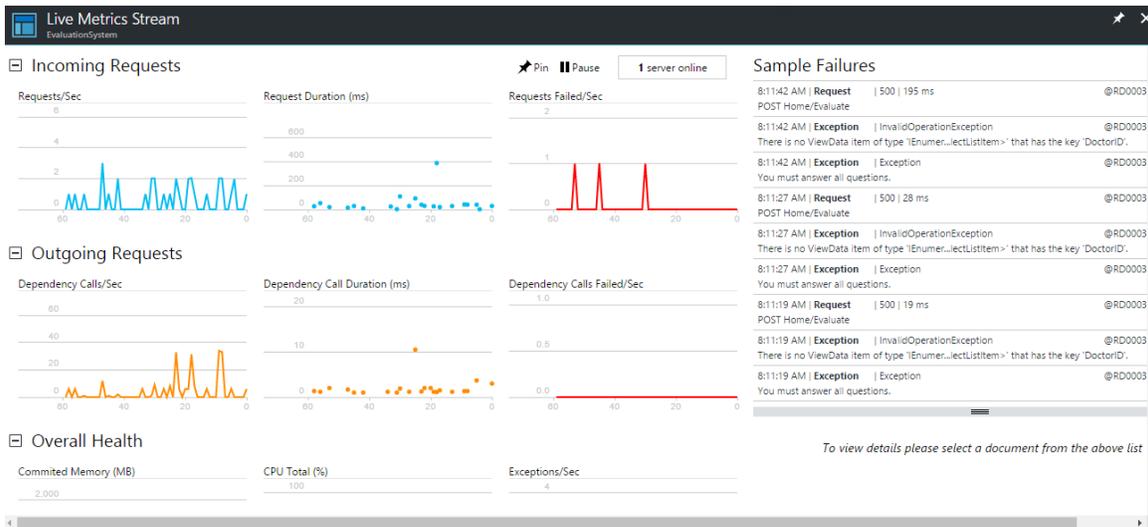


Fig. 7. Live detection.

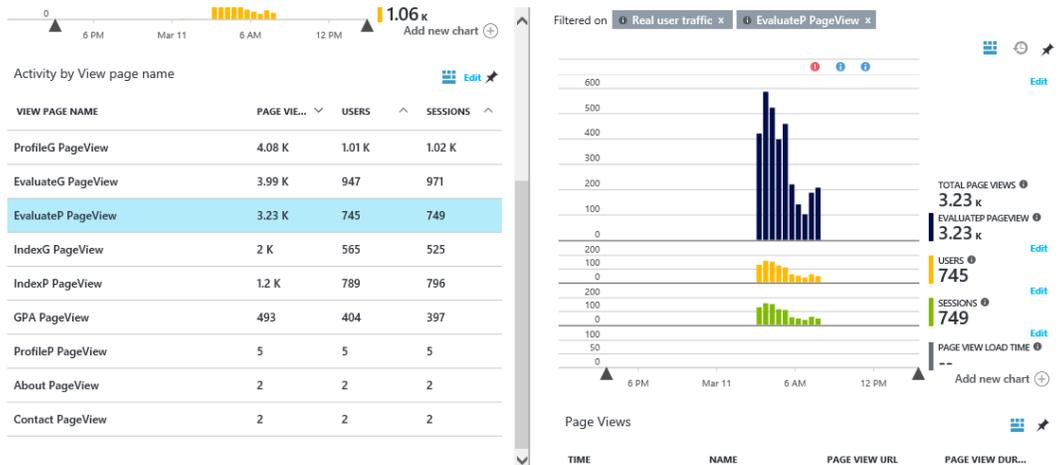


Fig. 8. Page views.

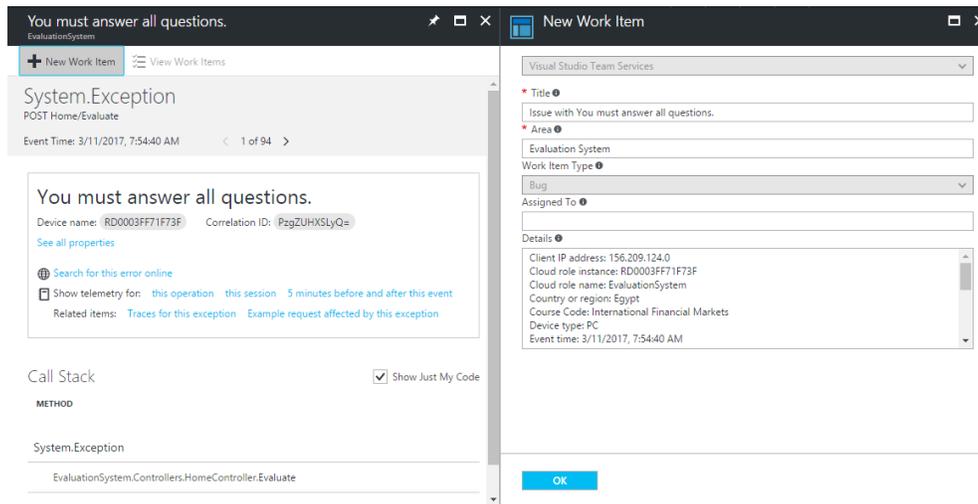


Fig. 9. New work item (bug).

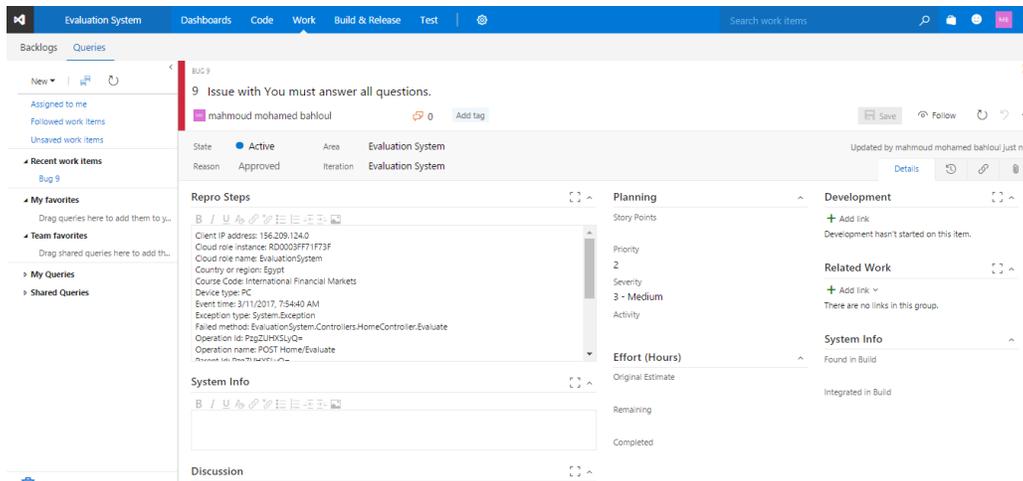


Fig. 10. Issue assigned to development team.

The success of any system not only depends on effective monitoring, but should be followed by a quick warning of any possible errors or problems in the system. Continuous monitoring leads to get immediate feedback on a deployment. When the operations team members discover the error, they send the data obtained to the developers to address and fix it as soon as possible as shown in Fig. 9.

The detected problem reaches to the developer in his queries as a bug to solve it and he gives it his priority as shown in Fig. 10.

### VIII. STUDY PHASE 3: VERIFYING APPLICABILITY OF RESEARCH WORK

The project was divided into 4 iterations. When using LSD the number of hours worked was expected to be 132 hour, but with the emergence of these errors and increased these hours became actual hours are 139 hour as shown in Fig. 11.

But when applying DevOps the number of hours worked was expected to be 132 hour, but with the emergence of these errors and increased these hours became actual hours are 135 hour as shown in Fig. 12.

Process	Start Date	Estimate End	Estimate Working Hours	Estimate Bugs	Actual End	Actual Hours
iteration 1	1/2/2017	4/2/2017	24		4/2/2017	24
iteration 2	6/2/2017	9/2/2017	24	2	9/2/2017	26
iteration 3	11/2/2017	19/2/2017	44	2	19/2/2017	46
iteration 4	21/2/2017	28/2/2017	40	3	28/2/2017	43

Fig. 11. Lean process.

Process	Start Date	Estimate End	Estimate Working Hours	Estimate Bugs	Actual End	Actual Hours
iteration 1	1/2/2017	4/2/2017	24		4/2/2017	24
iteration 2	5/2/2017	8/2/2017	24	1	8/2/2017	25
iteration 3	9/2/2017	17/2/2017	44	1	17/2/2017	45
iteration 4	18/2/2017	25/2/2017	40	1	25/2/2017	41

Fig. 12. FMI DevOps process.

Enhancement	Delivery Date	Resolved Date	# ofDays
I1: Writing characters in student id	5/2/2017	9/2/2017	4
I2: Repeation doctors and students data	20/2/2017	28/2/2017	8
I3: Postback evaluate page	2/3/2017	3/3/2017	1

Fig. 13. Enhancement issues in lean process.

Enhancement	Delivery Date	Resolved Date	# ofDays
I1: Writing characters in student id	5/2/2017	6/2/2017	0
Repeation doctors and students c	20/2/2017	21/2/2017	1
I3: Postback evaluate page	2/3/2017	2/3/2017	0

Fig. 14. Enhancement issues in DevOps process.

Issue	Number of affected	Total	Presentage
I3: Postback evaluate page	50	320	15.625

Fig. 15. Number of affected particular issue in lean process.

Issue	Number of affected	Total	Presentage
I3: Postback evaluate page	4	1200	0.33333333

Fig. 16. Number of affected particular issue in DevOps process.

In LSD the errors detected by the user is then reported to the operations team and the development team members is notified of this error by regular meetings or team reminder email. Consequently, it was a loss of a lot of time until the update of what is new and satisfy the needs of customers. As shown in Fig. 13 the number of days needed to solve some of the problems that appeared in the system.

Using DevOps errors are detected at the time they occur or even there are notifications of an error, this led to reduce the number of days to resolve the errors. As shown in Fig. 14 the number of days needed to solve some of the problems that appeared in the system

In LSD Delay in product improvement response lead to leaving the customer to the product as a result of the number of customers affected by the problem is large and consequently leads to the loss of a large number of customer loyalty to the company. As shown in Fig. 15 the number of

students affected by Postback Evaluate Page is 50 students from total 320.

In DevOps finding errors quickly led to reduce the number of students to be affected by errors and it becomes very small. As shown in Fig.16 the number of students affected by Postback Evaluate Page is 4 students from total 1200.

#### IX. DISCUSSION:

From the previous study, the results can be summed up, which displaying the role of DevOps to enhance and improve Lean Software Development as shown in Fig. 17 and 18.

Rapid feedback provides accurate expectations for customer needs that lead to lower levels of deployment pains and lower change fail rates. The percentage of students who were affected by specific issue through using lean software development was greater than the percentage of students who were affected by the DevOps.

	Lean	DevOps
Presentage of affected	15.625	0.333333333
Predictability	0.2	0.8
Delivery Time	10	90

Fig. 17. Summarizes up the results of using both lean software development and DevOps.

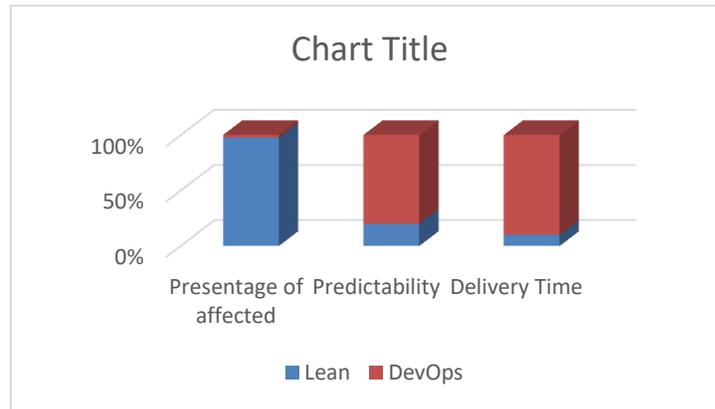


Fig. 18. Summarizes up the results of using both lean software development and DevOps.

Improve the company's time and market potentially from months and weeks to days and hours in addition to faster time to value through reduced cycle times. Consequently, the predictability percentage of problems that may occur or know and follow what the user is doing through DevOps is much higher than the lean software development.

Changing business needs always required to provide products for faster time to market due to Competition among software companies puts an increasing pressure to produce new features extremely fast. Consequently, DevOps increases the rate of software delivery more than lean software development.

#### X. CONCLUSION AND FUTURE WORK

As explained above in this paper how using practices of DevOps to enhance Lean Software Development that allows to cover the entire life-cycle from development to operations environments. Enhancing of Lean Software Development process was done through determine the causes of the lean software development wastes and how using DevOps practices in improving and addressing this wastes. The reasons identified in the Lean Software Development and the role of DevOps in the addressing or improvement this wastes lead to create a new lean and DevOps framework that used to enhance process of Lean through reduce time to market and increases the rate of software delivery.

Changing business needs always required to provide products for faster time to market due to Competition among software companies puts an increasing pressure to produce new features extremely fast so that need to rapid feedback that provides accurate expectations for customer needs that lead to lower levels of deployment pains and lower change fail rates.

In the light of that, it is suggested as future works: The field of software engineering is changing very quickly so more studies will help the companies to apply this framework in other large environments and to use DevOps Practices to enhance other process methods as Scrum or XP.

#### REFERENCES

- [1] MULLAGURU, Surendra Naidu. Changing Scenario of Testing Paradigms using DevOps—A Comparative Study with Classical Models. *Global Journal of Computer Science and Technology*, 2015, 15.2.
- [2] RODRÍGUEZ, Pilar, et al. Survey on agile and lean usage in finish software industry. In: *Empirical Software Engineering and Measurement (ESEM), 2012 ACM-IEEE International Symposium on*. IEEE, 2012. p. 139-148.
- [3] POPPENDIECK, Marv; CUSUMANO, Michael A. Lean software development: A tutorial. *IEEE software*, 2012, 29.5: 26-32.
- [4] WETTINGER, Johannes; ANDRIKOPOULOS, Vasilios; LEYMANN, Frank. Automated capturing and systematic usage of DevOps knowledge for cloud applications. In: *Cloud Engineering (IC2E), 2015 IEEE International Conference on*. IEEE, 2015. p. 60-65.
- [5] M. Poppendieck and T. Poppendieck, *Lean Software Development: An Agile Toolkit (The Agile Software Development Series)*. 2003, p. 240.
- [6] AL-BAIK, Osama; MILLER, James. Waste identification and elimination in information technology organizations. *Empirical Software Engineering*, 2014, 19.6: 2019-2061.
- [7] POPPENDIECK, Mary; POPPENDIECK, Tom. *Implementing lean software development: From concept to cash*. Pearson Education, 2007.
- [8] JONSSON, Henrik; LARSSON, Stig; PUNNEKAT, Sasikumar. Synthesizing a comprehensive framework for lean software development. In: *Software Engineering and Advanced Applications (SEAA), 2013 39th EUROMICRO Conference on*. IEEE, 2013. p. 1-8.
- [9] HITTERMANN, Michael. *DevOps for developers*. Apress, 2012.

- [10] DEBOIS, Patrick. Agile infrastructure and operations: how infra-gile are you?. In: *Agile, 2008. AGILE'08. Conference*. IEEE, 2008. p. 202-207.
- [11] HUMBLE, Jez; MOLESKY, Joanne. Why enterprises must adopt DevOps to enable continuous delivery. *Cutter IT Journal*, 2011, 24.8: 6.
- [12] HUMBLE, Jez; FARLEY, David. *Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation (Adobe Reader)*. Pearson Education, 2010.
- [13] LEHTOLA, Laura, et al. Linking business and requirements engineering: is solution planning a missing activity in software product companies?. *Requirements engineering*, 2009, 14.2: 113-128.
- [14] NEELY, Steve; STOLT, Steve. Continuous delivery? easy! just change everything (well, maybe it is not that easy). In: *Agile Conference (AGILE), 2013*. IEEE, 2013. p. 121-128.
- [15] CUKIER, Daniel. DevOps patterns to scale web applications using cloud services. In: *Proceedings of the 2013 companion publication for conference on Systems, programming, & applications: software for humanity*. ACM, 2013. p. 143-152.
- [16] WANG, Xiaofeng. The combination of agile and lean in software development: An experience report analysis. In: *Agile Conference (AGILE), 2011*. IEEE, 2011. p. 1-9.
- [17] WANG, Xiaofeng; CONBOY, Kieran; CAWLEY, Oisín. "Leagile" software development: An experience report analysis of the application of lean approaches in agile software development. *Journal of Systems and Software*, 2012, 85.6: 1287-1299.
- [18] MUJTABA, Shahid; FELDT, Robert; PETERSEN, Kai. Waste and lead time reduction in a software product customization process with value stream maps. In: *Software Engineering Conference (ASWEC), 2010 21st Australian*. IEEE, 2010. p. 139-148.
- [19] RODRÍGUEZ, Pilar. et al. Building lean thinking in a telecom software development organization: strengths and challenges. In: *Proceedings of the 2013 international conference on software and system process*. ACM, 2013. p. 98-107.
- [20] JONSSON, Henrik; LARSSON, Stig; PUNNEKAT, Sasikumar. Synthesizing a comprehensive framework for lean software development. In: *Software Engineering and Advanced Applications (SEAA), 2013 39th EUROMICRO Conference on*. IEEE, 2013. p. 1-8.
- [21] RODRÍGUEZ, Pilar, et al. Combining lean thinking and agile methods for software development: A case study of a finnish provider of wireless embedded systems detailed. In: *System Sciences (HICSS), 2014 47th Hawaii International Conference on*. IEEE, 2014. p. 4770-4779.

# SEUs Mitigation on Program Counter of the LEON3 Soft Processor

Afef KCHAOU

University of Tunis El Manar, Faculty of Sciences of Tunis  
Laboratory of Electronics and Microelectronics of Monastir  
Monastir, Tunisia

Wajih EL HADJ YOUSSEF

University of Monastir  
Laboratory of Electronics and Microelectronics of Monastir  
Monastir, Tunisia

Rached TOURKI

University of Monastir  
Laboratory of Electronics and Microelectronics of Monastir Monastir, Tunisia

**Abstract**—Analyzing and evaluating the sensitivity of embedded systems to soft-errors have always been a challenge for aerospace or safety equipment designer. Different automated fault-injection methods have been developed for evaluating the sensitivity of integrated circuit. Also many techniques have been developed to get a fault tolerant architecture in order to mask and mitigate fault injection in a circuit. Fault injection mitigation and repair techniques are applied together on LEON3 processor in goal to study the reliability of a soft-core. The so-called NETlist Fault Injection (NETFI+) tool is a fault injection techniques used in this paper. The prediction of Single Event Upset (SEU) error-rates between radiation ground testing and FPGA implementation have been done with good and accurate result. But no functional simulations have been performed. A Triple Modular Redundancy (TMR) is used in this paper as a repair technique versus fault injection. This paper analyses the effectiveness of fault tolerant method on LEON3 soft-core running a benchmark. It starts by evaluating the behavior of LEON3's program counter against Single Event Upset error-rate accuracy between the functional simulation and the FPGA emulation and an analysis of the LEON3 reliability in presence of fault tolerant technique. The objective is to offer, through the new version of NETFI+ with introducing a fault tolerant technique, the possibility to designers to evaluate the benefits of SEUs mitigation for the LEON3 processor on the program counter.

**Keywords**—NETFI+; fault injection; SEUs; LEON3; simulation; emulation; reliability; TMR

## I. INTRODUCTION

Embedded system undergoes several changes across the years, starting from simple mono CPU running applications to a complex system including co-processor, memory, input and output models. Using embedded systems in special applications, safety-critical or mission-critical, allows evaluating their dependability in presence of faults on the circuit or in the implemented application.

Fault injection can be used to evaluate embedded system running its own application [1].

Transient faults and soft errors lead to faults in a system without damaging the system under evaluation. Transient faults

are represented by a single or multiple nodes upset directly attributable to excess charge carriers created by an external source of radiation. Soft errors are defined by the impact of a transient fault that can be propagated beyond one clock cycle [2]. It flips one or more bits, modifying the data store of a memory cell, register, flip-flop and latch. SEU and Single Event Transient (SET) are soft errors that affect only one bit, other type of faults are used to modify more than one bit, and it's a Multiple Bit Upset (MBU) [3].

Areoflex Gaisler LEON3 processor has become more used in a critical and safety application, such as in automotive, multimedia system, wireless and more applications which require reliability. Fault injection in LEON3 soft-core is done in many works classified according to the type of faults, the methods used, the block under test, etc. LEON3 is characterized by its complexity and size. It's a reason to be a good design for evaluating the benefits of fault tolerant techniques [4].

In [5], injection of SEU, SET and MBU faults have been done in many components of LEON3, showing that integer unit and multiplier unit are more susceptible against SEU and MBU fault injection.

Emulation-based fault injection in LEON3 is done in [6], allowing a reduction in the experimental time.

In [7], SEU fault injection by FPGA emulation is made by applying an exhaustive fault injection in internal memory of LEON3. The results obtained show that the memory cell containing the data is the most sensitive to SEU.

In [8], a new methodology is proposed to evaluate the real cache sensitivity for a given application, and to calculate a more accurate SER. The methodology, based on monitoring the memory accesses, is applied to the LEON3 with several benchmarks showing that their proposed tool predicted all real errors with little over-estimation. Fault injection is done by radiation and emulation. The result shows that all the cache addresses are sensitive to SEU injection.

In [9], evaluating the effects of single bit errors at the memory and register locations is done using a high level error

injection technique. The results obtained show that this method is inaccurate in comparison with the techniques using a flip-flop error injection.

SEU fault injection, called bit-flip, upset or soft-error propagates in the design depending on the application [10] and it can cause a data corruption or a circuit malfunction. SEUs are random in space and time, they can modify any element on memory location also at any instant time. In [11], the application of a new methodology to attack a Program Counter (PC) of ARM is done by modifying the load-instruction of the PC.

In previous works, the PC of LEON3 was not evaluated in point of sensitivity because to its important function in system security, in addition, based on the work in [5] avowed that the integer unit is the most critical block of LEON3 to SEU and MBU fault injection. The principal goal in this paper is to give the benefits of SEU mitigation for the LEON3 processor on the PC by adding a repair technique of fault tolerant like a TMR in this work.

A new methodology is improved for fault attacks, NETFI+, in order to evaluate the behavior of LEON3 soft-core, the SEU error-rate reliability between the functional simulation and FPGA emulation is done by injecting an SEU fault on the flip-flop of LEON3's IU block precisely on the program counter register according to its importance in the instruction execution process. The principal goal in this paper is to present fault injection approach and analyze countermeasure effectiveness in circuit security.

The paper is organized as follows. In Section II a description of the NETFI+ principle will be done, overview of LEON3's integer unit is presented in Section III. Next section presents the NETFI+ flow, the analysis of the reliability of the LEON3's program counter by evaluating the SEU error-rate reliability between simulation and FPGA emulation is presented in Section V. Section VI provides a presentation of a repair technique used in order to evaluate the SEU mitigation for the LEON3 soft processor on the program counter. A conclusions and perspectives will be presented in the last section.

## II. NETLIST FAULT INJECTION PRINCIPLE

Fault injection approaches are based on injecting faults that can induce errors. Many researches separate between the methods of fault injection, it can be classified depending on two techniques based on hardware and software fault injection.

A software fault injection is presented by using a software program to inject faults in a physical model. Simulated fault injection can be observed and controlled while the system is simulated using HDL simulator.

A hardware fault injection allows evaluating a behavior of a system based on Commercial off-the-shelf (COTS) processor [12]. It's widely used and it can be classified on three categories [13]:

- Logical fault injection using debugging facilities: This type of injection allows to the processor logic resources to access their internal blocks and to add bit-flips.
- Physical fault injection: this method is accomplished using laser beams, electromagnetic interferences or a radiation in goal to induce faults in integrated circuits. This method offers actual hardware faults on real systems. It requires expensive material and the number of faults injected is limited, also a deep knowledge of the actual layout of the circuit [14].
- Logical fault injection: it can be made by circuit simulation using hardware description languages (HDL models) simulator or by circuit emulation using hardware emulation platforms. In simulation-based fault injection, the system under test is simulated in another system, while the emulation-based fault injection facilitates the injection on complex models by reducing time spent by a simulation-based fault injection.

The fault injection type used in this paper is the last one: Logical fault injection. Simulation-based fault injection allows the fault injection in high level models. In general, fault injection is presented by a bit-flip fault model when the content of a memory cell is inverted. It permits to evaluate the behavior of fault tolerance mechanisms [14].

The principle idea of the logical fault injection is the injection at the bit-flip model by inverting the content of a memory cell at the instant injection. Studying the reliability of an embedded system is a principal goal to define the capability of the system to run its function in abnormal condition for a given period of time [15]. Soft errors disturbing memory cells and registers in embedded system are called SEU was analyzed to evaluate the soft error-rate [13].

Sensitivity of the LEON3's integer unit against soft errors was estimated through two fault-injection campaigns. A first one was performed in a simulation in order to analyse a random SEU injection. A second one was performed in FPGA emulation to accelerate the fault injection campaign and mainly to evaluate the SEU error-rate reliability in the simulation campaign also the validation of a NETFI+ tool.

The NETFI+ tool used in this paper is based on netlist fault injection. It allows to inject SEU, MBU and SET faults in circuits at Register Transfer Level implemented on FPGA. This method enables to inject faults in all memory cell and at any clock cycle, exhaustive or randomly in time and location. In this paper, the principle idea is to study the reliability of LEON3's PC against SEU fault injection.

The HDL source code of the circuit is synthesized to get the correspondent netlist [16]. In next step, a MODNET (MODify NETlist) tool, described in [17], will be used to choose the type of faults which can be injected and give a modified netlist.

A NETFI+ tool in this work is improved to inject faults in all the memory cell of LEON3 also to inject all type of faults SEU, MBU and SET.

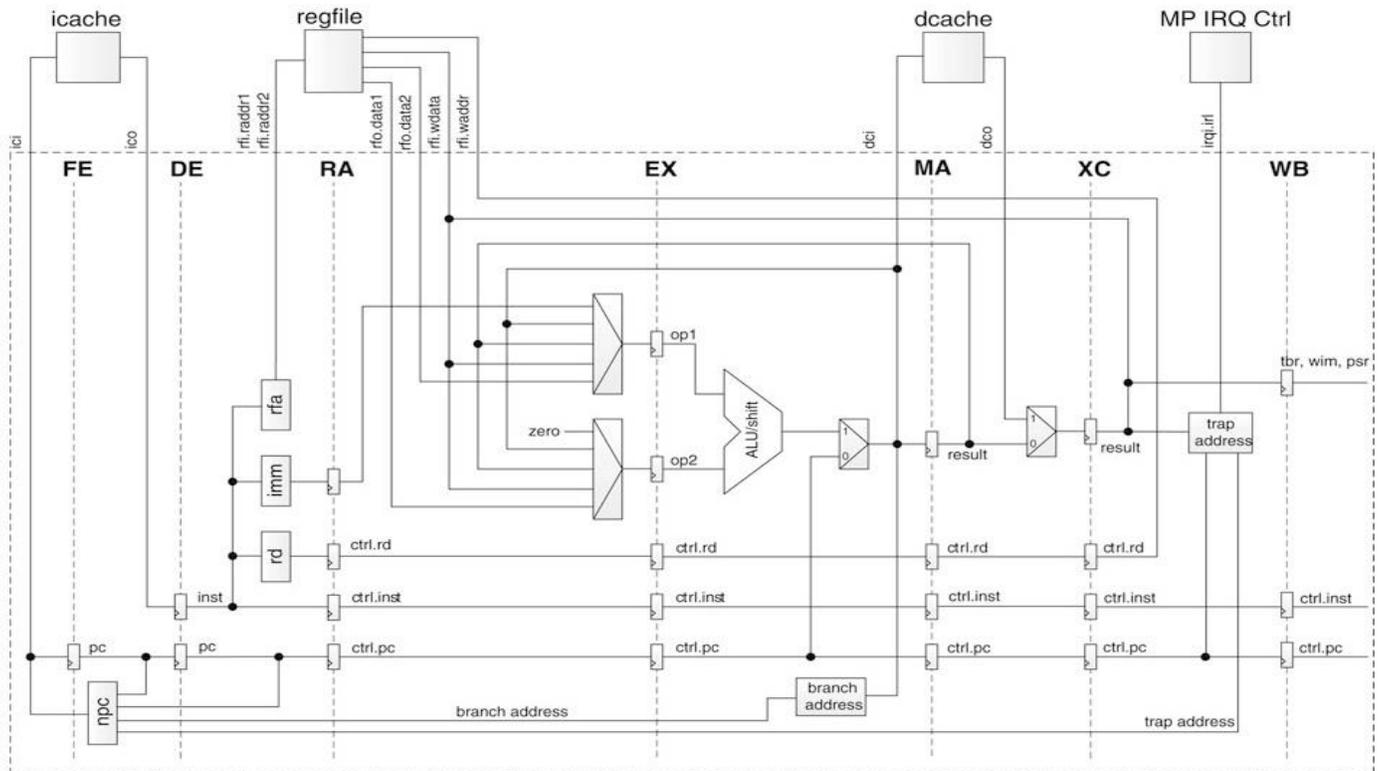


Fig. 1. The micro-architecture of the LEON3's Integer Unit [18].

### III. OVERVIEW OF LEON3'S INTEGER UNIT

The LEON3 Integer Unit (IU) is fully compliant with the SPARC V8 standards. SPARC is a CPU instruction set architecture derived from RISC. It comprises an integer unit (IU), an optional Floating-point unit (FPU) and a coprocessor (CP).

The IU executes the arithmetic instructions, computes memory addresses (load/store), maintains the Program Counter (PC) and controls the instruction execution for the FPU and the CP. "Fig. 1" shows the pipeline of the IU which consists of seven-stages with Harvard architecture.

IU integrates seven stage of pipeline, the FE stage (FEtch) fetches the instruction from the instruction cache through its address given by a PC. DE stage (DEcode) decodes the instruction. In the RA stage (Register Access), all operands are read from the register file or from the internal data bypasses and stored in EX stage (EXecute). ME stage (MEemory) stores the results and communication between IU and the other peripherals components which can be done. In XC stage (eXCeption) all traps and interrupts are resolved. In WR stage (WRite), a data not sent to the register file will be stored [19], [9].

Integer Unit controls, in general, all the operation of the processor and it includes two types of register: general-purpose registers and control/status registers. Whose General-purpose registers is a 32-bit registers, called r register.

An instruction can access the 8 global registers and a 24 registers window into r register. The register window contains 8 in and 8 local registers of a particular register set. The 8 in registers are addressable from the current window, the out registers.

The IU control/status registers include Processor State Register (PSR), Window Invalid Mask (WIM), Trap Base Register (TBR), Multiply/divide Register (Y), Program Counters (PC), Implementation-dependent Ancillary State Registers (ASRs) and Implementation-dependent IU Deferred-Trap Queue.

### IV. FAULT INJECTION FLOW

The emulation of SEU faults is done in the PC which is the overall security of any embedded system, in this case the LEON3 processor. PC gives the address of the instruction currently being executed by the IU.

Only 30-bit of PC will be used in the six stages of pipeline (FE, DE, RA, EX, ME, XC) because the LSB two bits of the PC are not used in the configuration, its implementation will cause a debug of the HDL model and an area waste in synthesis. A NETFI+ method allows injecting SEUs faults in a Flip-Flop of PC in all the stage of pipeline, in total 180 FF will be used to control the PC. "Fig. 2" illustrates the workflow adopted in the NETFI+.

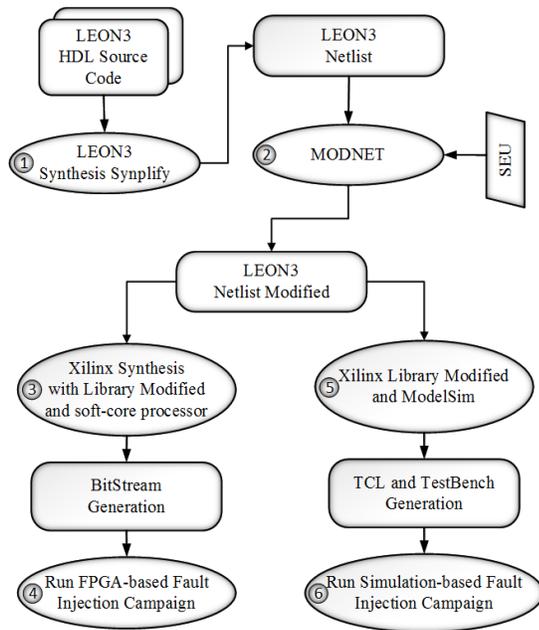


Fig. 2. Fault injection flowchart.

Initially, a Hardware Description Language (HDL) of the LEON3 is synthesized by Synplify tools to get the Verilog netlist in Step 1. This first step does not require any modification to the original design.

In Step 2, the first netlist resulting will be used as input for the MODNET tools, which adds a new signal “INJ” to all the Flip-Flop (FD and FDE) components used in the block of IU. After that, the new netlist obtained is then synthesized, by Synplify tools, in Electronic Design Interchange Format (EDIF) using a modified version of the sensitive components, which includes “INJ” signals to access them to fault injection. “Fig. 3” exhibits the addition of the ‘INJ’ signal in the design.

In Step 2, two possibilities of test injection can be applied, the first one is by FPGA emulation, steps 3 and 4, and the second one is by a simulation campaign, steps 5 and 6.

The FPGA emulation campaign is performed in Steps 3 and 4. In Step 3, the EDIF file obtained in Step 2 is then attached to the soft-core processor and the last synthesis is performed to generate a bitstream based on the target FPGA. Finally, in Step 4 the experiment is executed in hardware-based FPGA platform.

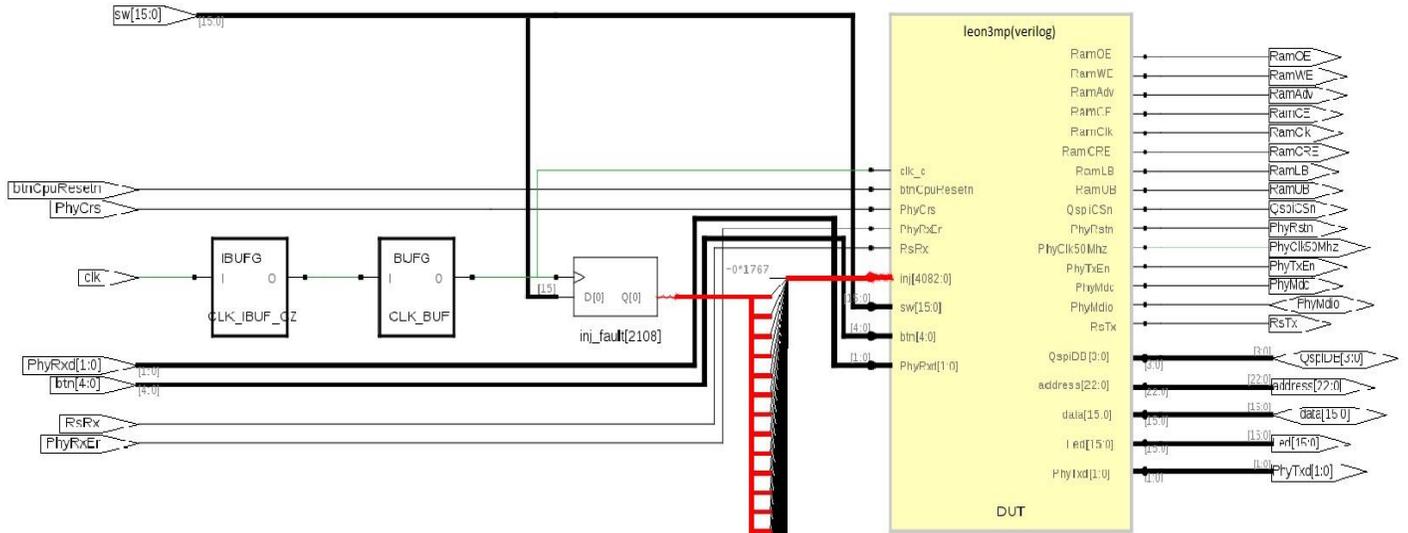


Fig. 3. Architecture of LEON3 block diagram which includes “INJ” signal.

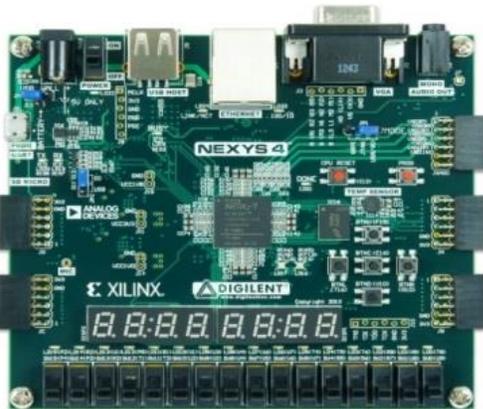


Fig. 4. The Nexys4 board.

In this work, a Nexys4 board, equipped with Xilinx Artix-7 XC7A100T-CS324, is used which is a complete circuit board. As shown in “Fig. 4”, the board is occupied by a diverse I/O, development connectors that allows a connection with the LEON3 implemented.

The setup and control of the fault injection experiments are performed by a Soft-Core Fault Injection Processor (SCFIP), embedded in the FPGA, and by a Tool Command Language (TCL) script in a personal computer, as can be seen in “Fig. 5”.

The SCFIP is used as the controller in charge of randomly selecting the time and in which registers will inject the faults. The results of LEON3 execution are sent to a personal computer connected through the UART interface.

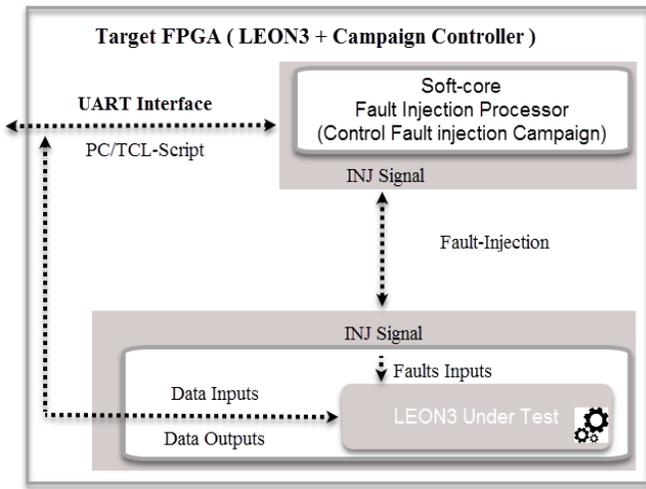


Fig. 5. FPGA fault injection strategy.

The simulation fault injection campaign following Step 5 and 6 of “Fig. 2”. In Step 5, the files obtained in step 2 are then attached to the Xilinx modified library in the ModelSim Mentor Tool [20]. Finally, in Step 6 the simulation, setup and the control of the fault injection campaign are performed by a TestBench in a dedicated server equipped with two Intel Xeon CPU E5-2620 and 64GB of RAM memory.

#### V. PROGRAM COUNTER RELIABILITY

The NETFI+ tool described on Section 2 will create a variety of SEU influence on LEON3. The first campaign of injection is done by a simulation fault injection. A benchmark, MulMatrix (Matrix product), is used to execute a simulation and to compare the results obtained with the standard results (Golden Results).

The benchmark used in this test is selected because it does not take enough time in RTL simulation also the number of repetitive instructions involved by its execution enables to guarantee a wrong behavior in presence of fault injection.

A random SEU fault injection in the PC register at the stages of pipeline is done. In total, 1080000 faults are injected in sensible FF of LEON3’s PC during one period, 100% of the flip-flop of FE and DE stage are sensitive to SEU, the FF of the other stages are non-sensitive.

The results obtained can be classified in four categories shows in “Table 1”: overwritten Faults, Failure results, Timeout and Stopped Execution.

The faults can be overwritten in some cases, it explains that the error is masked and cannot modify the result, this explains that at the moment of the injection, the PC does not point to the instruction used at the instant of simulation (about 100% faults undetected in RA, EX, ME and XC stage).

Failure results is procured when the simulation is done but giving a false result, this explains that at the moment of SEU injection, the PC does not point in the correct address of the currently instruction used in simulation (6.66% in FE stage and 3.33% in DE stage).

TABLE I. SEU FAULT INJECTION ANALYSIS

Type of Result	Stages of pipeline					
	FE	DE	RA	EX	ME	XC
%Overwritten	-	-	100	100	100	100
%Failure result	6.66	3.33	-	-	-	-
%Timeout	3.33	-	-	-	-	-
%Stopped execution	90	96.66	-	-	-	-

The simulation can exceed the approximate time of simulation (4365 us in this case) like a 3.33% in FE stage, this type of faults is named Timeout, the PC in this case stop incrementing and this explains a cause an infinite loop in simulation. Also the execution can be stopped (90% in FE stage and 96.66% in DE stage). In another way, the simulation cannot be finished normally, and the execution stops just after the moment of injection. In other words, the PC does not contain any address to be pointed.

The benefits of the simulation-based fault injection that it allows a fine-grained analysis through the assembly code of fault injection campaign. The assembly code contains all the instructions of simulation. SEU injection in PC occur some traps.

The results obtained are resumed in “Table 2”. The traps can be more detailed [21]:

- Illegal instruction:

When the simulation ends before a normal time of execution, an attempt is made to execute the instruction with an unimplemented opcode or an UNIMP instruction (Assembly code: unimp (trapped)). Other reason can be responsible for the stopped execution that the instruction would result an illegal processor state at the decode stage (Assembly code: save %sp, -0x0060, %sp (trapped)).

- Privileged instruction:

At the fetch stage, the PC stops to increment and remains constant and in another case, the PC does not contain any address to be pointed. The assembly code shows that an attempt was made to execute a privileged instruction (Assembly code: ldub [%o4], %o5 [0x0000XXXX]).

- Window overflows:

A SAVE instruction is responsible for this trap because at this instant the Current Window Pointer (CWP) will point to a window marked not valid in the WIM.

- Window underflows:

It’s caused by a RETT or a RESTORE instruction attempted in this case when the CWP would point to a window marked invalid in the WIM. In fact, RESTORE instruction has the same role of ADD instruction, it allows to increment by ‘1’ the CWP and to compare it with WIM, if the WIM bit corresponding to the new CMP is ‘1’, a window underflow trap is then generated.

TABLE II. TRAPS ANALYSIS

Type of Trap	Stages of pipeline	
	FE	DE
Illegal_instruction	49.5%	51.33%
privileged_instruction	14.4%	6.33%
Window_overflow	12.5%	-
Window_underflow	23.5%	42.33%

The NETFI+ tool applied in fault injection by simulation takes a lot of time, an FPGA implementation based fault injection is done to validate a NETFI+ tool and to evaluate the results obtained by simulation.

The simulation based fault injection offers the analysis of the reliability of any circuit, such as a microprocessor pipeline or cache memory using all the types of faults. The term exhaustiveness can be done only by this technique, but it consumes a long time. Emulation based fault injection allows to inject a high number of faults by winning the time of the fault injection process.

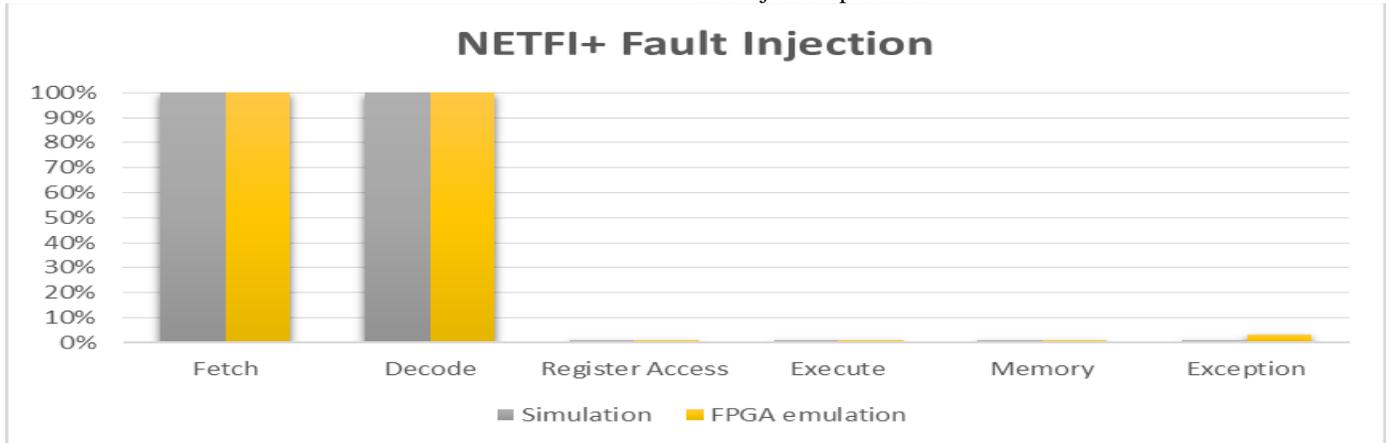


Fig. 6. NETFI+ analysis simulation versus FPGA emulation.

A NETFI+ tool is validated by simulation and FPGA implementation when a FF of FE and DE stage are 100% sensitive to SEU (“Fig. 6”). For the flip-flop non-sensitive of the other stages, the result remains the same except for the flip-flop of the XC stage when 0.02% is sensible to SEU.

The NETFI+ tool is limited in number of faults injected in FPGA emulation but it is faster than the fault injection by simulation. The analysis of the results obtained in simulation shows the benefits of the NETFI+ tool in simulation in its accuracy and the large number of faults which can be injected.

## VI. LEON3 SEU MITIGATION EVALUATION

### A. Principle of TMR

TMR is the most commonly used as a mitigation technique against SEUs for FPGA designs, used in radiation environments. The principle of TMR technique is done by triplicating a design and voting on the outputs of the three modules triplicated. TMR can be implemented on the latest commercial FPGA technologies, but it is costly in terms of area and power. It makes the circuit fault tolerant by masking and reducing the faults. It protects the design from errors propagated in LUT, internal state and control signals. “Fig. 7” shows the principle idea of the integer unit redundant with single voter.

The single voter with the triplicated logic will mask logic and errors created by SEUs. While two or three redundant copies of the design work correctly, errors will be masked and the output of the block will be correct [4].

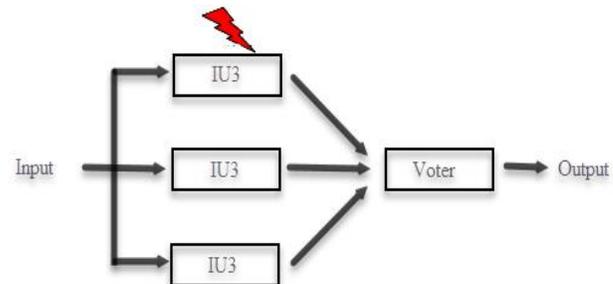


Fig. 7. TMR flow.

In [22], many approaches are used to detect a laser SEU faults for LEON3 on SRAM-based FPGA with the integration of several fault countermeasure techniques, the results obtained show that the modular triplication with single voter is the best one to mask errors. In [23], authors announced that a TMR presents a portable and robust solution.

TMR is generally used as a mitigation technique against a radiation fault injection. In [4], diverse repair techniques have been used to improve the SEUs mitigation for the LEON3 processor using two different approaches: Fault injection and Neutron radiation test. The results evince that using TMR with both CRAM (configuration memory) scrubbing and BRAM (internal block memory) scrubbing demonstrates that the reliability improvement is about 51.30x which used fault injection, and about 48.85x using Neutron radiation test.

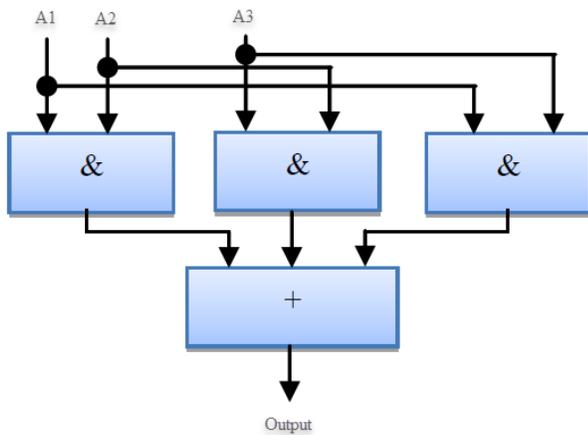


Fig. 8. Logic bloc of the voter.

The principal idea of TMR used in this work is by triplicating the integer unit of LEON3 and comparing the outputs with the golden operation (without countermeasure techniques) by voting the results of three redundant copies of the design to mask SEUs.

The voter is the important element in TMR technique. The importance of reliability in a majority voter is attributed to its application in both conventional fault-tolerant design and new Nano-electronic systems.

“Fig. 8” shows the logic bloc of the voter that masks faults in a single block of IU, such as A1, A2 and A3 which represent the output of the first, the second and the third copy of IU, respectively.

#### B. Analysis of SEU mitigation on LEON3’s Program Counter

In order to evaluate the behavior of LEON3 facing SEUs, fault injection is performed in the Program Counter register of the IU over its stage of pipeline. The proposed strategy of fault injection campaign NETFI+ is used and validated using two different campaign methods: simulation and FPGA emulation. The result obtained shows that 100% of FF in Fetch and Decode stage are sensitive to random SEUs fault injection by simulation and FPGA emulation. In goal to mitigate SEUs faults in the PC of LEON3, hardware integration of TMR fault tolerant technique in LEON3, by redundant the integer unit, is done. Only the bits of the PC register of Fetch and Decode stage will be used to mitigate de SEUs faults.

This section summarizes the TMR fault tolerant testing on LEON3 for SEUs fault injection. The SEU mitigation for the LEON3 processor on the program counter is shown in “Table 3”, two LEON3 design variations is shown in a table, the first design without TMR fault tolerant technique (unmitigated) and the second design with TMR.

The simulation of the design with TMR is made with success. Improvement is represented by a terms of sensitivity mitigation for fault injection.

The improvement in design sensitivity, according to the baseline design is enhanced when integrated a TMR module and injecting SEUs faults in one copy of IU, whose the TMR mitigates all the SEU injected to given about 100% of faults undetected.

TABLE III. SEUS FAULT INJECTION RESULTS WHICH INCLUDES TMR

Description	Unmitigated	One copy of IU affected	Two copies of IU affected	Three copies of IU affected
#faults injected	#18000	#18000	#36000	#108000
%sensitivity	100%	0%	0.1%	0.06%
#sensitive bits	#9000	#0	#2700	#3450
Improvement	1.00x	-	3.33x	2.60x

While the injection in three copies of IU demonstrates about 2.60x improvement over the injection within two copies of IU, the SEU mitigation technique used in this work, TMR, provides an important improvement in design sensitivity over the unmitigated baseline design. The percentage of sensitive bits within three copies of IU is about 3.19%, this is significant that the SEUs mitigation attained reaches about 96.81%.

The results reveal that using TMR in SRAM-based FPGA without scrubbing [5], the percentage of the sensitive bits is about 4.65% while in this paper using a TMR in IU, the sensitive bits represent 3.19% of the total bits.

#### VII. CONCLUSIONS AND PERSPECTIVES

In this paper, an extensive fault injection campaigns are done to evaluate the robustness of soft-core LEON3 processor against Single Event Upset. A new fault injection approach was improved in order to evaluate the susceptibility to soft errors, SEU, in LEON3’s program counter.

Two approaches to evaluate the sensitivity of integrated circuits to Single Event Upsets provoked by energetic particles present in the environment (space, Earth’s atmosphere) were explored. The first one is based on RTL simulations allowing evaluating IC sensitivity against SEU at early design phases, while the second one focuses on FPGA emulation which enables to obtain results closer to the ones of the hardware IC.

The accuracy/limitations of both approaches are studied by the analysis of experimental results. Fault injection based RTL simulation can be applied at early design phase, allowing fine-grained analysis also efficient local solution but it requires very important simulation time. For the fault injection based FPGA emulation is faster than fault injection based RTL simulation but it does not give detail information on reporting.

The results in this work put in evidence the importance of increasing the robustness of LEON3’s Program Counter register against soft-errors for critical applications. A hardware integration of a countermeasure unit is done in this work to give back the design fault tolerant. Analyzing the results shows that SEUs mitigation on the PC at the sensitive stage of pipeline, Fetch and Decode stages is improved about 99.979% using a repair technique, TMR.

Future work must be addressed to other types of faults like SET in the principal unit of LEON3, the integer unit, which including a SETs mitigation technique, TMR module.

REFERENCES

- [1] M. Murciano, M. Violante, "Validating the Dependability of Embedded Systems through Fault Injection by means of Loadable Kernel Modules," IEEE International High Level Design Validation and Test Workshop, 2007, pp 179-186.
- [2] A. Spilla, I. Polian, J. Müller, M. Lewis, V. Tomashevich, B. Becker, W. Burgard, "run-time Soft Error Injection and Testing of a Microprocessor using FPGAs," University of Freiburg, Jun 2013.
- [3] G.S. Rodrigues, F. Rosa, Á. Barros de Oliveira, F. Lima Kastensmidt, L. Ost, R. Reis, "Analyzing the Impact of Fault Tolerance Methods in ARM Processors under Soft Errors running Linux and Parallelization APIs," IEEE Transactions on Nuclear Science, 2017, pp. 1-7.
- [4] A M. Keller, M J. Wirthlin, "Benefits of Complementary SEU Mitigation for the LEON3 Soft Processor on SRAM-based FPGAs," IEEE Transactions on Nuclear Science, 2016, pp.519-528.
- [5] H. Abbasitabar, H. Zarandi, R. Salamat, "Susceptibility Analysis of LEON3 Embedded Processor Against Multiple Event Transients and Upsets," IEEE 15th International Conference on Computational Science and Engineering, 2012, pp. 548-553.
- [6] K. Chibani, M. Portolan, R. Leveugle. "Fast Register Criticality Evaluation in a SPARC Microprocessor," 10th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME), 2014, pp 1-4.
- [7] A. Kchaou, W. El Hadj Youssef, R. Tourki, F. Bouesse, P. Ramos, R. Velazco, "A deep Analysis of SEU Consequences in the Internal Memory of LEON3 processor," 17th IEEE Latin-American Test Symposium, 2016, pp 178.
- [8] S. Houssany, N. Guibbaud, A. Bougerol, R. Leveugle, F. Miller, N. Buard, "Microprocessor Soft Error Rate Prediction Based on Cache Memory Analysis," IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 59, NO. 4, 2012, pp 980-987.
- [9] H. Cho, S. Mirkhani, C-Y. Cher, J-A. Abraham, S. Mitra, "Quantitative Evaluation of Soft Error Injection Techniques for Robust System Design," 50th Annual Design Automation Conference, May 29 - June 07, 2013, ACM, Austin, TX, USA, pp: 1-10.
- [10] P. Peronnard, R. Ecoffet, M. Pignol, D. Bellin, R. Velazco, "Predicting the SEU Error Rate through Fault Injection for a Complex Microprocessor," IEEE International Symposium on Industrial Electronics, 2008, pp 2288-2292.
- [11] N. Alimi, Y. Lahbib, M. Machhout, R. Tourki, "An RTOS-based Fault Injection Simulator for Embedded Processors," International Journal of Advanced Computer Science and Applications, Vol. 8, No. 5, 2017, pp 300-306.
- [12] A. Clough, "Commercial Off-The-Shelf (COTS) Hardware and Software of Train Control Applications: System Safety Considerations," The national information service, Springfield, Virginia 22161. April 2003.
- [13] M. Portela-Garcia, C. Lopez-Ongil, M. Garcia Valderas, L. Entrena, "Fault Injection in Modern Microprocessors Using On-Chip Debugging Infrastructures," IEEE Transactions on dependable and secure computing, Vol.8, No.2; March-April 2011, pp 308-314.
- [14] J I. Gonzalez, A S. Cases, A M. Alvarez, S C. Asensi, H G. Miranda, M A. Aguirre, " Contrast of a HDL model and COTS version of a microprocessor for SOFT-ERROR Testing," 18th IEEE Latin American Test Symposium, 2017, pp. 1-6.
- [15] S. Ahmad, S. Sardar, B. Noor, A. Asar, "Analyzing Distributed Generation Impact on the Reliability of Electric Distribution Network," International Journal of Advanced Computer Science and Applications, Vol.7, No.10, 2016, pp 217-221.
- [16] W. Mansour, R. Velazco, "SEU fault-injection in VHDL-based processors: a case study," 13th Latin American Test Workshop (LATW), 2012, pp 1-5.
- [17] W. Mansour, R. Velazco, " An automated SEU fault-injection method and tool for HDL-Based Designs," IEEE Transactions on Nuclear Science, 60(4):2728-2733, Aug 2013.
- [18] M. Danek, L. Kafka, L. Kahout, J. Sykora, R. Bartosinski, "UTLEON3: Exploring Fine-Grain MultiThreading in FPGAs," Springer-Verlag New York, 2013, pp. 1-222.
- [19] J. Gaisler, M. Isomäki, "LEON3 GR-XC3S-1500 Template Design," Gaisler Research Based on GRLIB, October 2006.
- [20] "ModelSim Advanced Verification and Debugging," SE Tutorial, Version 6.0b, 15 November 2004.
- [21] Gaisler groups,"The SPARC Architecture Manual Version 8," SPARC International, 1991-1992.
- [22] F. Benevenuti, F L. Kastensmidt, " Evaluation of Fault Attack Detection on SRAM-based FPGAs," 18th IEEE Latin American Test Symposium, 2017, pp. 1-6.
- [23] P. Rech, C. Aguiar, C. Frost, and L. Carro, "An efficient and experimentally tuned software-based hardening strategy for matrix multiplication on gpus," IEEE Transactions on Nuclear Science, vol. 60, no. 4, Aug 2013, pp. 2797-2804.

# Deep Learning based Computer Aided Diagnosis System for Breast Mammograms

M. Arfan Jaffar

College of Computer and Information Sciences,  
Al Imam Mohammad Ibn Saud Islamic University (IMSIU),  
Riyadh, Saudi Arabia

**Abstract**—In this paper, a framework has been presented by using a combination of deep Convolutional Neural Network (CNN) with Support Vector Machine (SVM). Proposed method first perform preprocessing to resize the image so that it can be suitable for CNN and perform enhancement quality of the images can be enhanced. Deep Convolutional Neural Network (CNN) has been used for features extraction and classification with Support Vector Machine (SVM). Standard dataset MIAS and DDMS has been employed for testing the proposed framework by generating new images from these datasets by the process of augmentation. Different performance measures like Accuracy, Sensitivity, Specificity and area under the curve (AUC) has been employed as a quantitative measure and compared with state of the art existing methods. Results shows that proposed framework has attained accuracy 93.35% and 93% sensitivity.

**Keywords**—Classification; breast mammograms; computer aided diagnosis; deep learning

## I. INTRODUCTION

Cancer is the utmost precarious and foremost source of death in the entire world. It can effect different organs of human very badly and very fast. With the passage of time, it went to the stage where it is not possible to diagnose. It is possible to handle it at initial stages. Breast cancer is one of the conspicuous causes for deaths of women, and according to a survey and report published in 2016 statistics shows [1], that 61,000 new circumstances of breast cancer are prophesied. It has been well established and well reported in the literature that if breast cancer [2] is perceived at early stages by mammographic screening process then there are many chances to survive for these types of cancerous patients and even survival rate can be increased more than 90%. Therefore, for analysis and identification of breast cancer, the digital screening mammography [3] is broadly employed by radiologists. Computer aided diagnosis (CAD) is a process to diagnose automatically by using digital process by using some intelligent methods. In the literature, it has been reported that CAD systems has been used successfully for different types of cancer to diagnose at initial stages [4], [5]. It is very important to diagnose initially at early stages otherwise it is very difficult to handle it. Second the radiologist is very expensive in the whole world wide. Cancer is very dangerous disease so it is very important to have different opinions from different experts. It is very difficult for common person to consult more than one radiologist. Thus CAD can be used with radiologist as a second opinion [6]. Different intelligent techniques have

been used in the literature for CAD. Mostly existing methods used features extraction by extracting different types of features [7], [8]. These features include texture, statistical, shape or gradient based approaches. After extracting features, intelligent classifiers have been used. But the basic problem with these type of features is not robust. There is no standard way to check which features are most suitable for CAD and which classifier is most suitable. Thus features extraction is a big problem. To handle such type of issues, deep learning based methods has been proposed that used raw images as input [9], [10]. These deep learning approaches used raw images as input and extract and learn features directly from images. Then these features can be used for training and testing by fully connected deep layers of neural network. There are different models of deep learning are available like convolutional neural network (CNN) and deep belief network (DBN) [11], [12]. It has been presented in the literature that deep learning based methods perform well to solve pattern recognition problems specially in medical domain. Fig. 1 shows some images from DDMS dataset to show benign and malignant samples [13]. Table 1 show all abbreviation employed in this paper.

In the literature, it has been shown that in the previous decades, most of the researchers focus on some preprocessing and post-processing operation like segmentation, background removal, pectoral removal and some other methods for CAD [14]-[16]. These old techniques are now old and even required more processing time as well as some expert knowledge also required to highlight that which part is breast part and which part is not. Similarly, for post-processing some features extraction methods have been performed like texture features, some geometrical features, statistical features or some other features to classify images. Again there is no rule that which features are most suitable. Classification has been performed by using some old classifiers like artificial neural network, KNN, Bayesian [17], [18]. But most of the times, these features extraction and classifiers are not suitable. Even best features selection and classifiers selection is also a difficult task. With the concept of new advanced concept of deep learning, all these preprocessing and post-processing steps has been removed. Deep learning used original images as input, extract features automatically and learn it from direct pixels [19], [20]. It has been used in may real time applications and to solve complex problem. Basically, deep learning converts the original input images to features representation at different layers [21], [22]. These layers' work based upon some filters like edge detection filters or line detection. Most of the studies

in literature shows that the most suitable model is deep Convolutional Neural Network (CNN) that works just like Multi-Layer Perceptron (MLP) with some modifications. So in this paper, CNN has been used with SVM to extract features automatically and then classify by using deep concepts learning. Some preprocessing has been performed to resize the images. Most of the time the images used in mammography are 1024x1024 so it is very difficult to process it directly on such huge size of the image [23]. Therefore, resizing has been applied first. Then to smooth the images while retaining the edges preserved Bilateral filter has been employed with log transformation process. It enhances the quality of the images as shown in Fig. 2 and 3. The major contribution of the proposed method is to utilize the enhancement process by using hybrid bilateral filter with log transformation. Another one is to use deep convolutional neural network for the features extraction and classification.

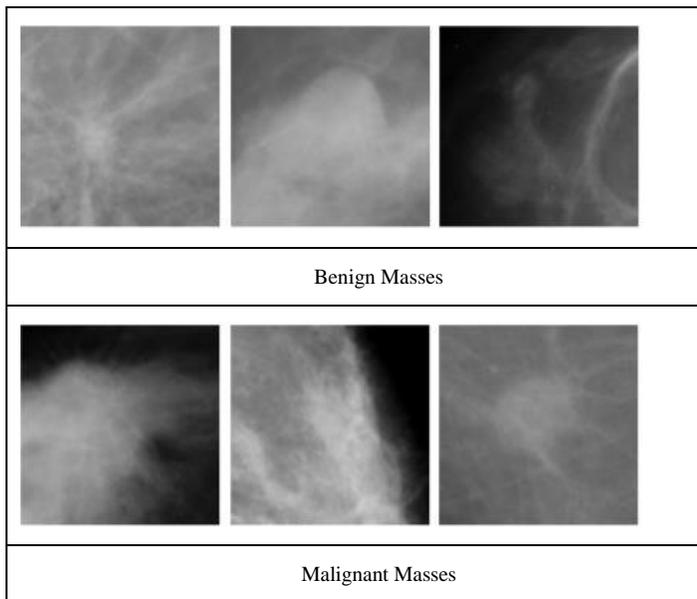


Fig. 1. Benign and Malignant Masses from DDMS dataset.

## II. PROPOSED METHOD

Deep convolutional neural network (CNN) has been employed in the proposed framework. The proposed framework used preprocessing step to enhance the quality of the images. Fig. 1 shows sample image taken from DDMS dataset and it clearly shows that this mammogram images are not high contrast. Some images are not clearly visible and good quality due to low resolution and sometimes due to conversion. To improve the quality of the images, enhancement process is required so that it can be improved and enhanced for good quality. Enhancement also improve quality so it can be more feasible for features extraction. Then, then extracted features are required to classify those regions. Deep CNN has been used with SVM for features extraction and classification. Details of all these phases has been given below in detail

### A. Dataset Description

We have used two standard datasets MIAS-mini [20] and DDMS [21] that are mostly used in the literature for breast mammograms classification and also augmented the images to

generate a large dataset that is suitable for deep convolutional neural network. Augmentation process has been explained in the preprocessing step in detail. In These both datasets are supervised that mean class labels of both datasets are given with data information. From MIAS dataset 322 images and from DDMS 1800 images has been selected to use in this study. Third dataset has been created artificially by augmentation process. For data augmentation rotation has been performed on both MIAS and DDMS images by 0, 45, 90, 135, 180, 225, 270, 315 and 360 degrees to generate huge dataset.

TABLE I. LIST OF ABBREVIATION USED IN THIS PAPER

Abbreviation	Complete Name
Log Transformation	Logarithmic Transformation
SVM	Support Vector Machine
Max Pooling	Maximum Pooling
CNN	Convolutional Neural Network
CAD	Computer Aided Diagnosis
MIAS	Mammographic image analysis society
DDMS	Digital database for screening mammography
AUC	Area under the receiver operating curve
ReLU	Rectified Liner Unit

### B. Preprocessing

After taking images from these datasets, it is required to enhance the visual quality of mammogram images. So a framework or procedure is required so that quality and visibility of the images can be enhanced. There are different enhancement methods available in the literature but most suitable enhancement method for mammograms is to improve the breast part in such a way that it also preserves the edges information. First, to decrease the size of mammograms, resizing has been applied so that it can be suitable for convolution neural network. So mammogram images are resized to suitable size 96x96 using a special technique bicubic interpolation. After resizing to proper size, bilateral filter employed on the images. Bilateral filter enhances the images but also preserve edges. We have tried many other methods based upon histogram but bilateral filter performs well as compare to other methods. So first original images processed to enhance the quality by incorporating bilateral filter [7]. This filter improves the image quality and make it smooth. Then logarithm transformation [8] employed on the filtered images returned by bilateral filter processed images. The resultant images displayed in Fig. 2. It improves the dynamic range especially for dark portion in the images as shown in Fig. 3.

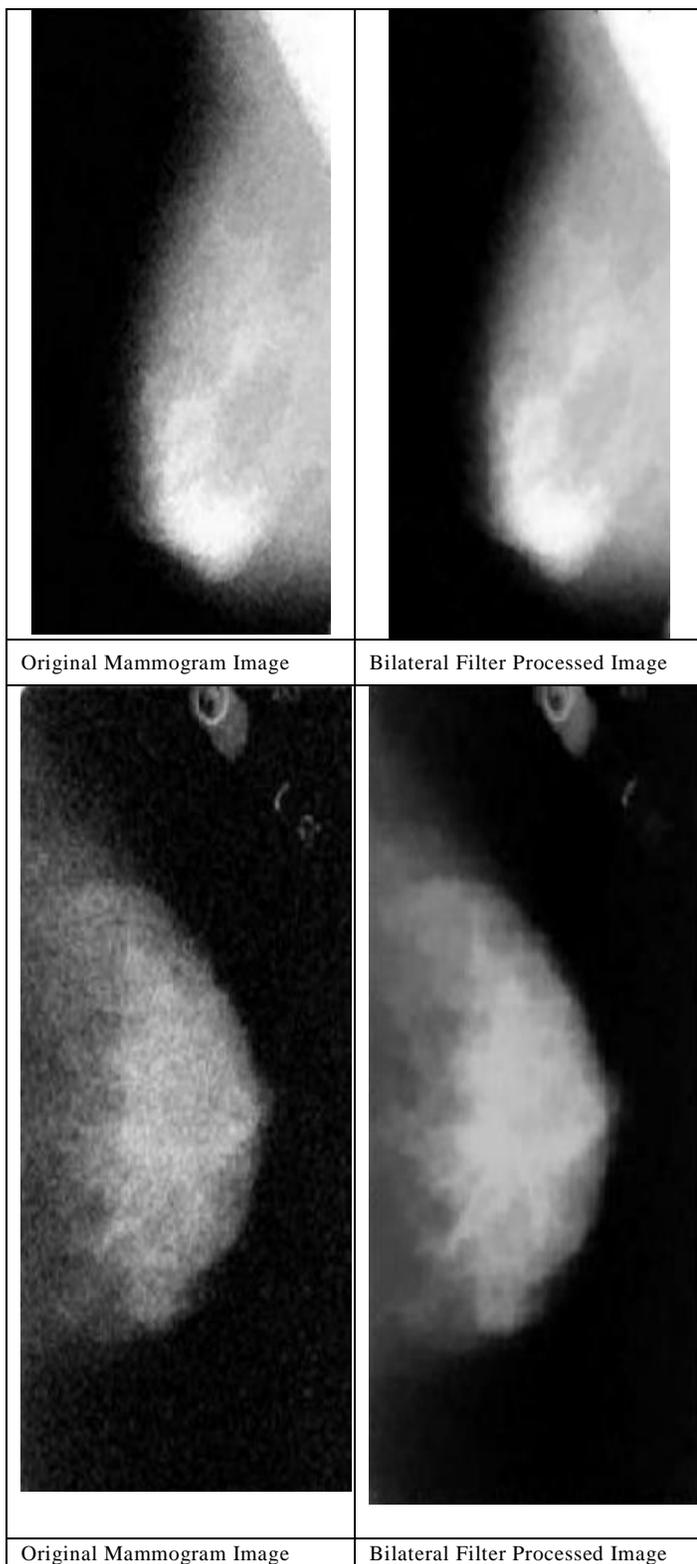


Fig. 2. Bilateral filter processed images results.

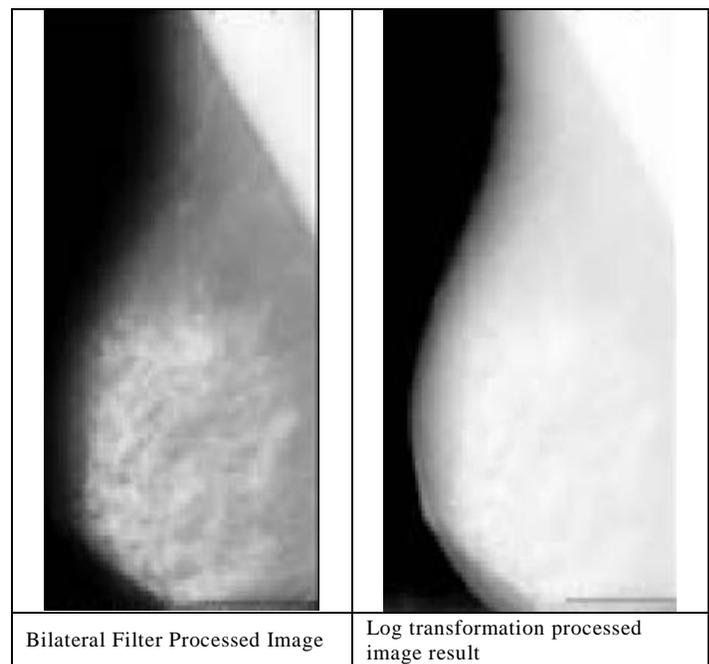


Fig. 3. Log Transformation processed images results.

Results of Bilateral filter and log transformation has been shown in Fig. 2 and 3. These images show clearly improvement by applying this process. At the end, to increase the training data, we created many images from existing data available in the datasets so that it can be suitable for convolutional neural network. For this purpose, rotation has been performed on all existing images at different angles by 0, 45, 90, 135, 180, 225, 270, 315 and 360. So in this way, a large training dataset has been created. Finally, we augmented the images to have a bigger training set. For data augmentation we rotated the images by 0, 45, 90, 135, 180, 225, 270, 315 and 360 degrees. After creating all these images, horizontal flip operation has been performed on all these artificially created images by rotation.

### C. Deep Convolutional Neural Network

For classification, feed-forward convolutional neural network has been applied for the classification of mammogram images. The network used in this paper has been shown in Fig.4. Three convolution layers has been used for experimentation. Three layers has three filters 46, 128 and 256 respectively and each filter has a size of 5x5. Rectified Liner Unit (ReLU) has been used after these filters as activation functions. After using two convolutional layers at first two layers, max pooling has been employed at the third layer and after third layer, quadrant pooling has been applied. After that, fully connected layers has been used with 400 hidden neurons and at the end softmax layer has been used for classification. Max pooling concept has been used for filtering the process at each layer as shown in Fig. 5 and this figure has been download from internet. Biases has not been used during experimentation. The major advantage of not using biases is that it increases the speed of the network so that it can be trained very efficiently and it also decreases the total amount of parameters that required learning during training [10].

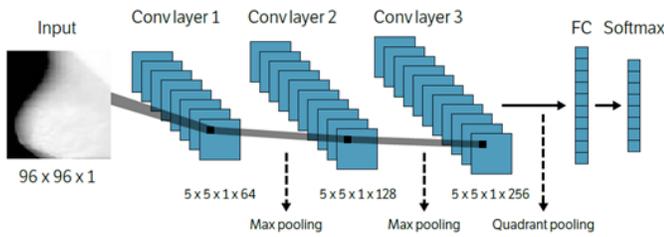


Fig. 4. Deep convolutional neural network model.

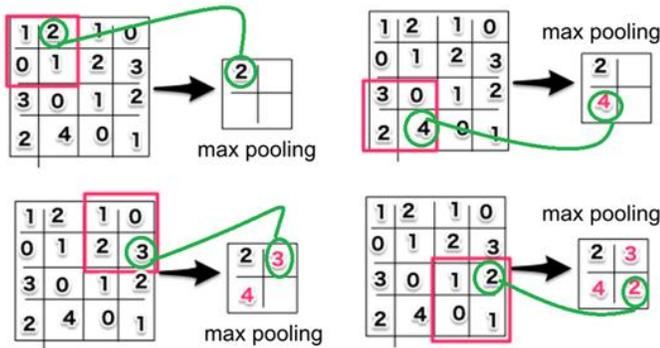


Fig. 5. Process of Max Pooling applied in CNN.

For training the network, stochastic gradient descent method has been used. Batch size for this stochastic gradient descent set to 64 with momentum 0.8. And decay parameter set to  $1e-5$ . Dropout mechanism also used to the fully-connected layer by setting probability to 0.5. Training has been performed by using the model of Anna software library. After training and features extraction along with corresponding labels of mammograms, it is used to train the Support Vector Machine (SVM) classifier. Training has been performed on different classifiers for each network. Finally, at the end, average has been used to calculate the average class score of all these different classifiers for the classification result as output.

### III. RESULTS AND DISCUSSION

Different performance measures have been used for the validation of proposed method. Classification results have been shown in the form of accuracy in Table 2. To calculate accuracy, a probability of 50% was used as threshold. For experimentation and testing, laptop with NVIDIA GTX 840M 2GB, 6 GB installed memory and Intel Core i7-4500U 1.8 GHz CPU has been used. MATLAB 2016b version has been for implementation and MatConvNet framework has been used for CNN. It takes around 120 minutes for training and for testing; it takes only 7 minutes for all test images.

Proposed method has been tested on large number of images taken from two standard datasets MIAS (322 images) and DDMS (1800 images) and also generated artificially by augmentation process explained in section 2.1 so total approximately 19000 image dataset has been generated by using this process. To generate the training and testing data, 10-fold cross validation process has been used with 50% training and 50% testing. Four different measures have been used to validate the results. Different performance measures like Accuracy, Sensitivity, Specificity and area under the

receiver operating curve (AUC) [23] has been employed to check the output of the proposed framework. These are the standard measures used for classification and specially AUC is the most common to assess the overall discrimination. Greater value of AUC shows good performance that should be close to 1. Proposed method has been compared with state-of-the-art methods that used the same concept like deep learning classification algorithms. The most common method used in this area are DeepCAD [17], CNN-Max-CAD-Qiu [15] and CNN-CAD-Jiao [16] by using same parameters and datasets. Table 2 shows results of the proposed method on different datasets like MIAS-mini, DDMS, combination set of MIAS-mini and DDMS and combination of MIAS-mini, DDMS and artificially generated datasets. This table shows the if CNN has been trained on large dataset then it also enhances the performance of the proposed method. It clearly shows that on new artificially generated dataset the accuracy is 93.72 percent that is maximum against all other datasets. Even other measures like sensitivity is 94.19, specificity is 92.24 and AUC is 0.93. All these values shows that it performs well on big dataset that contains approximately 19000 images.

TABLE II. DIFFERENT CLASSIFIERS RESULTS WITHOUT ENHANCEMENT

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
MIAS-mini	92.85	93.25	90.56	0.92
DDMS	93.02	92.84	91.35	0.91
MIAS-mini + DDMS	93.35	93.18	91.04	0.93
New Artificially generated dataset	93.72	94.19	92.24	0.93

TABLE III. COMPARISON OF PROPOSED METHOD WITH DIFFERENT EXISTING METHODS

Methodology	Dataset	Accuracy %	AUC %
Ball et al. [21]	DDMS	87	NA
Varela et al. [22]	DDMS	81	NA
DeepCAD [17]	DDMS + MIAS	91	0.91
CNN-Max-CAD-Qiu [15]	DDMS + MIAS	80	0.76
CNN-CAD-Jiao [16]	DDMS + MIAS	73	0.67
Proposed	DDMS + MIAS	93.35	0.93

Table 3 shows comparison of proposed method with other existing methods available in the literature. All these results for other methods has been taken reported in other research papers. The obtained results are reported in Tables 2 and 3. Table 3 results have been taken from paper [17] for DeepCAD, CNN-MAX-CAD-Qiu and CNN-CAD-Jiao and for Ball and Varela has been taken from [18]. Ball shows 87% accuracy on DDMS dataset, Varela shows 81% on DDMS dataset.

These methods have been not reported AUC in the paper therefore, I did not show results for these methods. DeepCAD has accuracy 91% and AUC 0.91 percent and author has tested on 600 ROI images collected from MIAS and DDMS. Author also shows results for other methods on the same collection of MIAS + DDMS dataset that CNN-Max-CAD-Qiu shows 80% and 0.76% accuracy and AUC accordingly and CNN-CAD-Jio shows 73% and 0.67% on the same collection of MIAS+DDMS dataset. Proposed method shows 93.35% accuracy and 0.93% AUC on combination dataset of MIAS-mini + DDMS. These values shows that proposed method works better as compare to all these existing methods. The proposed system does some preprocessing to enhance the performance of images like enhancement and resizing therefore it also performs well. Second, it has been trained on large set of images generated by artificially that is also helpful for good performance.

#### IV. CONCLUSION

In this paper, I have proposed a computer aided diagnosis system based upon deep convolutional neural network. There are two different tasks that has been applied. In the first task, breast preprocessing has been performed to make it suitable for CNN. During preprocessing, resizing has been performed on the images so that it can be suitable for CNN. Then dataset has been increase so that it is suitable for CNN. Different rotations have been used to generate artificially data by using the same data available in MIAS and DDMS dataset. After that, enhancement has been performed by using Bilateral filter and log transformation so that images can be smooth while preserving edges. After that CNN model has been used for features extraction classification. Results shows that it has improved the performance of CAD system by incorporating CNN. Due to these contributions, proposed system perform well. In the future, some other model of deep learning will be explored.

#### REFERENCES

- [1] Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2016. *CA Cancer J. Clin.* 2016, 66, 7–30.
- [2] Kalager, M.; Zelen, M.; Langmark, F.; Adami, H.O. Effect of screening mammography on breast-cancer mortality in Norway. *N. Engl. J. Med.* 2010, 363, 1203–1210.
- [3] Posso, M.; Carles, M.; Rué, M.; Puig, T.; Bonfill, X. Cost-Effectiveness of double reading versus single reading of mammograms in a breast cancer screening programme. *PLoS ONE* 2016, 11, 1–13.
- [4] Bird, R.E., Professional Quality assurance for mammography screening programs, *Radiology* 177, pp.587. 1990.
- [5] Bellotti, R., De Carlo, F., Tangaro, S., Gargano, G., Maggipinto, G., Castellano, M., ... De Nunzio, G. (2006). A completely automated CAD system for mass detection in a large mammographic database. *Medical Physics*, 33(8), 3066–3075.
- [6] Varela, C., Tahoces, P. G., Méndez, A. J., Souto, M., & Vidal, J. J. (2007). Computerized detection of breast masses in digitized mammograms. *Computers in Biology and Medicine*, 37(2), 214–226.
- [7] F. Sahba, A. Venetsanopoulos, Breast mass detection using bilateral filter and mean shift based clustering, in: *Proceedings of the 2010 International Conference on Signal Processing and Multimedia Applications (SIGMAP)*, 2010, pp. 88–94.
- [8] Urvashi Manikpuri, Yojana Yadav, Image Enhancement Through Logarithmic Transformation, *International Journal of Innovative Research in Advanced Engineering (IJRAE) ISSN: 2349-2163 Volume 1 Issue 8, 2014.*
- [9] Wei, D., Chan, H. P., Helvie, M. a, Sahiner, B., Petrick, N., Adler, D. D., & Goodsitt, M. M. (1995). Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis. *Medical Physics*.
- [10] Pooya Khorrami, Tom Le Paine, Thomas S. Huang, Do Deep Neural Networks Learn Facial Action Units When Doing Expression Recognition?, *Computer Vision and Pattern Recognition*, 2017 arXiv:1510.02969.
- [11] T.L. Economopoulos, P.A. Asvestas, G.K. Matsopoulos, Contrast enhancement of images using Partitioned Iterated Function Systems, *Image and Vision Computing Volume 28, Issue 1, January 2010, Pages 45–54.*
- [12] Freund Y, Schapire RE, Experiments with a new boosting algorithm. In: *Machine learning: proceedings of the thirteenth international conference*, 1996, pp 325–332.
- [13] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [14] Lior Rokach, Ensemble-based classifiers, *Artificial Intelligence Reviews* (2010) 33:1–39.
- [15] Qiu, Y.; Wang, Y.; Yan, S.; Tan, M.; Cheng, S.; Liu, H.; Zheng, B. An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology. In *Proceedings of the SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis*, San Diego, CA, USA, 24 March 2016.
- [16] Jiao, Z.; Gao, X.; Wang, Y.; Li, J. A deep feature based framework for breast masses classification. *Neurocomputing* 2016, 197, 221–231.
- [17] Qaisar Abbas, DeepCAD: A Computer-Aided Diagnosis System for Mammographic Masses Using Deep Invariant Features, *Computers*, 2016, 5, 28; doi:10.3390/computers5040028.
- [18] Wentao Zhu, Qi Lou, Yeeleng Scott Vang, Xiaohui Xie, Deep Multi-instance Networks with Sparse Label Assignment for Whole Mammogram Classification, arXiv preprint arXiv:1612.05968 (2016).
- [19] Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S.; et al. The mammographic image analysis society digital mammogram database. In *Excerpta Medica: International Congress Series*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 1069, pp. 375–378. Available online: <http://peipa.essex.ac.uk/info/mias.html> (accessed on 28 October 2016).
- [20] Heath, M.; Bowyer, K.W.; Kopans, D. The digital database for screening mammography. In *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, ON, Canada, 11–14 June 2000.
- [21] Ball, J.E., Bruce, L.M.: Digital mammographic computer aided diagnosis (cad) using adaptive level set segmentation. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 4973–4978. IEEE (2007).
- [22] Varela, C., Timp, S., Karssemeijer, N.: Use of border information in the classification of mammographic masses. *Physics in Medicine and Biology* 51(2), 425 (2006).
- [23] Metz, C.; Herman, B.; Shen, J. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Stat. Med.* 1998, 17, 1033–1053.

# ODSA: A Novel Ordering Divisional Scheduling Algorithm for Modern Operating Systems

Junaid Haseeb

Department of Computer Science  
University of Management and  
Technology, Sialkot, Pakistan

Khizar Hameed

Department of Computer Science  
University of Management and  
Technology, Sialkot, Pakistan

Muhammad Junaid

Department of Computer Science  
University of Management and  
Technology, Sialkot, Pakistan

Muhammad Tayyab

Department of Computer Science  
University of Management and  
Technology, Sialkot, Pakistan

Samia Rehman

Department of Computer Science  
COMSATS Institute of Information  
Technology, Islamabad, Pakistan

Agha Muhammad Musa Khan

Department of Computer Science  
University of Management and  
Technology, Sialkot, Pakistan

**Abstract**—CPU scheduling is defined as scheduling multiple processes that are required to be executed in a specific time period. A large number of scheduling algorithms have been proposed to achieve maximum CPU utilization/throughput and minimizing turn around, waiting and response time. Existing studies claim that Round Robin (RR) is providing best results in terms of above-mentioned factors. In RR, a process is assigned to CPU for a fixed time quantum then the process starts its execution, in case that assigned time quantum greater than CPU's capacity then remaining section of that process waits for its next turn. Although RR schedules processes in an efficient manner, however, it has certain limitations such as if time quantum is too small or large, it causes frequent context switching and response time can increase. To address these identified problems, various improved versions of RR also exist. The purpose of this paper is twofold: 1) a comparison between different improved versions of RR; and 2) a new algorithm named Ordering Divisional Scheduling Algorithm (ODSA) is also proposed that combines various features of different algorithms and is actually an improvement to RR. Our results show that ODSA can schedule processes with less turn around and average waiting time as compared to existing solutions.

**Keywords**—CPU scheduling; round robin scheduling algorithm; turnaround time; waiting time; context switching

## I. INTRODUCTION

Operating System (OS) is an essential part of a computer system that acts as an intermediary between input commands and hardware. Among various functions performed by OS, one is processed scheduling, as Central Processing Unit (CPU) has to manage concurrently executing processes. Some processes are concerned with OS and others are originated by users. For the execution, each process requires specific time duration of CPU. Required execution time is totally dependent on the type of process to be executed; it may fall into the category of engaging CPU's resources for a long time or either short. In the

context where multiple processes are available in the ready state against only one CPU than OS has to decide which process needs to be executed first. For this purpose, many scheduling algorithms have been proposed and this management of ordering processes is known as process scheduling [1]. These proposed algorithms have been designed with various goals such as better utilization of CPU's resources, less turnaround time, waiting and response time of processes.

CPU plays a vital role in the execution of processes as it has to assign required resources by OS to a specific process. In the case of multiple processes to be executed, scheduling of processes requires a careful and must ensure fairness so that process starvation is minimized [2]. Scheduling process can be performed using software like scheduler or dispatcher [3]. Round Robin (RR) is most commonly known algorithm that helps in scheduling processes for OS [1], [4], [5]. In RR, CPU splits OS's time into multiple slices that are known as time quantum. Then these time intervals are assigned to processes so that their execution over OS can be performed. RR scheduler is mainly concerned with following dimensions.

- **CPU Utilization** – By keeping the CPU as busy as possible.
- **Throughput** – Number of tasks completed in unit time.
- **Turnaround** – Time required completing a job after the submission.
- **Waiting Time** – Time required waiting in a ready queue.
- **Response Time** – Time required to response a particular job after the submission.
- **Fairness** – Time is given by CPU to each thread.

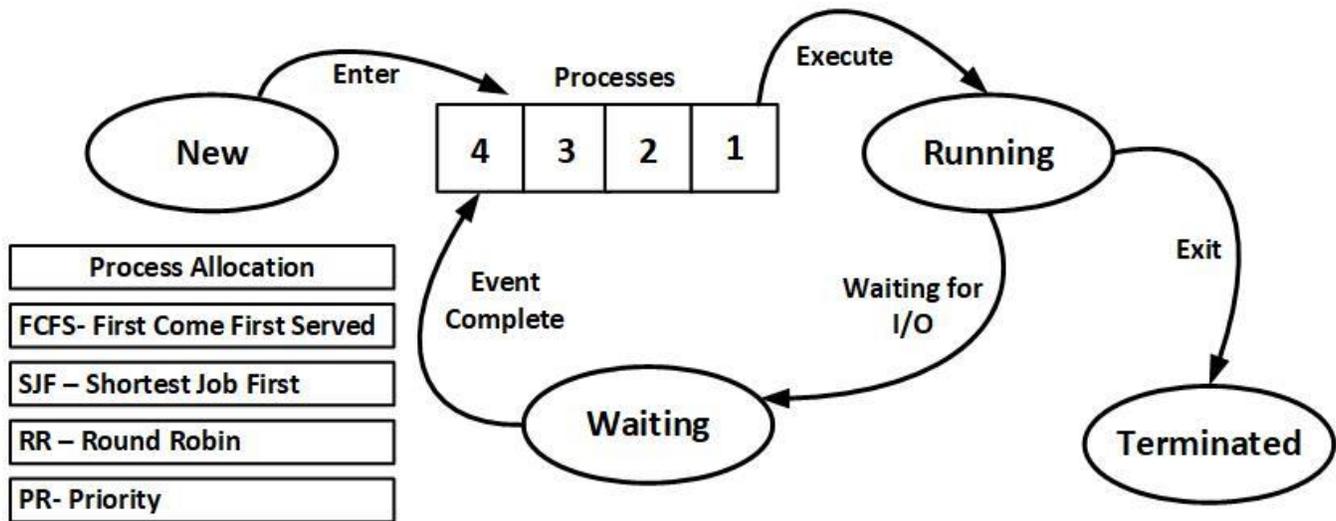


Fig. 1. Process creation and allocation in OS.

Another scheduling algorithm is First-come, first served (FCFS) [4]. In FCFS scheduling algorithm, the process request the CPU and CPU in return execute the process in same order. A single queue is maintained for ready processes in this algorithm. It's a non-preemptive scheduling algorithm i.e. once the CPU has been allocated to a process, that process keeps the CPU until it releases the CPU, either by terminating or by requesting I/O. The next algorithm is SJF [5]. In Shortest Job First (SJF), advanced knowledge of time taken by the processes is required. The process having less time is executed by CPU prior to that having a large amount of time. Another is priority based algorithm [6]. In priority based algorithm, every process is assigned a fixed priority by OS and the scheduler arranges the processes in the ready queue in order of their priority. Fig. 1 briefly explains the process creation and allocation procedure for all the above-mentioned algorithms.

Although these scheduling algorithms consider performance parameters very well but still have definite problems. In FCFS, when one process is completed than CPU switches to another process, therefore, scheduling overhead is minimal and no reorganization of the process queue is required [6], [7]. The FCFS algorithm has low throughput as the process executes in the same order as they come and there is a possibility that long processes hold the CPU for a long duration. Resultantly turnaround time, waiting time and response time can be high [7], [8]. In SJF there is an additional context switching if a shorter process arrives during another process execution. This halts the currently running process, execute the shortest job and then resume the previous one. This creates additional overhead. The algorithm is giving maximum throughput [4], [9]. The major flaw is starvation which occurs when there are a large number of processes are being run by CPU. In priority based algorithm, overhead is not minimal. Waiting and response time is interlinked with priority. Higher priority processes have smaller waiting and response times. In [5] the starvation problem also exists. In RR scheduling algorithm, if time quantum is too small it gives extensive overhead. In addition, average response time, waiting time is dependent on a number of processes, its length and value of

time quantum. Starvation has been reduced almost to zero [2], [10].

As mentioned earlier that RR is considered as one the most commonly used algorithm. In order to overcome its concerned problems, various improvements have been made in RR algorithm [7]. However, which algorithm is giving maximum result is still questionable. In this paper, we provide an overview of improvements made up till now in RR using simulations. We also propose new algorithm Ordering Divisional Scheduling Algorithm (ODSA) to achieve maximum performance in terms of scheduling criteria.

Rest of the paper is structured as follows. Section II provides a detailed discussion about existing similar works done in this field like PBDRR [11], Time quantum using fuzzy logic [10], an improved RR algorithm [12] and dynamic quantum [5], [13]. Section III presents the analysis & experiment discussion along with results. Section IV contains the detail of proposed algorithm. In section V, comparison results are discussed with rest of the algorithms. Section VI provides conclusion and future implications.

## II. RELATED WORK

Round Robin has become one of the most important and widely used scheduling algorithms, despite its problems such as fixed quantum size [2], [14]. As Round Robin (RR) is used almost in all like in Windows, UNIX, BSD etc., So, to overcome its shortcomings, many types of research have been carried out [2], [12], [13], [15].

Mohanty et al. [3] proposed priority based Dynamic Round Robin algorithm to combine the dynamic time quantum and priority based selection of processes. In this way, time slice for each process is calculated and it changes after every round of execution. In [7], an algorithm was designed which take input sequence and assign priority, and then it sets the value of original time slice (OTS). The components of priority are calculated using short components [8]. In first round for processes having SC as 1, assign time quantum same as intelligent time slice whereas the processes having SC as 0

given the time quantum equal to the ceiling of the half of the intelligent time slice. In next round, the processes having SC as 1 assign double the time slice of its previous round whereas the processes with SC equal to 0 given the time quantum equal to the sum of the previous time quantum and ceiling of the half of the previous time quantum. After various examples, it reflected that as time quantum is dynamic therefore reducing no of context switching, average waiting and response time [3].

Alam et al., [1] used the concept of fuzzy logic to determine time quantum. Fuzzy logic is basically an extension of Boolean logic dealing with the concept of partial truth that denotes the degree of truth. In real everything, it can be expressed in binary terms. He used fuzzy interface system for finding the time quantum [16]. The fuzzy inference system accepts two numbers as input and produce only single number as output. The input numbers specify as the total number of processes resides in the ready queue and average burst time of processes [17]. Time quantum is the fixed output value generated by that system. The advantage of fuzzy logic is that each process in the system assigned a fixed time quantum according to average burst time. In addition, the performance of the system is not declined due to gratuitously context switches[1].

Mohanty et al., [18] proposed new algorithm as a combination of the shortest job first (SJF) and RR. In the first process, it is assigned to CPU using Round Robin scheduling algorithm. In a second step, it selects the shortest job from the waiting queue and it shortest job assign to the CPU [9]-[14], [19]. The process will be terminated after the successful execution of all the processes. Another improvement is the combination of SJF and RR algorithms [1]. This algorithm runs as normal RR in the first cycle and then selects the SJF from waiting for the queue and so on. From a number of experiments present in this paper, it is obvious that total waiting time and average turnaround time both are reduced [16], [17], [20]. The reduction of total waiting time and turnaround time shows maximum CPU utilization and minimum response time [18].

Noon et. al [2] proposed the new concept of time quantum which is based on dynamic allocation of time to processes. The given time quantum allocated to process on their burst time. To solve the problem of time quantum, AN algorithm is proposed that adjusts the time quantum of processes which resides in ready queue. Experimental evaluations claimed that the efficiency of AN algorithm is higher than the existing RR algorithm[10], [16]. In context of efficiency, the dynamic scheduling algorithm is reliable and scalable for wide variety of OS's as it provides the support of self- adaptation OS. The self-adaptation property automatically fills the requirement of end user [17].

Above discussion can be summarized as many improvements to RR have been already proposed. One question that is important yet not answered is which algorithm provides better results in terms of CPU performance. To answer this question, we have performed simulations of various algorithms. Further discussion about experimental setup, selected algorithms and criterion for declaring best algorithm is part of section III and results of our simulation are discussed in key findings section.

TABLE I. PROCESSES ALONG WITH BURST TIME AND PRIORITY

PROCESS ID	BURST TIME (ms)
P0	12
P1	49
P2	20
P3	60
P4	30
P5	9

TABLE II. PROCESSES ALONG WITH BURST TIME AND PRIORITY

PROCESS ID	BURST TIME (ms)
P0	10
P1	2
P2	1
P3	15

### III. EXPERIMENTAL SETUP AND ANALYSIS

In this section of the paper, different improved versions of RR are compared to answer the question of the best scheduling algorithm. Three most common known algorithms (2, 3, and 4) are selected and best performance criterion is based on turnaround and average waiting time. Turnaround time of a process is defined as the time a process has to wait for getting its turn so that its remaining execution can be completed and the waiting time of a process is the time it has to wait in queue before going into execution mode. To perform the simulations, we have taken two examples. Details about processes included in examples along with burst time are presented in Tables 1 and 2 and fixed time quantum respectively and simulation results according to already selected algorithms are as follows.

We have tested three algorithms on the above two examples. Each algorithm has the output in term of average Turnaround time and the average waiting time. In these examples, the overall waiting time for a process is too high in such away the process has to wait for a log until its execution time. To overcome these problems, we have proposed a new algorithm named a novel: Ordering Divisional Scheduling Algorithm (ODSA), in which we have achieved very low throughput time as well as very low waiting time for a process to be executed. In the following section IV, detailed information ODSA is provided.

### IV. ORDERING DIVISIONAL SCHEDULING ALGORITHM (ODSA)

A new algorithm Ordering Divisional Scheduling Algorithm (ODSA) is designed providing better efficiency as compared to all discussed algorithms and also overcoming the problems mentioned in introduction part. As the name depicts it first arrange the processes in ascending order based on their burst time and then divide the processes into two halves. Complete steps involved in this algorithm are as follows:

- First, arrange the processes as per ascending order of burst time.
- Then divide the processes into two halves (In case of odd values use ceiling in first half).
- First, run the process from second half having shortest burst time.
- Then run all the first half processes using SJF algorithm.
- At the end run all the remaining processes of second half using RR algorithm.

After dividing when we run the process from second half then using two different algorithms i.e. SJF and RR this will reduce the processes waiting time and response time hence improving the efficiency.

**Algorithm 1 ODSA**

```

1: procedure ODSA
2: Input: Processes,
3: initialization Processes = φ
4: ODSA ( )
5:  Foreach (Processes)
6:     processes = Order(processes)    ▷ Ascending order of Burst
6:                                            ▷ time of each processes
7:     Split(processes') HalfA ^ HalfB    ▷ Splits the processes
7:                                            into two halves
8:  Foreach (HalfB)
9:     ExecutedProcesses [] = B.processesExecute (BurstHalfB)
9:                                            ▷ As per shorted Burst time
10:  Foreach (HalfA)
11:     ExecutedProcesses [] = A.processesExecute (SJFHalfA)
11:                                            ▷ As per SJF
12:  If (A.processes ∉ ExecutedProcesses [] && B.processesExecute ∉ executed[])
13:     ExecuteProcesses [] = Execute(RRHalfA&B)    ▷ As per RR
13:                                            remaining processes
14: end procedure
    
```

**A. Verification of ODSA**

To validate our proposed algorithm, we have again used example 1 and 2. This time these examples are scheduled according to newly proposed ODSA algorithm. Results are presented as follows.

**1) Step – I & II:**

Arrange the processes as per ascending order of burst time and divide into two halves as presented in Table 3.

TABLE III. PROCESSES AFTER ORDERING AND DIVIDING

PROCESS ID	BURST TIME (ms)
P5	9
P0	12
P2	20
P4	30
P1	49
P3	60

**2) Step – III, IV & V:**

- First, run the process from second half having shortest burst time i.e. P4.

- Then run all the first half processes using SJF algorithm.
- At the end run the remaining process of second half using RR algorithm.

Results of example 3 after scheduling processes according to ODSA are as following:

**3) Step – I & II**

- Arrange the processes as per ascending order of burst time and divide into two halves as presented in Table 4.

TABLE IV. PROCESSES AFTER ORDERING AND DIVIDING

PROCESS ID	BURST TIME(ms)
P2	1
P1	2
P0	10
P3	15

**4) Step – III, IV & V**

- First, run the process from second half having shortest burst time i.e. P0.
- Then run all the first half processes using SJF algorithm.
- At the end run the remaining process of second half using RR algorithm.

**V. RESULTS – COMPARISON**

Results obtained from simulations are compared with existing algorithms are available in Tables 5 and 6.

It reveals that newly designed algorithm has less turnaround as well as waiting time as compared to all the improvements made in RR algorithm up till now. The number of context switches is also less hence ODSA improves the efficiency and throughput.

TABLE V. COMPARISONS OF ODSA WITH EXISTING ALGORITHM USING EXAMPLE 3

Scheduling Criteria	An Improved RR algorithm	AN algorithm	PBDRR	ODSA
Turnaround time	99.33	119.33	106.16	85.1
Waiting time	69.33	89.33	76.16	55.1
Context Switching	11	10	11	8

TABLE VI. COMPARISONS OF ODSA WITH EXISTING ALGORITHM USING EXAMPLE 4

Scheduling Criteria	An Improved RR algorithm	AN algorithm	PBDRR	ODSA
Turnaround time	17.5	17.25	23	15.5
Waiting time	10.5	10.25	18	8.5
Context Switching	6	6	6	5

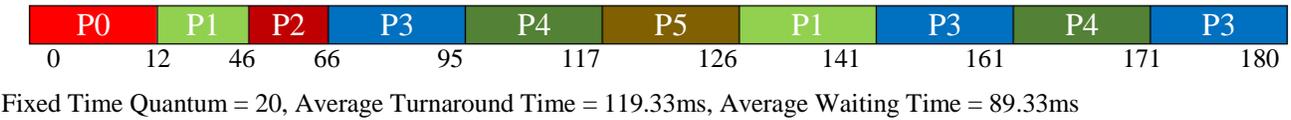
ODSA is showing less turnaround / waiting time and also less number of context switching. Graphical comparison of ODSA with all discussed algorithm is shown in Fig. 2.

**Example 1**

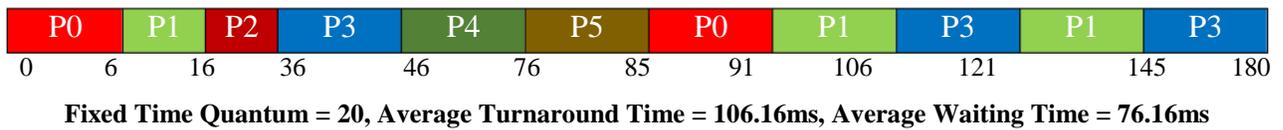
An improved RR algorithm [4]



A New Round Robin Algorithm [2]



PBDRR algorithm [3]

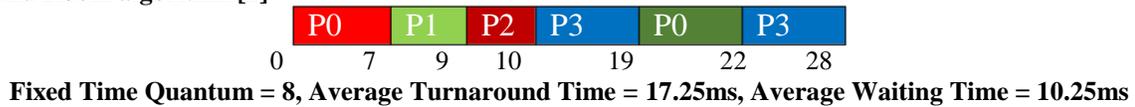


**Example 2:**

An improved RR algorithm [4]



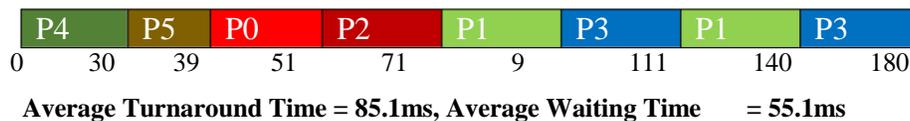
A New Round Robin algorithm [2]



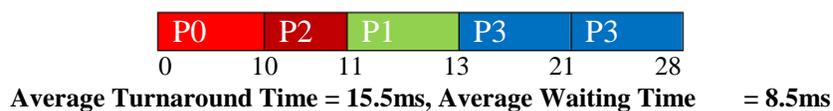
PBDRR algorithm [3]



**Example 3:**



**Example 4:**



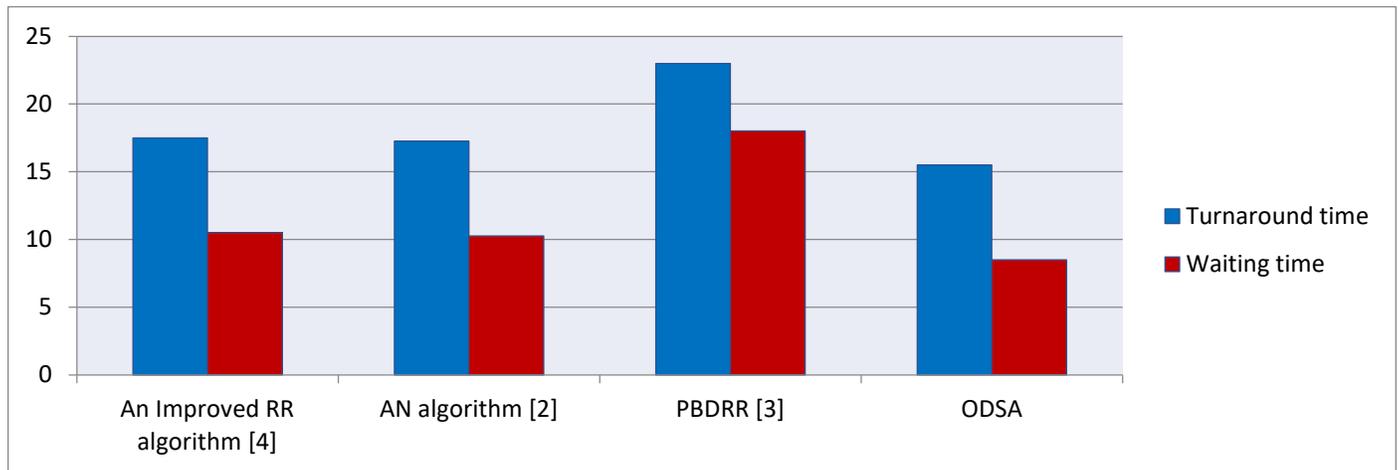


Fig. 2. Comparison of proposed and existing algorithms.

## VI. CONCLUSION

Process scheduling is important in a multiprogramming environment. Various algorithms have been designed so far for better process management. Different researchers have made improvements in RR algorithm to overcome the problems like starvation, context switching etc. ODSA has been proposed in this paper which fairly allocates the resources to CPU. This algorithm is improving the CPU scheduling criteria and processes turnaround time is very less. It is also improved the efficiency in terms of response time and no of context switches. Deficiencies in existing algorithms like context switching, starvation, convoy effect etc. have also overcome by ODSA. Therefore, we can say efficiency of proposed scheme in terms of CPU scheduling criteria is better than all the existing improvements made till now in RR algorithm. In future, we are interested to enhance ODSA by considering more performance parameters. Comparing ODSA with other scheduling algorithms in terms of energy utilization and management of hardware resources are future implications.

## REFERENCES

- [1] B. Alam, M. N. Doja, and R. Biswas, "Finding Time Quantum of Round Robin CPU Scheduling Algorithm Using Fuzzy Logic," IEEE International Conference on Computer and Electrical Engineering (ICCEE), pp. 795–798, 2008.
- [2] A. Noon, A. Kalakech, and S. Kadry, "A New Round Robin Based Scheduling Algorithm for Operating Systems : Dynamic Quantum Using the Mean Average," IJCSI International Journal of Computer Science Issues, vol. 8, no. 3, pp. 224–229, 2011.
- [3] P. R. Mohanty, P. H. S. Behera, K. Patwari, M. Dash, and M. L. Prasanna, "Priority Based Dynamic Round Robin ( PBDRR ) Algorithm with Intelligent Time Slice for Soft Real Time Systems," vol. 2, no. 2, pp. 46–50, 2011.
- [4] W. Stallings, "Operating Systems: Internals and Design Principles," Pearson, vol. 8, 2014.
- [5] L.-X. Wang, "A Course in Fuzzy Systems and Control," Prentice-Hall, Inc., 1997.
- [6] A. S. Tanenbaum, A. S., & Woodhull, "Operating systems: design and implementation," Englewood Cliffs, NJ Prentice-Hall, vol. 2, 2006.
- [7] S. Mamdani, E. H., & Assilian, "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller," International Journal of Man-Machine Studies, vol. 7, no. 1, pp. 1–13, 1975.
- [8] M. Sugeno, "Industrial applications of fuzzy control.," Elsevier Sci. Inc., 1985.
- [9] J. R. Jang, "ANFIS : Adaptive Network Based Fuzzy Inference System," IEEE transactions on systems, man, and cybernetics, vol. 23, no. 3, 1993.
- [10] D. J. Simon, "Training fuzzy systems with the extended Kalman filter," Fuzzy sets and systems, vol. 132, no. 2, pp. 189–199, 2002.
- [11] W. Tong and J. Zhao, "Quantum Varying Deficit Round Robin Scheduling Over Priority Queues," IEEE International Conference on Computational Intelligence and Security, pp. 252–256, 2007.
- [12] R. J. Matameh, "Self-adjustment time quantum in round robin algorithm depending on burst time of the now running processes," American Journal of Applied Sciences, vol. 6, no. 10, pp. 1831–1837, 2009.
- [13] T. Helmy, "Burst Round Robin as a Proportional-Share Scheduling Algorithm," IEEE International Conference on Grid and Cooperative Computing, 2007.
- [14] S. M. Mostafa, S. Z. Rida, and S. H. Hamad, "Finding time quantum of round robin CPU scheduling algorithm in general computing systems using integer programming." International Journal of Research and Reviews in Applied Sciences (IJRRAS), vol. 5, no. 1, pp. 64–71, 2010.
- [15] Liu, C. L., & Layland, J. W "Scheduling Algorithms for Multiprogramming in a Hard- Real-Time Environment Scheduling Algorithms for Multiprogramming," Journal of the ACM (JACM), vol.20, no. 1, pp. 46–61, 1973.
- [16] H. S. Behera, "A New Proposed Dynamic Quantum with Re-Adjusted Round Robin Scheduling Algorithm and Its Performance Analysis," International Journal of Computer Applications, vol. 5, no. 5, pp. 10–15, 2010.
- [17] Yang, L., Schopf, J. M., & Foster, I, "Conservative scheduling: Using predicted variance to improve scheduling decisions in dynamic environments" In Proceedings of the 2003 ACM/IEEE conference on Supercomputing, pp. 1–16, 2003.
- [18] R. K. Yadav, A. K. Mishra, N. Prakash, and H. Sharma, "An Improved Round Robin Scheduling Algorithm for CPU scheduling ," International Journal on Computer Science and Engineering vol. 2, no. 4, pp. 1064–1066, 2010.
- [19] S. Suranawarat, "A CPU Scheduling Algorithm Simulator," IEEE International Conference on Frontiers In Education Conference-Global Engineering: Knowledge Without Borders, Opportunities Without Passports, pp. 19–24, 2007.
- [20] P. R. Mohanty, P. H. S. Behera, K. Patwari, M. R. Das, and M. Dash, "Design and Performance Evaluation of a New Proposed Shortest Remaining Burst Round Robin (SRBRR) Scheduling Algorithm," In Proceedings of International Symposium on Computer Engineering & Technology (ISCET), Vol. 17, pp. 126-137.

# Customized Descriptor for Various Obstacles Detection in Road Scene

Haythem Ameer

Laboratory of Micro-Optoelectronics and Nanostructures  
(LMON), University of Monastir,  
Monastir, Tunisia

Abdelhamid Helali

Laboratory of Micro-Optoelectronics and Nanostructures  
(LMON), University of Monastir  
Monastir, Tunisia

J. Ramírez

2Dept. of Signal Theory, Networking and Communications,  
University of Granada,  
Granada, Spain

J. M. Gorriz

2Dept. of Signal Theory, Networking and Communications,  
University of Granada,  
Granada, Spain

Ridha Mghaieth

Laboratory of Micro-Optoelectronics and Nanostructures  
(LMON), University of Monastir  
Monastir, Tunisia

Hassen Maaref

Laboratory of Micro-Optoelectronics and Nanostructures  
(LMON), University of Monastir  
Monastir, Tunisia

**Abstract**—Recently, real-time object detection systems have become a major challenge in the smart vehicle. In this work, we aim to increase both pedestrian and driver safety through improving their recognition rate in the vehicle's embedded vision systems. Based on the Histogram of Oriented Gradients (HOG) descriptor, an optimized object detection system is presented in order to achieve an efficient recognition system for several obstacles. The main idea is to customize the weight of each bin in the HOG-feature vector according to its contribution in the description process of the extracted relevant features. Performance studies using a linear SVM classifier prove the efficiency of our approach. Indeed, based on the INRIA datasets, we have improved the sensitivity rate of the pedestrian detection by 11% and the vehicle detection by 5%.

**Keywords**—ADAS; customized HOG; linear SVM; obstacle detection

## I. INTRODUCTION

Nowadays, pattern recognition has become an interesting task in several applications such as Advanced Driver Assistance System (ADAS) specifically for pedestrian and vehicle detection. The need for such system is motivated, unfortunately, by the number of pedestrians killed in road accidents each year. With 1.25 million deaths, each year [1], the World Health Organization describes traffic accidents as one of the major causes of death and injuries around the world.

To enhance pedestrian's safety and prevent vehicles collision, several pairs of sensors were used in ADAS applications such as, camera and RADAR [2], camera and LIDAR [3], thermal camera [4], stereovision [5], [6], etc. Most of the ADAS are based on one vision sensor with generally another active sensor. Currently, the recent advances

in image resolution and power computing platforms, computer vision systems are becoming increasingly available for ADAS. Some new high-end cars are already equipped with several on-vehicle sensors to prevent danger cases. In this context, our application is integrated in order to detect and recognize different obstacles in an urban environment and aimed at helping drivers to see the road environment and reduce traffic accidents with an automotive monocular camera.

Pedestrian and vehicle detection tasks have dominated the recent works in ADAS. They represent the most complex objects, since they have a significant inter-variability in the shape, size, color, and appearance found in typical driving scenarios. This type of obstacle has made the detection process a major challenge so far. Consequently, it is necessary to investigate more powerful feature extraction methods to address the obstacle recognition challenge. The main idea in this work is to build a dedicated descriptor for each type of obstacle without changing the process of recognition. Personalizing the parameters of single descriptor to extract features and recognizing several type of objects, makes it possible to gain in speed and area consumed by implementation process.

The structure of the detection task for typical computer vision systems using a monocular camera is illustrated in Fig. 1. In the obstacle detection chain, images are acquired through a camera: a sliding window function scans the entire image and generates several sub-windows named Regions of Interest (ROIs). First, the descriptor extracts the significant features, namely shape, local distribution of gradient intensity and edge directions presented in each sub-window. Second, it supplies the classifier to decide whether the desired obstacle is present or not for each ROI.

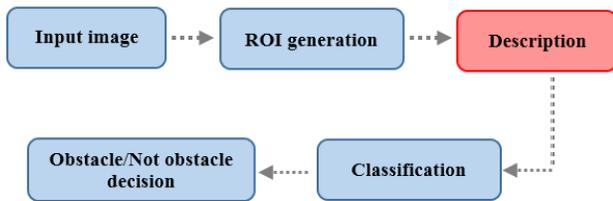


Fig. 1. Obstacle detection chain for typical computer vision systems.

Our obstacle detection process is included in the conventional passive supervised machine learning. Supervised learning method takes a known dataset (images in our case) and known responses to the data named labels (positive examples/negative examples), and tries to build a predictive system which can be used for mapping a new unknown image. In literature, various combinations of descriptor/classifier pairs are commonly used to recognize a special obstacle. Furthermore, we find some descriptors that are more suitable to characterize an object among others. We mention for examples, the HOG descriptor developed by N. Dalal and B. Triggs [7] for pedestrian detection, Haar wavelets by viola et al.[8]for face detection, the LBP descriptor T. Ojala et al. [9] characterized by their low computational cost, and finally the combination between several descriptors as in [10]-[12].

According to the state of the art, the two best known classifiers are Adaboost and Support Vector Machine (SVM). Several weak classifiers are combined into a stronger one to define an Adaboost classifier, while the SVM constructs one or a set of hyperplanes in a high dimensional space in order to achieve a good separation (the largest margin) between the positive and negative training dataset. In more general overviews, most of the proposed works that focus on an obstacle detection system (based on supervised learning machine) combine the HOG descriptor with the SVM classifier [13]-[17] and the Haar features with AdaBoost classifier [18]-[20]. These combinations achieve a better result owing to the logarithmic adaptation between the constituents of every pair. In this paper, we will be interested by the pedestrian and vehicle detection at once; something that is not enough developed in recent works. We will use a modified HOG [21] with a linear SVM as a descriptor / classifier pair to detect and identify the desired obstacles.

This paper, present an analyzing and customizing of the HOG model presented by N. Dalal et al. [7] in order to create a dedicated descriptor for each type of obstacle without changing the process of recognition. The remaining of this paper is organized as follows: in the next two sections, we summarize some related works and describe briefly the computation steps of the standard HOG descriptor. The proposed customized HOG for each obstacle will be presented and discussed in Section 4. Experimental results for pedestrian and vehicle recognition are given and discussed in section 5. Finally, Section 6 will conclude this work.

## II. RELATED WORK

Numerous studies have been conducted in order to address

the detection of pedestrians, vehicles, road signs or other objects that can be presented in a road scene. But, only a few of them considered the detection of various obstacles at once by the same technique, specifically for a pedestrian and vehicle detection tasks. In this work, we will focus on the pedestrian and vehicle detection problem simultaneously, something that has not been explored enough in recent works on computer vision systems.

Over the past few years, several feature extraction processes have been done. We will mention a few of them based on the HOG descriptor for pedestrian and vehicle detection. The HOG descriptor has been initiated by N. Dalal et al. [10], it is a powerful feature extraction method dedicated for the human shape. A modified approach proposed by G. Ballesteros et al. [22] yielded a reduced set of HOG features. In this way, the dimensionality of the feature vector was decreased significantly. The mechanism proposed by Zhang et al. [23] adapts the cell size in the descriptor entry by a limiting ratio (length / width = 2), then each image is divided in  $8 \times 16$  cells per average. Jia et al. [24] integrated the HOG descriptor in the Viola's face detection Framework (viola et al. [8] at the end to achieve the descriptor effectiveness and the Framework speed. X. Wang et al. [25] combined a Local Binary Patterns descriptor (LBP) with the HOG algorithm to define a new descriptor called HOG-LBP. The performance of their algorithm for the pedestrian detection exceeds that of standard HOG. Q. Zhu et al. [26] developed a real-time system by integrating a cascade of rejectors with HOG features to achieve a fast and accurate human detection system. In [27] a new descriptor called Scale Space Histogram of Oriented Gradients (SS- HOG) was considered. The authors have used the multiple scale property to describe an object.

Influenced by the high performance reported by the HOG descriptor, some other research have considered the advantages and extracted features for other objects like face, head, bicycle, car, etc. Some works related to vehicle detection are mentioned below. A typical preceding system for vehicle detection using a standard HOG descriptor and SVM classifier has been presented by M. Ling et al. [28] and X. Li et al. [29]. While, Arróspide et al. [30] have proposed an HOG-like gradient-based descriptor for vehicle verification with an exploitation of the known rectangular shape of vehicle rears. To detect vehicles in videos, a combination of Haar-features and HOG-features has been presented by H. Youpan et al. [31]. The authors have expressed that their method can classify and detect the vehicles in multi-orientations with good classification results. The same procedure was proposed by P. Negri et al [32], but with a comparative study between the Haar-like features, the HOG features and their fusion. The results show that the fusion combines the advantages of the first two detectors. Known that the standard parameters of HOG are optimized for human recognition, a re-optimization of the HOG parameters for vehicle detection has been presented by G. Ballesteros et al. [22]. They have tested various combinations in their experiments, and the results show that  $[-\pi, \pi]$  as orientation range, (n=4) as the number of cells, (p=16) as the number of orientation bins and a nonlinear kernel on SVMs are the most suitable choice for vehicle detection.

In this paper, an innovative technique is proposed to customize the standard HOG for each obstacle and then a comparison between our approach and other works will be presented.

### III. OVERVIEW OF HOG FEATURES DESCRIPTOR

The HOG-features extraction approach could be used to describe a specific gradient orientation in local parts of the image. Such algorithm calculates the gradient direction in small areas of an image, then it assembles the information obtained from all regions into a single vector. N. Dalal et al. [7] have subdivided the image into regions of  $8 \times 8$  pixels that are named cells. Indeed, the HOG feature extraction method consists of calculating the cell-histogram vectors (each vector contains 9 bins and represent the histogram of orientated gradients in one cell), then concatenating them in a single vector. To increase immunity against light variations and lighting conditions, the authors in [7] have normalized all  $2 \times 2$  neighboring cells (which were called a block) to an L2-norm using the following equation:

$$V \rightarrow \frac{v}{\sqrt{\|v\|^2 + \epsilon^2}} \quad (1)$$

where  $V$  is the normalized vector,  $v$  is the non-normalized vector and  $\epsilon$  is a very small constant.

The final HOG feature vector is the collection of the normalized vectors for all the blocks, with an overlapping of 50% per cell. Considering a sliding window of  $64 \times 128$  pixels presented in Fig. 2, it contains  $7 \times 15$  blocks. The assembly of normalized vectors for all blocks into a single 1-D vector then gives 3780 components ( $36 \times 7 \times 15 = 3780$ ). The first observation reveals that this feature extraction is a dense representation that maps local image regions to high-dimensional feature spaces. They will be used to train a linear SVM classifier.

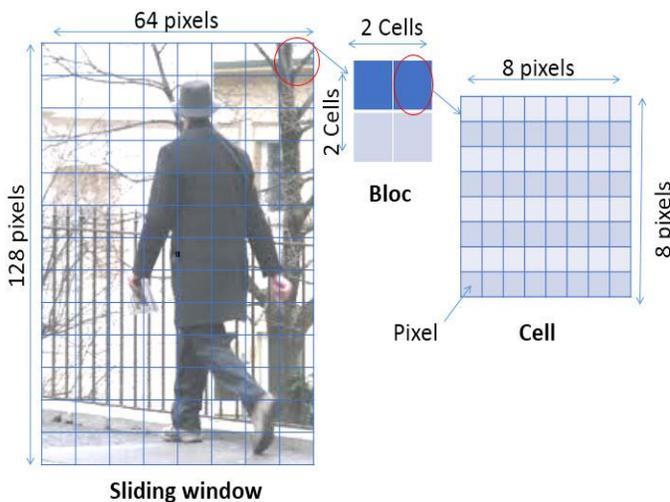


Fig. 2. Overview of HOG feature configuration: window, blocks and cells configuration.

### A. Gradients and oriented gradients computation

Gradient computation is the first step to extract the HOG features. To calculate the pixel gradients, several techniques have been previously presented. Among these techniques, the use of a centered derivative mask  $[-1, 0, 1]$  turns out to be the best result [7]. The application of the selected gradient operator provides the edge intensity and orientation value for each pixel. The horizontal gradient  $dx(x,y)$  and vertical gradient  $dy(x,y)$  of the pixel  $I(x,y)$  are calculated through equations (2) and (3), while the magnitude  $M(x,y)$  is calculated through equation (4).

$$dx(x, y) = [-1, 0, 1] * I(x, y) \quad (2)$$

$$dy(x, y) = [-1, 0, 1] * I(x, y) \quad (3)$$

$$M(x, y) = \sqrt{dx(x, y)^2 + dy(x, y)^2} \quad (4)$$

Furthermore, the gradient orientation  $\theta(x,y)$  is given by equation (5):

$$\theta(x, y) = \arctan\left(\frac{dx(x, y)}{dy(x, y)}\right) \quad (5)$$

### B. Spatial / Orientation Binning

The histograms show the partition of the orientated gradient elements over the cells. In [7], authors have divided the gradients orientation  $[0^\circ-180^\circ]$  “unsigned gradient” into 9 intervals with the same range ( $20^\circ$  for everyone) as showing in Fig. 3.

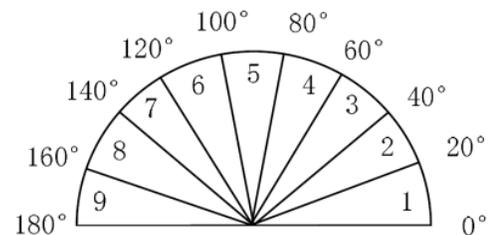


Fig. 3. Orientation range of each bin.

Each interval will be represented by a bin that codes the occurrence’s frequency of the gradients orientation in a cell. In practice, each pixel in the cell contributes with a vote to the two closest histogram channels, weighted according to the gradient magnitude at the location  $(x, y)$ . To summarize, the histogram of oriented gradients is a histogram of neighborhood pixels according to their gradient orientation and is weighted by their gradient magnitude.

### C. SVM classifier

By using Supervised Learning Machine in obstacle detection systems, the common characteristics of the samples belong to the same class (training phase) can be determined, which allows the system to subsequently recognize the class of a new unknown sample (decision phase). The SVM

classifier belongs to the class of Supervised Learning Machines. Such algorithm tries to build an optimal hyper-plane in order to separate the examples of two different classes during the learning phase (B.E. Boser et al [33] and V. Vapnik [34]). Thus, the decision is made using the previously constructed hyper-plane.

Considering the following set of learning examples and associated class labels  $\{X_k, Y_k\}$

Where,  $X_k$  denote the HOG vectors and  $Y_k \in \{-1, 1\}$  the class labels.

Initially, the method ensures the transformation of  $X_k$  in a larger space using a kernel function  $\phi(x)$ . Then it tries to find a decision function which is given by equation (6):

$$f(x) = w * \phi(x) + b \tag{6}$$

Where, the decision function  $f(x)$  is optimal in the sense that it maximizes the distance between the nearest point  $\phi(x_i)$  and the hyper-plane. The class label of the HOG vector is then obtained by considering the sign of  $f(x)$ . Solving the optimization problem is obtained by using the following equations:

$$\min_{w, \xi} \frac{1}{2} \|w\|^2 + c \sum_{i=1}^m \xi_i \tag{7}$$

subject to the constraints:

$$y_i (w^T x_i + b) \geq 1 - \xi_i \quad \text{for } i = (1 \dots m) \tag{8}$$

Where, the variables are  $\xi_i$  known as slack variables. The regularization parameter C is a positive constant that controls the relative influence of the two competing terms. In our experiments, we will use the linear SVM as our binary classifier due first to the large number of HOG features (one may not need to map data to a higher dimensional space) and second to its faster computation.

#### IV. IMPROVED HOG APPROACH

In the following subsections, we describe in depth the complete framework of the proposed detection system. The main goal of our approach is to increase the accuracy of the road-obstacle detection system. Our study has presented an improvement for the two most common obstacles in the road (pedestrian and car), but not only limited to these two types. In fact, the method can be applied to other road obstacles such as; buses, bikes, animals (dogs, cats, antelope...) or to recognize the traffic signs. Pedestrians and cars are the most complex obstacles for the detection and identification task, due to their change in appearance and position previously mentioned. The steps involved in the proposed approach are the following.

Firstly, we modified the histogram building method of the standard HOG algorithm to get an average histogram of oriented gradients for each selected eigenvector. Secondly, we apply a new procedure to extract the bins that better

characterize the desired object features. Finally, we amplify the selected bins in the new customized HOG algorithm that will be included in the main chain of the vision system. A general overview of the complete framework can be seen in Fig. 3. More explanations for each phase are presented in the following subsection (Fig. 4).

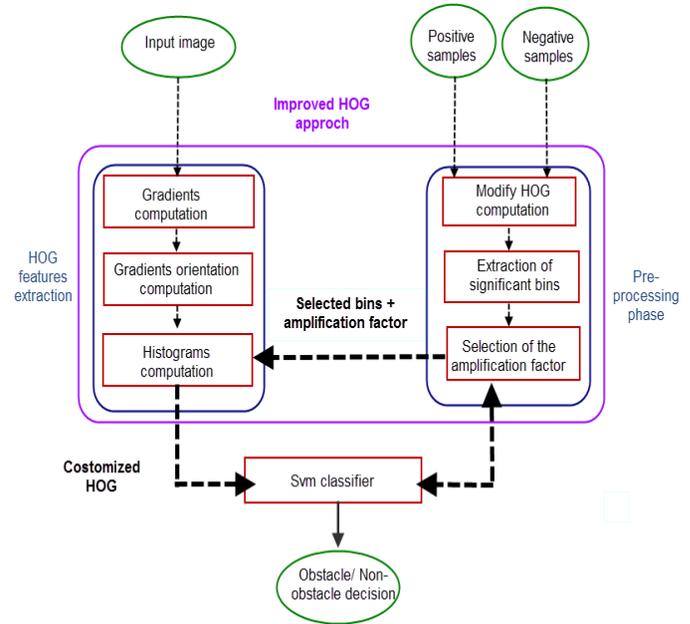


Fig. 4. Whole system of the customized HOG approach.

#### A. Modified computational method

The histograms that better characterize the desired object are selected by adding all the obtained cell-histogram vectors instead of its concatenation as in the original HOG algorithm. This process gives an average vector that containing 9 bins in the whole image instead of 3780 components. Based on the adopted Dalal's approach [10], the local normalization block was maintained in order to guarantee the immunity against lighting conditions. Nevertheless, the overlap of the cells was removed, having negligible effect in this stage.

Therefore, the first phase provides 9-components mean vector  $X_1$ , characterizing the image of the object to be detected. Regarding the large inter-variety between pedestrians, we must now generalize this vector through averaging it in the whole dataset that contains n pedestrian images. This step can be obtained through equation (9).

$$X = \frac{\sum_{k=1}^n X_k}{n} \tag{9}$$

where  $X$  represents the mean vector of the whole positive database,  $X_k$  represents the vector for the image number  $k$  and  $n$  is the total number of the positive examples in the dataset.

MIT CBCL and INRIA pedestrian datasets are the two most commonly used databases in the field of computer vision machine for the pedestrian recognition task. All of them can

be publicly accessed. In our experiments, the mean vectors were obtained by averaging all the pedestrian images of the entire INRIA and MIT datasets. The same procedure was performed for the calculation of the average vector for the negative examples (not pedestrian images) in the INRIA dataset (all training and test negative examples). Fig. 5 illustrates the mean vectors calculated for the two datasets (pedestrian and non-pedestrian images).

**B. Extraction of the significant bins**

At present, we have two main vectors that define a pedestrian image and a random image through 9 bins for each one. Thereafter, we calculated the difference between the two histograms in order to extract the most frequent orientations presented in pedestrian images. The result vector is represented in Fig. 6.

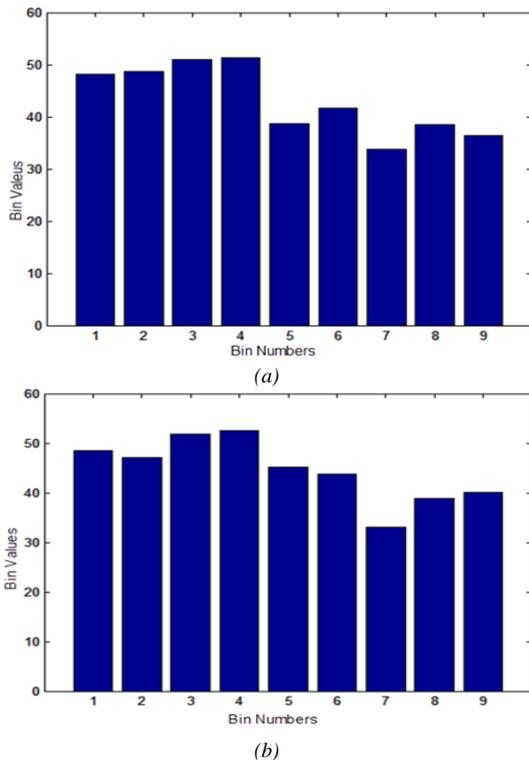


Fig. 5. Mean vectors of the used dataset: (a) Mean vector of the pedestrian images, (b) Mean vector of the random images.

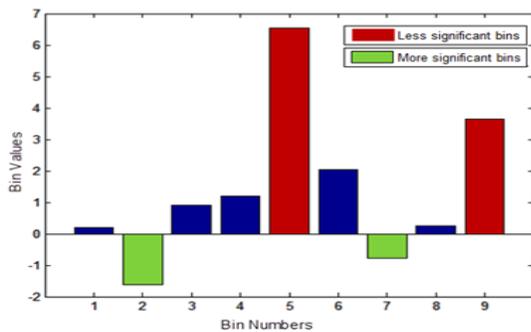


Fig. 6. Difference between the two mean vectors of the used datasets.

As shown in Fig. 6, the subtraction between the mean vectors gives two special bins (2 and 7) whose values are reversed when compared to other bins. Indeed, these bins have larger gradient density in a pedestrian image than a random image in traffic environments. In other words, these bins encode the edge orientations that describe the shape of the human bodies. Thus, we called them the most significant bins. On the other hand, the bin numbers 5 and 9 represent the highest values in this histogram. These bins encode the least frequent oriented gradients for pedestrian images. Then, we called them the less significant bins. In the last phase of our proposed algorithm, a modification will take place in the vote partition of the oriented gradient elements that may be very promising.

**C. Amplifying the extracted bins**

The main idea is to amplify the most significant bins using an alpha parameter ( $\alpha > 1$ ) in the cell-histograms building step. The physical significance of this amplification is to highlight the contrast of the contour for some specific orientations that describes the shape of the human body. Actually, the different bins of the HOG-feature vector will not share the same weight, and an amplification factor will be distributed for each bin in order to increase the weight bins that describe the relevant obstacle features.

**V. EXPERIMENTAL RESULTS AND DISCUSSION**

In order to well assess the measures performance, the experimental results are evaluated based on the three statistical measures test of a binary classification: Accuracy, sensitivity and specificity. Accuracy measures the proportion of actual positives and negatives samples which are correctly identified. Sensitivity measures the proportion of actual positive samples which are correctly identified (e.g. the percentage of pedestrian images which are identified as a true pedestrian image). Specificity measures the proportion of negative samples which are correctly identified (e.g. the percentage of non-pedestrian images that are identified as a true non-pedestrian image). Their expressions are:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

$$Sensitivity = \frac{TP+TN}{TP+FN} \tag{11}$$

$$Specificity = \frac{TP+TN}{TP+FP} \tag{12}$$

where  $TP$  is the number of true positives; number of pedestrian images correctly classified;  $TN$  is the number of true negatives; number of non-pedestrian images correctly classified;  $FP$  is the number of false positives; number of pedestrian images classified as non-pedestrian;  $FN$  is the number of false negatives: number of non-pedestrian images classified as pedestrian.

**A. Pedestrian detection**

In this section, the impact of the proposed algorithm is analyzed. The used datasets to evaluate our approach are INRIA [35] and MIT [36]. The first dataset contains 2416 positive examples (1208 pictures with their reflections of

horizontal axis) and 1218 negative examples. It contains pedestrians in various postures, clothing as well as wide variety of backgrounds and lighting condition. This makes it one of the most complex databases for pedestrian detection. The MIT dataset contains only positive examples (709 images). The bodies of pedestrians are centered and they have almost the same size in the image. Additionally, a pedestrian is shown alone in a front or rear position. Therefore, these characteristics make the MIT dataset less complicated than the INRIA dataset. Fig. 7 shows some images from the datasets.

The fusion of the two databases in our learning system provides a greater efficiency in the general detection system, which is assessed in terms of 2% increase in the recognition rate of the system.

1) Alpha parameter study

The system has been trained with INRIA and MIT training datasets and tested with INRIA Test Dataset. The global recognition system, presented in Fig. 8, was increased by amplifying the most significant bins extracted in the first phase (bin No 2 and 7). This comes as expected due to the amplification of the characteristics concerning the pedestrian shape in the image. By varying the value of the amplification factor  $\alpha$ , the sensitivity rate changes significantly, while the specificity still globally unaffected.

On the other hand, the amplification and the attenuation of the least significant bins (bin No 5 and 9) reduce respectively the recognition rate by 1.5% and 0.5%. Therefore, we have maintained their values without modifications.

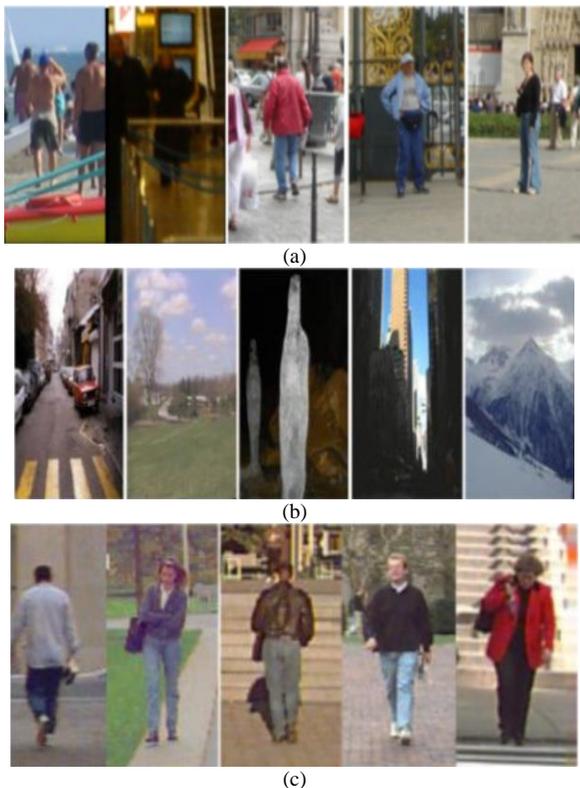


Fig. 7. Image examples from datasets (a) Positive examples of INRIA Dataset, (b) Negative examples of INRIA Dataset and (c) Positive examples of MIT Dataset.

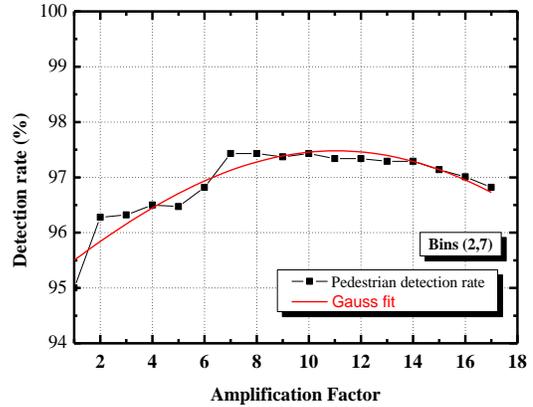


Fig. 8. The system performance according to  $\alpha$  value.

2) Experimental results

The details of the used database for learning and testing the performance of the pedestrian detection system are presented in Table 1.

A comparison between the results of our approach and further works based on HOG descriptor is shown in Table 2.

TABLE I. DETAILS OF THE USED DATABASE

Data set	Learning		Test	
	Positive Examples	Negative Examples	Positive Examples	Negative Examples
INRIA	2416 (128×64)	1218 (320×240)	1126 (128×64)	453(different)
MIT	709 (128×64)	--	214 (128×64)	--
Used	3125 (128×64)	6090 (128×64)	MIT214 (128×64) INRIA 1126 (128×64)	INRIA 2265 (128×64)

TABLE II. COMPARISON OF THE EXPERIMENTAL RESULTS

Descriptor	Datasets			
	Sensitivity MIT	Sensitivity INRIA	Specificity INRIA	Accuracy MIT + INRIA
HOG N. Dalal et al. (2005)	81.39%	85.9%	99.10%	88.79 %
Optimized HOG G. Zhang et al. (2010)	100%	99.07%	98.89%	99.32%
Integral HOG Y. Said et al. (2012)	---	94.48%	97.56%	96.02%
<b>This work</b>	<b>100%</b>	<b>97.42%</b>	<b>100%</b>	<b>99.14%</b>

It can be concluded from Table 2 that a perfect recognition rate (100%) for the Negative INRIA dataset and the MIT dataset is obtained, together with a respectable percentage for the Positive INRIA dataset. As a conclusion, the proposed

system yields a significant performance in the characterization of pedestrian features, when compared to the other works [7], [40].

**B. Vehicle detection**

*1) Database*

To build a vehicle recognition system in the conventional supervised learning, the positive training examples consist of vehicles, and the negative training examples consist of random non-vehicles. The datasets used in our system are MIT cars [37], INRIA cars [38] and Markus cars [39] as positive examples and non-pedestrians INRIA datasets as negative examples. Fig. 9 shows some positive and negative examples. We manually delete the images for non-pedestrian examples that contain cars in order to use them as negative examples for learning. We have obtained 988 car images with their reflections (1 976 samples in total) as positive examples and 4 236 samples extracted from 1 059 not-car images as negative examples. 1/3 of each database was intended for test and 2/3 were intended for learning the system.

*2) Re-optimizing the HOG parameters for vehicle detection*

After the validation of our approach in pedestrian detection system by increasing the system’s accuracy, we will now generalize the proposed approach through the detection of other various obstacles in an urban environment. Based on the same principle (increase of significant bins of each specific obstacle), we will be interested in detecting and identifying the vehicle obstacles. In an image, pedestrian and car have various different characteristics. The HOG descriptor is primarily built for pedestrian detection. Therefore, we need to re-optimize several parameters of the standard HOG descriptor to get the best results for the car detection. Then we added the process of our approach.

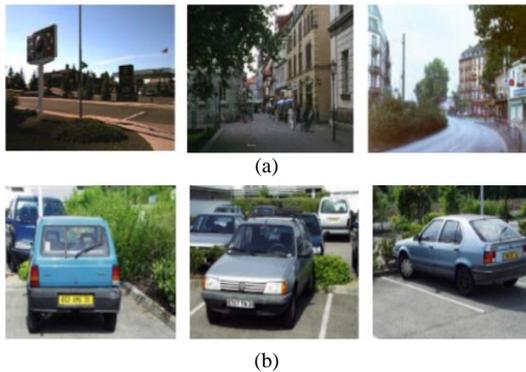


Fig. 9. Image examples from datasets: (a) Negative examples, (b) Positive examples.

Primary, most of the vehicles have rectangular shapes and they have a larger size than a pedestrian, justifying the choice of (128×128) pixels per window in the learning system. Second, changing the number of pixels per cell, the number of cells per block and the overlapping ratio does not affect the system’s performance.

Then we will keep the same parameter’s values (8 × 8 Pixels per cell, 2×2 Cells per block) proposed in the standard HOG, that turns out to be effective to express the features of

cars in images. Finally, vertical orientations for a car are characterized by an acute and accurate angle, which does not change within its movement at variance with the pedestrians. That leads to minimizing the scale of bins by increasing her number in [0 Π] plan. The simulation results for different values of bin’s number are shown in Table 3.

The simulation provides the best result for 60 bins through an accuracy rate equal to 97.3% with an enhancement of 1.47%. However, it represents the most complex and greedy simulation: resource intensive, memory consuming, execution time... Since such application target an automotive embedded system, working with higher feature-dimension will slow the learning step and may be risking the over-fitting of the SVM classifier in the hardware implementation. Therefore, in the following, we will apply our approach for a number of bins equal to 18; first, in order to save simulation time, second to target an efficient hardware implementation of real time vehicle detection and finally to demonstrate the efficiency of our approach since this case represents the lower sensitivity.

The sample size used in our experiment is a window of 128 × 128 pixels that define the car and non-car images. The HOG-feature extraction process based on 18 bins gives a features vector of 16200 dimensions as shown Fig. 10.

TABLE III. RECOGNITION RATE ACCORDING TO THE NUMBER OF BINS IN THE HISTOGRAMS

Number of bins	6	9	18	24	36	60	72
HOG Feature length	5400	8100	16200	21600	32400	54000	64800
Sensitivity rate (%)	93.44	93.88	91.26	93.8865	96.94	97.4	96.3
Specificity rate (%)	97.50	97.78	99.12	98.32	97.58	97.2	96.7
Accuracy rate (%)	95.47	95.83	95.19	96.1	97.26	97.3	96.5

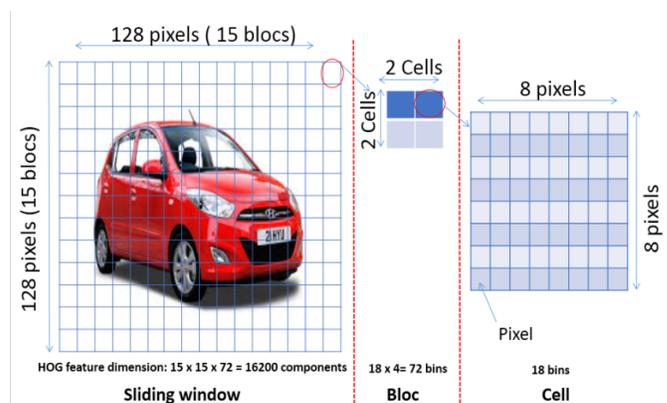


Fig. 10. Overview of the HOG features configuration for the vehicle detection: window, blocks and cells.

*3) Experimental results*

To better extract the car features, we have applied our approach on the whole datasets through these three steps:

- Select the significant bins that better describe a car feature from other obstacles

- Optimize the amplification factor for each selected bin
- Inject the proposed change in the standard HOG algorithm

The results of each step are presented below.

#### 4) Extraction of significant bins

As shown in Fig. 11, the subtraction between the two mean vectors of the negative and positive examples for the car datasets used in our experiments (INRIA, MIT, MARKUS) give four bins (2,7,15 and 17) whose values are reversed compared to the others bins. By the same logic we explain that these bins have larger gradients orientation density in a car image than a random image in traffic environments. On the other hand, the bin numbers (6, 10 and 14) represent the highest values in this histogram. These bins encode the least frequent oriented gradients for a car image.

Afterwards, we will amplify several combinations between the most significant bins, the least significant bins and their mixture, in the purpose of extracting the best possible system's accuracy.

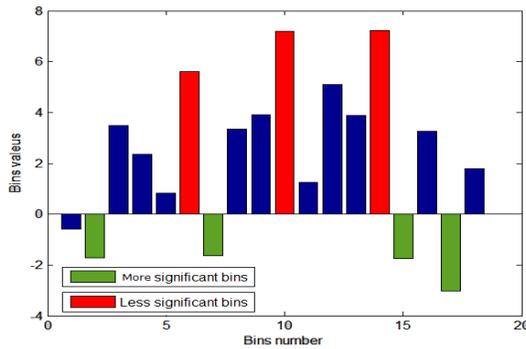


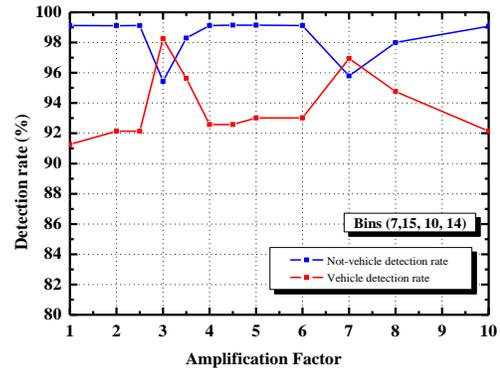
Fig. 11. Most significant bins for a car features extraction based on 18 bins.

#### 5) Select the amplification factor

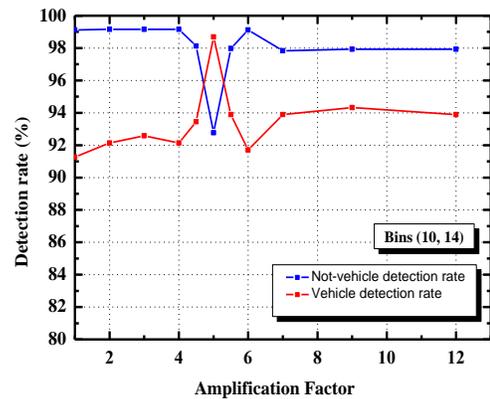
In Fig. 12, we represent the best three combinations of the different significant bins. Equally to the case of the pedestrian detection system, we swept the amplification factor at the end to get the best recognition rate for the car detection. Through experimentations, we can reveal that the significant bins are sensitive to the amplification process. The best sensitivity rate has reached 98.69%, by amplifying the bins 10 and 14 with a factor equal to 5. However, we note a clear degradation for the specificity rate that attains 92.76%. The amplification of bins 2,7,15 and 17 has achieved the accuracy in all simulations; in fact, we attain a sensitivity non-vehicle recognition system. The amplification factor getting the highest rate is equal to 3. A comparison between our results and other ones (presented in Table 4), shows that the proposed approach outperforms recent works [28], [30], [31], [41]. However, we cannot rely on this comparison because we do not share the same database, seeing that a growing number of on-road equal to 98.25% and a specificity equal to 95.43% for the vehicle studies are reporting results from private video datasets.

Through the whole experimentation of the pedestrian and the vehicle detection, we have observed that the customized

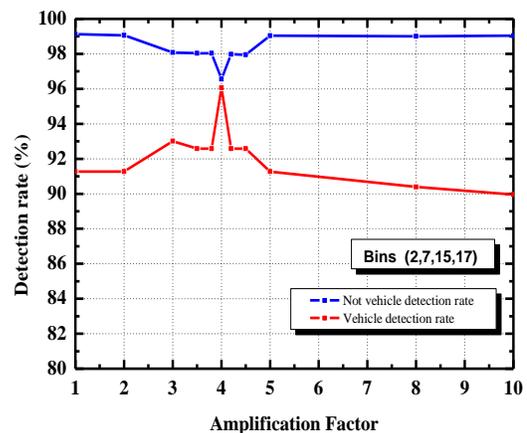
HOG-feature extraction method goes well with several types of obstacles. In addition, a tracking technique can be introduced to supply missing and false detection.



(a)



(b)



(c)

Fig. 12. Detection rate of the best three combinations for the different significant bins: (a) Amplification process for bins (7, 15, 10, 14), (b) Amplification process for bins (10, 14), (c) Amplification process for bins (2,7,15,17).

TABLE IV. COMPARISON OF EXPERIMENTAL RESULTS

Methods	Sensitivity rate	Specificity rate	Accuracy rate
HOG-Like Gradient <i>J. Arrosipide et al. (2012)</i>	92.48%	--	--
Haar-features and HOG-features <i>H. Youpan et al. (2013)</i>	97.2%	96.8%	97%
Standard HOG <i>Ling Mao et al. (2010)</i>	96.87%	97.33%	97.1%
HOG-HCT <i>S. Li et al. (2012)</i>	85.2%%	--	--
<b>This work</b>	<b>98.25%</b>	<b>96.32%</b>	<b>97.28%</b>

## VI. CONCLUSION

In this paper, we have proposed an improved version of the HOG feature extraction called Customized HOG. The main contribution is to extract and amplify the most significant bins that describe particularly the desired object. This technique presents a potential solution to the emerging problems related to the obstacle detection for ADAS as well as other applications. The performance evaluation shows that the proposed approach yields significant improvements for the characterization of pedestrians and vehicles features compared to other approaches. Future research works will focus on real-time object detection and its implementation on Field Programmable Gate Arrays (FPGAs) using the proposed customized HOG and some techniques to reduce the feature dimensionality.

## REFERENCES

- [1] Global status report on road safety 2015. World Health Organization, 2015
- [2] Matthias Serfling, Otto Loehlein, Roland Schweiger and Klaus Dietmayert, 2009. Camera and Imaging Radar Feature Level Sensor fusion for Night Vision Pedestrian Recognition. IEEE Intelligent Vehicles Symposium, 597 – 603, Xi'an, 2009, .
- [3] Daniel Clarke, Alois Knoll, Feihu Zhang: Vehicle Detection Based on LiDAR and Camera Fusio. 2014. Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on (October. 2014)
- [4] Hao Sun, Cheng Wang and Boliang Wang.2011. Night vision pedestrian detection using a forward looking infrared camera. International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping, 1-4, (Jan. 2011).
- [5] Leu. A, Aiteanu. D and Graser. A. 2011. A novel stereo camera based collision warning system for automotive applications, Applied Computational Intelligence and Informatics (SACI), 2011 6th IEEE International Symposium on, 409 – 414 (May.2011).
- [6] Mesmakhosroshahi. M, Joohee Kim, Yunsik Lee, and Jong-bok Kim. 2013. Stereo based region of interest generation for pedestrian detection in driver assistance systems. Image Processing (ICIP), 2013 20th IEEE International Conference on, 3386 – 3389, (Sept. 2013).
- [7] N. Dalal and B. Triggs. 2005. Histograms of Oriented Gradients for Human Detection. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 886-893 (June. 2005)..
- [8] Viola. P, Jones. M.2001. Rapid Object Detection using a Boosted Cascade of Simple Features. Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on , vol 1, 511-518, 2001.
- [9] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, Jul 2002
- [10] Y. Cao, S. Pranata, M. Yasugi, Z. Niu and H. Nishimura, "Staged multi-scale LBP for pedestrian detection," 2012 19th IEEE International Conference on Image Processing, Orlando, FL, pp. 449-452, 2012.
- [11] Jingjing Li, Yong Zhao, Dongbing Quan. 2013. The combination of C SLBP and LBP feature for pedestrian detection. International Conference on Computer Science and Network Technology (ICCSNT), 543 – 546, (Oct. 2013).
- [12] Cosma. C, Brehar. R, Nedeveschi. S. Pedestrians detection using a cascade of LBP and HOG classifiers. 2013. International Conference on Intelligent Computer Communication and Processing (ICCP), 69 – 75, (Sept. 2013).
- [13] Laoprasa. N, Thongkrau. T, Sunat. K, Songrum. P, Chamchong. 2014. Improving vehicle detection by adapting parameters of HOG and kernel functions of SVM. International Computer Science and Engineering Conference (ICSEC), 372 – 377, (July. 2014)..
- [14] N. Dalal, B Triggs, and C Schmid. 2006. Human Detection Using Oriented Histograms of Flow and Appearance. 9th European Conference on Computer Vision, Graz, Austria. Computer Vision – ECCV, 428–441, 2006.
- [15] XianbinCao, Changxia Wu, Pingkun Yan, Xuelong Li.2011. Linear SVM classification using boosting HOG features for vehicle detection in low-altitude airborne videos. 18th IEEE International Conference on Image Processing (ICIP), 2421 – 2424, (Sept. 2011).
- [16] Lorca. D.F, Arroyo. R, Sotelo. M.A.2013. Vehicle logo recognition in traffic images using HOG features and SVM. International Conference on Intelligent Transportation Systems - (ITSC), 2229 – 2234, (Oct. 2013)..
- [17] ChengbinZeng, Huadong Ma, Anlong Ming. 2010. Fast human detection using mi-sVM and a cascade of HOG-LBP features. 17th IEEE International Conference on Image Processing (ICIP) 3845 – 3848, (Sept. 2010).
- [18] M. P. Hoang, Le Dung, De Souza-Daw. T, Nguyen Tien Dzung, Thang Manh Hoang. 2012. Extraction of human facial features based on Haar feature with Adaboost and image recognition techniques. International Conference on Communications and Electronics (ICCE), 302 – 305, (Aug. 2012).
- [19] Van-Dung Hoang, Vavilin. A, Kang-Hyun Jo.2012. Pedestrian detection approach based on modified Haar-like features and AdaBoost. 2012. International Conference on Control, Automation and Systems (ICCAS). 614 – 618, (Oct. 2012).
- [20] Rakate, G.R, Borhade. S.R, Jadhav, P.S, Shah. M.S.2012. Advanced Pedestrian Detection system using combination of Haar-like features, Adaboost algorithm and Edgelet-Shapelet, International Conference on Computational Intelligence & Computing Research (ICCIC). 1-5, (Dec. 2012).
- [21] H. Ameur, A. Helali, M.Nasri, H. Maaref. (2014), Improved feature extraction method based on Histogram of Oriented Gradients for pedestrian detection, Computer & Information Technology (GSCIT), pp 1 – 5, 2014
- [22] Ballesteros. G, Salgado. L. 2014. Optimized hog for on-road video based vehicle verification. Signal Processing Conference (EUSIPCO), 805 – 809, (Sept. 2014).
- [23] G. Zhang, F. Gao, C. Liu, W. Liu, H. Yuan. 2010. A Pedestrian Detection Method Based on SVM Classifier and Optimized Histograms of Oriented Gradients Feature, International Conference on Natural Computation (ICNC), 3257 – 3260, (Aug. 2010).
- [24] Hui-Xing Jia, Yu-Jin Zhang. 2007. Fast Human Detection by Boosting Histograms of Oriented Gradients. International Conference on Image and Graphics, 683 – 688, (Aug. 2007).
- [25] X. Wang, Tony X. Han, S. Yan. 2009. An HOG-LBP Human Detector with Partial Occlusion Handling. Computer Vision, 2009 IEEE 12th International Conference on, 32 – 39, (Sept. 2009)..
- [26] Q. Zhu, S. Avidan, Mei-Chen Yeh, and K. Cheng. 2006. Fast Human Detection Using a Cascade of Histograms of Oriented Gradients.2006. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on (Volume 2), 1491 – 1498, 2006.

- [27] N. He, J. Cao and L. Song. 2008. Scale Space Histogram of Oriented Gradients for Human Detection, International Symposium on Information Science and Engineering. 167 – 170 (Dec. 2008)..
- [28] Ling Mao, Mei Xie, Yi Huang, Yuefei Zhang. 2010. Preceding Vehicle Detection Using Histograms of Oriented Gradients, International Conference on Communications, Circuits and Systems (ICCCAS), 354 – 358, (July. 2010).
- [29] Xing Li, Xiaosong Guo. 2013. A HOG Feature and SVM Based Method for Forward Vehicle Detection With Single Camera, International Conference on Intelligent Human-Machine Systems and Cybernetics. 263 – 266. (Aug. 2013).
- [30] Arrospide. J, Salgado. L, Marinas. J. 2012. HOG-like Gradient-based Descriptor for Visual Vehicle Detection, Intelligent Vehicles Symposium Alcalá de Henares, Spain. 223 – 228, (June. 2012).
- [31] Youpan Hu, Qing He, Xiaobin Zhuang, Haibin Wang, Baopu Li, Zhenfu Wen, Bin Leng, Guan Guan, Dongjie Chen. 2013. Algorithm for vision-based vehicle detection and classification, International Conference on Robotics and Biomimetics (ROBIO), 568 – 572, (Dec.2013).
- [32] Pablo Negri, Xavier Clady, Shehzad Muhammad Hanif, and Lionel Prevost. 2008. A Cascade of Boosted Generative and Discriminative Classifiers for Vehicle Detection, EURASIP Journal on Advances in Signal Processing, 12, 2008.
- [33] B.E. Boser, I. M. Guyon, and V. N. Vapnik. 1992. A training algorithm for optimal margin classifiers, Proceeding COLT '92 Proceedings of the fifth annual workshop on Computational learning theory 144-152, 1992..
- [34] V. N. Vapnik. 1998. Statistical Learning Theory, September 1998.
- [35] M. Oren and C.P. Papageorgiou and P. Sinha and E. Osuna and T. Poggio, Pedestrian Detection Using Wavelet Templates, cvpr, PP 193-99, 1997.
- [36] N. Dalal, "INRIA Person Dataset". <http://pascal.inrialpes.fr/data/human/>
- [37] CBCL Car database. <http://cbcl.mit.edu/projects/cbcl/software-datasets/cars128x128.tar.gz>
- [38] <http://www.cs.ubc.ca/~pcarbo/objrecls/data/cars.tar.gz>
- [39] Michael Fink and Pietro Perona, Technical report Caltech CSTR:2003.008, 2003.
- [40] Y. Said, M. Atri, T. Saidani and R. Tourki. 2012. Détection d'Individus dans les Images Couleurs à base d'Histogramme Intégral de Gradient Orienté et SVM. Sciences of Electronics Technologies of Information and Telecommunications Sousse. (March. 2012).
- [41] Sun Li, Bo Wang, ZhiHui Zheng, HaiLuo Wang. 2012. Multi-view vehicle detection in traffic surveillance combining HOG HCT and deformable part models. Wavelet Analysis and Pattern Recognition International Conference on. Xian, 202 – 207, (July. 2012).

# A Mathematical Model for Comparing Memory Storage of Three Interval-Based Parametric Temporal Database Models

Nashwan Alromema<sup>1</sup>

<sup>1</sup>Department of Computer Science, Faculty of Computing and Information Technology-Rabigh, Rabigh, Saudi Arabia

<sup>1</sup>Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia

Mohd Shafry Mohd Rahim

Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia

Ibrahim Albidewi

Department of Computer Science, Faculty of Computing and Information Technology, Jeddah, Saudi Arabia

**Abstract**—Interval-Based Parametric Temporal Database Model (IBPTDM) captures the historical changes of database object in single tuple. Such data model violates 1NF and it is difficult to be implemented on top of conventional Database Management Systems (DBMS). The reason behind that, IBPTDM cannot directly use relational storage structure or query evaluation technique that depends on atomic attribute values as well as it is unfixed attribute size. 1NF model with its features can be used to solve such challenge. Modeling time-varying data in 1NF model raise a question about memory storage efficiency and ease of use. A novel approach for representing temporal data in 1NF model and compare it with other main approaches in literature is the main goal of this research. To this end, a mathematical model for comparing a three different storage models is demonstrated to illustrate that the proposed model is more efficient than other approaches under certain conditions. The simulation results showed that the proposed model overcomes the needless redundancy of data, achieves saving in memory storage, and it is easy to be implemented in relational data model or to be adapted with a production systems that need to track temporal aspects of functioning database Systems.

**Keywords**—Valid-time data model; N1NF; interval-based timestamping; temporal data model; 1NF

## I. INTRODUCTION

Modeling temporal database is considered a vital and highly demanding problem. That is why varieties of techniques have been proposed to address this problem from different viewpoints [1]-[3]. Modeling temporal database in relational framework differs in many dimensions [4]-[11]. The most frequently stated approaches are tuple timestamping with First Normal Form (1NF), and attribute timestamping with Non-First Normal Form (N1NF). Based on the timestamp of the data, the first approach (1NF) has two distinctions namely, Tuple Timestamping Single Relation (TTSR), and Tuple Timestamping Multiple Relations (TTMR).

Models under TTSR approach are discussed by [1], [4], [5], [12]-[14]. An example of some of these temporal data models are LEGOL 2.0 by Jones [15], Temporally Oriented Data Model by Ariav [16], HSQL by Sarda [17], and TQuel by Snodgrass [18]. TTSR approach introduces redundancy, where attribute values that change at different time are repeated in multiple tuples. Furthermore, Steiner in [19] stated that, the main disadvantage of this approach is that the fact about a real world entity is spread over several tuples, where each tuple represents a state during a certain time period in the real world.

Models under TTMR approach have solved the problem of data redundancy in TTSR by decomposing the temporal relation as follows: time-varying attributes are distributed over multiple relations, and time-invariant attributes are gathered into separates relation. Many temporal data model discussed in literature are categorized under this approach [5]. An example of some of these temporal data models are Temporal Relational Model by Navathe and Ahmed [20], Snodgrass [18], Tansel [10], and Kvet [14]. The data models under this approach need a variation of join -known as temporal intersection join- that is used for combining the information for an object. Temporal intersection join is generally expensive to be implemented.

The second approach (N1NF) violates the atomicity of single data representations and based on the timestamp, the data can be timestamped in the level of tuple or in the level of attributes [5], [8]-[10]. An example of this approach is the parametric temporal data model that is based on attributes-timestamping and that uses a temporal element as a timestamp [21]. The bitemporal conceptual data model (BCDM) is another example of such approach that forms the basis for the temporal structured query language (TSQL) proposed by Jensen [5]. BCDM is based on tuple-timestamping and it uses interval-based timestamping [2].

Due to the needless redundancy of data in TTSR approach, expensive implementation in TTMR approach, and the implementation difficulty in parametric temporal data model in

top of relational data model, a new approach to model, implements, and query TDB in relational framework is proposed. The proposed approach is referenced as Tuple Timestamp Historical Relation (TTHR) [22]. This temporal data model (TTHR) is based on a tuple timestamping for the lifespan time of database objects, and it is also based on attributes timestamping for the historical valid time changes of time varying attributes. TTHR is in 1NF and it is an extension and reducible of Snodgrass (Tquel) temporal data model [14, 19]. TTHR mimics the features of TTSR and TTMR as well as the most common temporal database models discussed in literature.

Storage efficiency of temporal database systems has a direct impact to the system performance; therefore, in this study we will compare the three approaches (TTSR, TTMR, and TTHR) in terms of memory storage point of view. To measure the storage costs, we will establish mathematical model (formulas) for the three approaches. It can give us a reasonable judgment to determine whether TTHR is suitable for the implementation of the parametric temporal data model in top of conventional DBMS. Throughout our investigations into storage efficiency, we will show that the TTHR approach is comparable to the other approaches and it is even better under certain conditions. A similar study for calculating the efficiency of memory storage has been done by Atya in [2], this study compared the Snodgrass model (which is under TTMR-based approach in our study) with Tansel N1NF relational nested model. A study by Noh in [23] has introduced a new platform for modeling temporal database under XML-based platform. He compared the relational model as he named it (which is under TTSR-based approach in our study) with XML-based, and object-oriented based approach.

## II. INTERVAL-BASED PARAMETRIC TEMPORAL DATABASE MODEL

Interval-based parametric temporal database model uses N1NF and attribute-timestamping data model with interval-based timestamps. The time interval  $[t_1, t_2]$  consists of finite time points between  $t_1$  and  $t_2$  such that  $t_1^2 = [t_1, t_2] = \{t \in T \mid t_1 \leq t \leq t_2\}$ , where  $T$  is defined as a set of time points in the domain  $D$ . Interval start point  $t_1$  and interval end point and  $t_2$  are the minimum and maximum boundary of the interval, both belong to the interval. Intervals can be defined as open, half-closed, or closed. In this study, Left-Right bounded (closed) representation for periods of validity is considered. Intervals can be compared to show their relative positions using Allen's interval logic [24], [25]. Fig. 1 shows *Emp* relation which represents the historical changes of employees' time-varying data with *Address*, *Tel\_no*, *Supervssn*, *Dno* (department number), *Salary*, and *Rank*. *SSN*, *name*, and *Birth\_date* is considered as time invariant attributes. It can be shown in Fig. 1 that the information about database object is modeled in one tuple and each time-varying attribute is timestamped by one or more time interval. An example of time intervals is  $[0, 4]$  and  $[5, 10]$  which timestamp the valid change time of address of employee *Nashwan*. The single time instance can be represented in interval-based, as an example time instance 10 has  $[10, 10]$  representation.

<i>Emp</i>						
<i>SSN</i>	<i>Name</i>	<i>Birth_date</i>	<i>Address</i>	<i>Tel_no</i>	<i>Supervssn</i>	
2091	Ali	1/10/1990	$[0, \infty]$ Rabigh	$[0, \infty]$ 098778776	$[0, \infty]$ Null	
2089	Nashwan	1/5/1980	$[0, 4]$ Jeddah $[5, 10]$ Rabigh	$[0, 10]$ 059987767	$[0, 6]$ Null	$[7, 10]$ 2091

<i>Dno</i>	<i>Salary</i>	<i>Rank</i>	<i>T<sub>is</sub></i>
$[0, \infty]$ 10	$[0, \infty]$ 13000	$[0, \infty]$ C	$[0, \infty]$
$[0, 8]$ 20	$[0, 2]$ 2000	$[3, 9]$ 2500	$[0, 10]$ A
$[9, 10]$ 30	$[10, 10]$ 3000		

Fig. 1. *Emp* relation in interval-based parametric temporal data model.

The three data model (TTSR, TTMR, and TTHR) are based on schema extension approach of conventional relational data model and it can be implemented in conventional RDBMS.

## III. TEMPORAL DATA REPRESENTATIONS IN THREE APPROACHES

In this section, the *Emp* relation shown in Fig. 1 is going to be mapped to TTSR approach, TTMR, and TTHR approach (proposed model), respectively.

### A. TTSR data representation approach

In TTSR, temporal database relations are in 1NF model with snapshot relations. To model *Emp* relation using TTSR representation, let  $R_{TTSR}$  represents the relation in TTSR representation that has the schema structure  $R_{TTSR} = (A_K, A_U, A_C, T_{Is}, T_{Ie}, T_{Vs}, T_{Ve})$ .

<i>Emp</i>												
<i>SSN</i>	<i>Name</i>	<i>Birth_date</i>	<i>Address</i>	<i>Tel_no</i>	<i>Supervssn</i>	<i>Dno</i>	<i>Salary</i>	<i>Rank</i>	<i>T<sub>v</sub></i>	<i>T<sub>ve</sub></i>	<i>T<sub>is</sub></i>	<i>T<sub>ie</sub></i>
2091	Ali	1/10/1990	Rabigh	098778776	2090	10	13000	C	0	$\infty$	0	$\infty$
2089	Nashwan	1/5/1980	Jeddah	059987767	Null	20	2000	A	0	2	0	10
2089	Nashwan	1/5/1980	Jeddah	059987767	Null	20	2500	A	3	4	0	10
2089	Nashwan	1/5/1980	Rabigh	059987767	Null	20	2500	A	5	6	0	10
2089	Nashwan	1/5/1980	Rabigh	059987767	2091	20	2500	A	7	8	0	10
2089	Nashwan	1/5/1980	Rabigh	059987767	2091	30	2500	A	9	9	0	10
2089	Nashwan	1/5/1980	Rabigh	059987767	2091	30	3000	A	10	10	0	10

Fig. 2. *Emp* relation in TTSR approach.

Fig. 2 demonstrates the *Emp* relation after transformation form parametric interval-based representation shown in Fig. 1. The evolution of data in *Emp* relation represented in TTSR approach is shown in Fig. 2. The semantic of update operation follows the temporal update operation introduced in literature [2], [3], [26]. The consequence of updating any time-varying attribute results in inserting a new tuple with the new updated values and new time points as shown in Fig. 2. Deleting any tuple is accomplished by updating  $T_{Is}$  to instant time point. The highlighted tuple with red color in Fig. 2 is an example of logical delete.

**B. TTMR data representation approach**

In TTMR, the relations are represented by: snapshot relation  $R_{TTMR} = (A_K, A_U, T_{ls}, T_{le})$ , for each time varying attribute there are separate relations  $R_{A_{C_1}} = (A_K, A_{C_1}, T_{vs}, T_{ve}) \dots R_{A_{C_i}} = (A_K, A_{C_i}, T_{vs}, T_{ve})$ , and  $R_{LS} = (A_K, T_{ls}, T_{le})$  for the lifespan time that are all in 1NF relations [3], [18], [21], [27], [28]. The relations in Fig. 3 show the representation of *Emp* relation in Fig. 1 using TTMR representation.

<i>Emp</i>				<i>Emp_LS</i>			
SSN	Name	Birth_date		SSN	$T_{ls}$	$T_{le}$	
2091	Ali	1/10/1990		2091	0	$\infty$	
2089	Nashwan	1/5/1980		2089	0	10	

<i>Emp_Address</i>				<i>Emp_Telno</i>				<i>Emp_Superssn</i>			
SSN	Address	$T_{vs}$	$T_{ve}$	SSN	Tel_no	$T_{vs}$	$T_{ve}$	SSN	Superssn	$T_{vs}$	$T_{ve}$
2089	Jeddah	5	4	2089	098778776	0	10	2089	Null	0	6
2089	Rabigh	5	10	2091	059987767	0	$\infty$	2089	2091	7	10
2091	Rabigh	0	$\infty$					2091	Null	0	$\infty$

<i>Emp_Dno</i>				<i>Emp_salary</i>				<i>Emp_Rank</i>			
SSN	Dno	$T_{vs}$	$T_{ve}$	SSN	Salary	$T_{vs}$	$T_{ve}$	SSN	Rank	$T_{vs}$	$T_{ve}$
2089	20	0	8	2089	2000	0	2	2089	A	0	10
2089	30	9	10	2089	2500	3	9	2091	C	0	$\infty$
2091	10	0	$\infty$	2089	3000	10	10				
				2091	13000	0	$\infty$				

Fig. 3. Emp relation in TTMR approach.

The temporal relation schema (*Emp*) in Fig. 3 that is corresponding to  $R_{TTMR}$  in TTMR, is decomposed into  $i + 2$  relations, where  $i$  (number of time-varying attributes) is equal to 6 and the 2 other relations are the lifespan relation and the relation that holds the time-invariant attributes. The 6 relations corresponding to each time-varying attribute that will be used to record the valid-time of the time-varying attributes in *Emp*. The lifespan relation will be used to track the changes of the lifespan of the objects in *Emp*. Finally, the non-temporal database relation is used to record the data of non-time-varying attributes.

**C. TTHR data representation approach**

In TTHR the general representation of  $R_T$  (temporal relational schema) is accomplished as two relations namely,  $R_T$  and  $R_T\_VT$ , Where  $R_T = (A_K, A_U, A_C, A_T)$ , and  $R_T\_VT$  is a new auxiliary relation schema that is created as  $R_T\_VT = (A_K, Att\_index, \alpha, T_{vs}, T_{ve})$ . Semantically the attributes of  $R_T\_VT$  have the following meaning, Att\_index: is a variable to identify the time-varying attribute  $A_{C_m}$  which begins updated such that  $1 \leq m \leq j$ .  $\alpha$  is a new attribute that corresponds to attribute Updated\_V as shown in Fig. 4. This attribute stores the updated value of any attribute in  $A_C$  set.  $T_{vs}$ : represents the Valid Start Time (VST).  $T_{ve}$ : represents the Valid End Time (VET).

The purpose of this representation is to keep the latest (current snapshot data) updated data in one relation  $R_T$ , and the historical changes of the validity of the time-varying data in the auxiliary relation  $R_T\_VT$  [22, 29]. A relation instance is denoted by  $r_i$ , and  $r_i\_vt$ , where  $r_i(R_T)$  means  $r_i$  is an instance of  $R_T$ , and  $vt\_r_i(R_T\_VT)$  means  $r_i\_vt$  is an instance of  $R_T\_VT$ . For tuples the symbols  $x, y$  and  $z$  can be used, thus a tuple,  $T_{ls}$  and  $T_{le}$  of that particular object (tuple), whilst the tuple(s)  $r_i\_vt[x] = \langle a_K, Att\_index, \alpha, a_T \rangle$  in the relation instance  $r_i\_vt(R_T\_VT)$  is/are referencing to tuple  $x$  in  $r_i$ .

The tuple(s) in  $r_i\_vt$  consist of the primary key of  $x$ , the identity(Att\_index) of the time-varying attribute in  $x$ , the updated time-varying attributes value  $\alpha$  in  $x$ , and the time validity of the updated attribute  $T_{vs}$  and  $T_{ve}$ . A subset of the domain of lifespan time is associated with each tuple in  $R_T$  shows that the existence of the object recorded by the tuple is true in the modeled reality during each lifespan chronon in that subset. A subset of domain of valid times is associated with each tuple in  $R_T\_VT$ , represents the fact that the tuple  $r_i\_vt[x]$  records the change of the validity of  $a_{C_m}$  in  $x$ . This fact is considered true in the modeled reality, such that the time of validity strictly contained in the time of the lifespan of  $x$ . Thus, the associated time with a tuple in TTHR is interval-based temporal timestamp. The tuples in  $r_i$  are timestamped by the lifespan time of the object denoted by  $t_{ls}$ , whereas the tuples in  $vt\_r_i$  are timestamped by the valid-time denoted by  $t_v$ , both consisting of a temporal chronon in the time dimension spanned by lifespan and valid time.

<i>Emp</i>	1	2	3	4	5	6	7	8		
SSN	Name	Birth_date	Address	Tel_no	Superssn	Dno	Salary	Rank	$T_{ls}$	$T_{le}$
2091	Ali	1/10/1990	Rabigh	029876554	2090	10	17000	C	0	$\infty$
2089	Nashwan	1/5/1980	Rabigh	059987767	2091	30	3000	A	0	10

<i>Emp_VT</i>					
SSN	Att_index	Updated_V	$T_{vs}$	$T_{ve}$	
2089	7	2000	0	2	
2089	3	Jeddah	0	4	
2089	5	Null	0	6	
2089	6	20	0	8	
2089	7	2500	3	9	

Fig. 4. Emp relation in TTHR approach.

The example in the Fig. 4 uses two relations: *Emp*, describing employees information, such that, this relation is corresponding to  $R_T$  in TTHR, and the auxiliary relation *Emp\_VT* that is used to record the changes of the validity of the time-varying attributes in *Emp* as well as the changes of the lifespan of the objects in *Emp*. The different types of attributes

of  $Emp$  and  $Emp\_VT$  are:  $Emp$  relation:  $A_K = \{SSN\}$ ,  $i = 1$ ;  
 $A_U = \{Name, Birth\_date\}$ ,  $n = 2$ ;  
 $A_C = \{Address, Tel\_no, Supervssn, D\_no, Salary, Rank\}$   
 $j=6$ ;  $A_T = \{T_{ls}, T_{le}\}$  and  $Emp\_VT$  relation:  
 $Employees\_VT = \{SSN, Att\_index, Updated\_V, T_{vs}, T_{ve}\}$

As shown in Fig. 4,  $\alpha$  is equivalent to  $Updated\_V$  that stores the old value of the updated time-varying attributes.  $Att\_index$  attribute stores the position of time-varying attributes location in the main relation  $Employees$ , such that the domain ( $Att\_index$ ) =  $\{0,3,4,5,6,7,8\}$ , where 0 is used to index the object's lifespan time and 3, 4, 5, 6, 7, and 8 are used to index the time-varying attributes as shown in Fig. 4. Granularity of chronon is assumed one month for both lifespan time and valid time. Integers are used as timestamp components that can be thought as dates, for example the integer 7 represents the date of 'April 2012'.

#### IV. DISCUSSION OF MEMORY STORAGE COSTS

In this section, we will formalize the storage costs of the three different approaches for representing the interval-based temporal data models in relational framework. The notations uses in this study are given in Table 1. Let  $R_T$  be a temporal relational schema with an arbitrary set of attributes  $\{A_1, A_2, \dots, A_n, T\}$ , where these attributes can be classified into 4 groups: **key** attributes, Time-invariant attributes (**Unchangeable**), Time-varying attributes (**Changeable**), and **Timestamps** (temporal) attributes. These groups can be represented by  $K, U, C$ , and  $T$  respectively.

Thus the schema of temporal relation can be redefined as  $\{A_K, A_U, A_C, A_T\}$ , where

$$A_K = \{A_{K_1}, A_{K_2}, \dots, A_{K_j}\}$$

$$A_U = \{A_{U_1}, A_{U_2}, \dots, A_{U_n}\}$$

$$A_C = \{A_{C_1}, A_{C_2}, \dots, A_{C_i}\}$$

$$A_T = \{A_{T_1}, A_{T_2}\}$$

**Definition 1:** The cost of different attribute types are defined as:

$$\text{cost}(A_k) = \sum_{i=1}^j \text{cost}(A_{k_i}) = K \text{ byte} \quad (1)$$

$$\text{cost}(A_U) = \sum_{i=1}^n \text{cost}(A_{U_i}) = U \text{ byte} \quad (2)$$

$$\text{cost}(A_C) = \sum_{m=1}^i \text{cost}(A_{C_m}) = C \text{ byte} \quad (3)$$

$$\text{cost}(A_T) = \sum_{i=1}^2 \text{cost}(A_{T_i}) = T \text{ byte} \quad (4)$$

**Definition 2:** The update frequency of time-varying attributes  $A_{C_m} \in \{A_{C_1}, A_{C_2}, \dots, A_{C_i}\}$  in a period of time is calculated as:

$$f(A_C) = \sum_{m=1}^i f(A_{C_m}) = \delta \text{ times} \quad (5)$$

TABLE I. NOTATIONS

Symbol	Meaning
$R_T$	A temporal relational schema with an arbitrary set of attributes $\{A_K, A_U, A_C, A_T\}$
$A_K$	Set of key attributes
$A_U$	Set of Time-invariant attributes( <b>Unchangeable</b> )
$A_C$	Set of Time-varying attributes ( <b>Changeable</b> )
$A_T$	The interval-based timestamp attribute.
$J$	total number of key attributes ( $A_K$ )
$N$	total number of Time-invariant attributes
$I$	total number of Time-varying attributes ( $A_C$ )
$f(A_{C_m})$	Update frequency of $m$ -th time-varying attribute. Such that $A_{C_m} \in \{A_{C_1}, A_{C_2}, \dots, A_{C_i}\}$ and $1 \leq m \leq i$
$S(A_{\chi\gamma})$	A function to be defined on all the attributes in $R_T$ , where $S(A_{\chi\gamma})$ returns the size of attribute $A_{\chi\gamma}$ in bytes. $\chi \in \{K, U, C, T\}$ and $\gamma \in \{1, 2, \dots, j\}$ (key attributes), $\gamma \in \{1, 2, \dots, n\}$ (time-invariant attributes), $\gamma \in \{1, 2, \dots, i\}$ (time-varying attributes) or $\gamma \in \{1, 2\}$ (timestamping attributes)
$Cost(A_\chi)$	A function to be defined on the subset attributes $\chi$ , where $\chi \in \{K, U, C, T\}$ and return the size of the attributes group in byte.
$Cost(z)$	The cost of a tuple(row) $z$ in relation instance $r_t$ is the summation of the cost of all subsets attributes equals to $\text{cost}(A_K) + \text{cost}(A_U) + \text{cost}(A_C) + \text{cost}(A_T)$ .
$K$	Cost of <b>key</b> attributes
$U$	Cost of Time-invariant attributes( <b>Unchangeable</b> )
$C$	Cost of Time-varying attributes ( <b>Changeable</b> )
$T$	Cost of <b>Timestamps</b> (temporal) attributes
$x$	A tuple in a temporal relation
$z$	A tuple in a temporal relation

#### A. TTSR data representation approach

The temporal relation in TTSR can be represented as:

$$R_{TTSR} (A_{k_1}, \dots, A_{k_j}, A_{U_1}, \dots, A_{U_n}, A_{C_1}, \dots, A_{C_m}, A_T)$$

To calculate the memory storage efficiency of interval-based temporal database relation represented by TTHR approach, a general formula is constructed for calculating the size of a single tuple in a temporal relation. The cost of representing one tuple  $x$  in relation instance  $r(R_{TTSR})$  is calculated as:

$$\text{Cost}(x) = \text{cost}(A_K) + \text{cost}(A_U) + \text{cost}(A_C) + \text{cost}(A_T) \quad (6)$$

$$= K + U + C + 2T \text{ byte}$$

as stated in (1), (2), (3) and (4)  $\text{cost}(A_T) = 2T$ , because the tuple in TTSR will be timestamped by valid time and lifespan time. The cost of storing the history of the changes of  $A_C$  with  $f(A_C) = \delta$  times in a period (lifespan interval) of time  $\lambda$  is calculated as:

$$= \delta(K + U + C + 2T) \quad (7)$$

An update in any  $A_C$  requires the insertion of a new row with all attributes. Using (6) and (7), the memory storage cost of one object represented by TTSR approach can be defined as:

$$\text{Cost}(TTSR) = (K + U + C + 2T) + \delta(K + U + C + 2T) \quad (8)$$

### B. TTMR data representation approach

The temporal relation in TTMR is represented as:

$$R_{TTMR} (A_{k1}, \dots, A_{kj}, A_{u1}, \dots, A_{un})$$

$$R_{A_{c1}} (A_{k1}, \dots, A_{kj}, A_{c1}, A_T)$$

$$R_{A_{c2}} (A_{k1}, \dots, A_{kj}, A_{c2}, A_T)$$

$$R_{A_{c3}} (A_{k1}, \dots, A_{kj}, A_{c3}, A_T)$$

$$R_{A_{c4}} (A_{k1}, \dots, A_{kj}, A_{c4}, A_T)$$

.....

$$R_{A_{ci}} (A_{k1}, \dots, A_{kj}, A_{ci}, A_T).$$

To calculate the memory storage efficiency of interval-based temporal database relation represented by TTHR approach, a general formula is constructed for calculating the size of a single tuple in a temporal relation. The cost of storing one tuple  $x$  in relation instance  $r(R_{TTMR})$  is calculated as stated in (1), (2), (3) and (4), as follows:

$$\begin{aligned} \text{Cost}(x) &= \text{Cost}(A_k) + \text{Cost}(A_u) + \text{Cost}(A_T) + \\ &\quad \sum_{m=1}^i [S(A_{c_m}) + \text{Cost}(A_k) + \text{Cost}(A_T)] \\ &= K + U + T + \sum_{m=1}^i S(A_{c_m}) + \sum_{m=1}^i K + \sum_{m=1}^i T \\ &= (K + U + T) + i(K + T) + \sum_{m=1}^i S(A_{c_m}) \end{aligned}$$

Since

$$\begin{aligned} \text{Cost}(A_C) &= i(K + T) + \sum_{m=1}^i S(A_{c_m}) \\ &= i(K + T) + C \end{aligned}$$

Then the  $\text{Cost}(x)$  can be represented as:

$$\begin{aligned} \text{Cost}(x) &= K + U + T + C + i(K + T) \\ &= K(i+1) + U + C + T(i+1) \text{ byte} \quad (9) \end{aligned}$$

The variable  $i$  represents the total number of time varying attributes  $A_C$ . The cost of storing the history of the changes of each  $A_{c_m}$  with  $f(A_{c_m}) = \delta_m$  times in a period/interval of time  $\lambda$  can be calculated as:

$$\begin{aligned} &= \sum_{m=1}^i \delta_m (S(A_{c_m}) + K + T) \\ &= \sum_{m=1}^i \delta_m S(A_{c_m}) + (K + T) \sum_{m=1}^i \delta_m \end{aligned}$$

Since  $\sum_{m=1}^i \delta_m = \delta$ , as in Eqns. 5, then the equation becomes:

$$= \sum_{m=1}^i \delta_m S(A_{c_m}) + \delta(K + T) \quad (10)$$

Using (9) and (10), the memory storage cost of one object represented by TTMR approach can be defined as:

$$\begin{aligned} \text{Cost}(TTMR) &= K(i+1) + U + C + T(i+1) + \\ &\quad \sum_{m=1}^i \delta_m S(A_{c_m}) + \delta(K + T) \quad (11) \end{aligned}$$

### C. TTHR data representation approach

In TTHR model, the temporal relation schema is represented by  $R_{TTHR}$  and  $R_{VT}$  as shown below:

$$R_{TTHR} (A_{K_1}, \dots, A_{K_j}, A_{U_1}, \dots, A_{U_n}, A_{C_1}, \dots, A_{C_m}, A_T),$$

$$R_{VT} (A_{K_1}, \dots, A_{K_j}, Att\_index, \alpha, A_T).$$

To calculate the memory storage efficiency of interval-based temporal database relation represented by TTHR approach, a general formula is constructed for calculating the size of a single tuple in a temporal relation. The cost of representing one tuple  $x$  in relation instance  $r(R_{TTHR})$  is calculated as:

$$\begin{aligned} \text{Cost}(x) &= \text{cost}(A_k) + \text{cost}(A_u) + \text{cost}(A_c) + \text{cost}(A_T) \\ &= K + U + C + T \text{ byte} \quad (12) \end{aligned}$$

As stated in (1), (2), (3) and (4), the cost of storing the history of changes of  $A_C$  with  $f(A_C) = \delta$  times in a period/interval (lifespan interval) of time  $\lambda$  can be calculated as:

$$= \delta(K + S(Att\_index) + S(\alpha) + T),$$

since the size of  $Att\_index = 1$ , and

$$S(\alpha) = \text{Max}(S(A_{C_1}), S(A_{C_2}), \dots, S(A_{C_i})) \quad , \quad \text{let}$$

$$\beta = S(\alpha), \text{ then} \quad (13)$$

Such that,  $Att\_index$ : is an attribute to index the time-varying attributes with one byte size.  $\alpha$ : is a new added attribute of variant data type to hold data from different types. Its size is assumed to be the same size of the largest field size in  $A_C$ . The size of  $\alpha$  in byte is  $S(\alpha) = \text{Max}(S(A_{C_1}), S(A_{C_2}), \dots, S(A_{C_i}))$ , using (12) and (13), the memory storage cost of one object represented by TTHR approach can be defined as:

$$\text{Cost}(TTHR) = (K + U + C + T) + \delta(K + 1 + \beta + T) \quad (14).$$

## V. COMPARISONS OF MEMORY STORAGE COST AND RESULT ANALYSIS OF THE THREE APPROACHES

In this section, we will mimic the storage cost of the three models based on various settings of the parameters that have direct impact to the temporal data storage. The Default values are initiated with consideration of general cases as follows:  $K =$

9,  $U=110$ ,  $C=37$ ,  $T=20$ ,  $\beta=9$ . Fig. 5 shows the memory storage cost for the initial values for the different parameters that construct the temporal relation. For these values, TTSR-based approach shows worse storage costs comparing to TTMR-based and TTHR-based approaches. However, the graph shows a positive indication that TTHR can be used as an efficient storage that is better than TTMR-based approach until the value of  $\delta = 40$ . After this point it seems that both TTHR and TTMR have the same storage efficiency.

Fig. 6 shows the storage costs of the temporal relational approach after freezing all the parameters and varying the sizes of the time-varying attributes ( $C$ ). For these values, TTSR-based approach shows worse storage costs comparing to TTMR-based and TTHR-based approaches. However, the graph shows a positive indication that TTHR can be used as an efficient storage that is better than TTMR-based approach until the value of  $C = 150$  byte. After this point it seems that both TTHR and TTMR have the same storage efficiency.

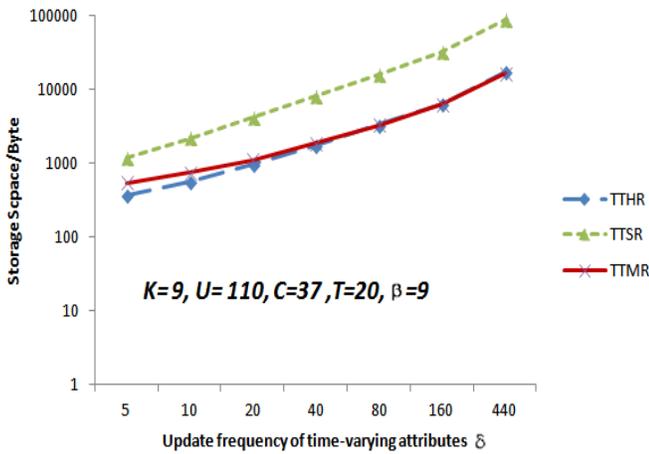


Fig. 5. Storage cost for the update frequency ( $\delta$ ) variations.

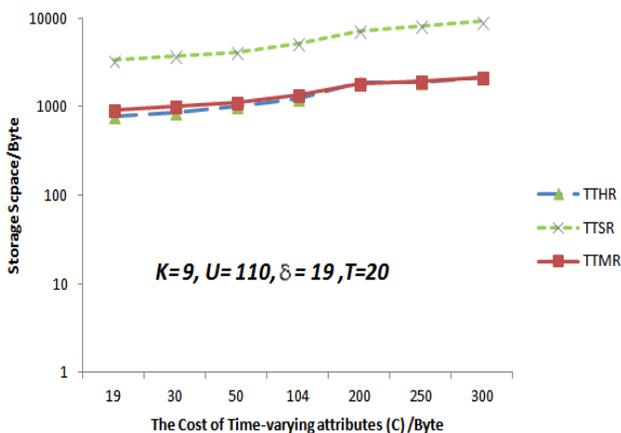


Fig. 6. Storage Cost for the Time-varying attributes' size ( $C$ ) variations.

Fig. 7 shows the storage efficiency after freezing all the parameters and varying the sizes of key attributes ( $K$ ) value variations.

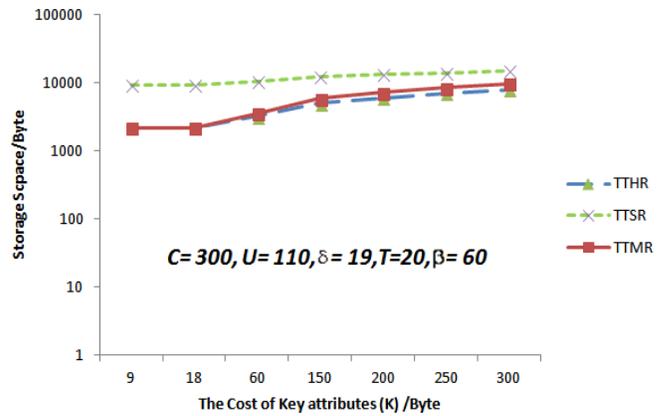


Fig. 7. Storage Cost for the Key attributes' size ( $K$ ) variations.

We increase  $K$  value from 9 to 300 bytes. As we can see, the TTHR-based approach shows the best storage efficiency than the others. However, it is shown that the difference of storage efficiency is marginal between the TTHR-based approach and the TTMR-based approach.

A similar study for calculating the efficiency of memory storage has been done by Atya in [2], this study compared the Snodgrass model (which is under TTMR-based approach in our study) with Tansel N1NF relational nested model. A study by Noh in [23] has introduced a new platform for modeling temporal database under XML-based platform. He compared the relational model as he named it (which is under TTSR-based approach in our study) with XML-based, and object-oriented based approach.

## VI. CONCLUSION

A new approach for representing temporal database in relational data model has been demonstrated in this research work. A comparison study of the proposed model (TTHR) with the main models in literature (TTSR and TTMR) with respect to the memory storage efficiency has been mathematically illustrated. To measure the storage costs, we have established a mathematical model (formulas) for the three approaches. The measurement of the performance is represented by the size of the whole stored temporal data as stated in [22], [29], [31]. It has been proved that TTHR has achieved significant saving in memory storage that ranges between 68%-81% over TTSR approach, and 10%-32% over TTMR. The memory storage save is based on the average change of the time varying attributes [29], [30], [31]. A validation and verification study of the correctness and the expressiveness of TTHR model has been depicted in [32]. Finally, TTHR mimics TTMR in data representation by removing the needless redundancy of data. Moreover, TTHR mimics TTSR in representing the current valid data in one relation, to benefit from querying the current snapshot data which costs a lot in TTMR as stated in [22].

## ACKNOWLEDGMENT

This paper was supported by the Deanship of Scientific Research (DSR), King Abdulaziz University. The authors, therefore, acknowledge with thanks to DSR's technical support.

REFERENCES

- [1] Anselma, L., Bottrighi, A., Montani, S., & Terenziani, P. Extending BCDM to cope with proposals and evaluations of updates. IEEE T KNOWL DATA EN, IEEE Transactions on 2013; 25(3), 556-570.
- [2] Atay, C. An attribute or tuple timestamping in bitemporal relational databases. TURK J ELECTR ENG CO 2016; 24: (pp. 4305 – 4321). doi:10.3906/elk-1403-39
- [3] Elmasri, R., and Navathe. Database Systems 6<sup>th</sup> edition. Pearson, 2011..
- [4] Arora, S. A comparative study on temporal database models: A survey. In Advanced Computing and Communication (ISACC), International Symposium on 2015; (pp. 161-167). IEEE.
- [5] Jensen, C. S., Snodgrass, R. T., & Soo, M. D. *The tsq2 data model*. Springer US 1995; (pp. 157-240).
- [6] Tansel, A. Adding time dimension to relational model and extending relational algebra. INFORM SYST 1986; 11(4), 343-355.
- [7] Tansel, A., and Snodgrass, R. *Temporal databases: Theory, design and implementation*. Redwood City, CA: Benjamin Cummings, 1993.
- [8] Tansel, A. Temporal data modeling and integrity constraints in relational databases. In *Computer and Information Sciences-ISCIS 2004*; (pp. 459-469). Springer Berlin Heidelberg.
- [9] Tansel, A. On handling time-varying data in the relational data model. INFORM SOFTWARE TECH 2004; 46(2), 119-126.
- [10] Tansel, A. Modeling and Querying Temporal Data, Idea Group Inc 2006.
- [11] Terenziani, P. Coping with events in temporal relational databases. . IEEE T KNOWL DATA EN, IEEE Transactions on 2013; 25(5), 1181-1185.
- [12] Garani, G. (2003). *A temporal database model using nested relations* (Doctoral dissertation, BIRKBECK COLLEGE).
- [13] Garani, G. A generalised relational data model. International Journal of Computer Systems Science and Engineering 2007; 11, 43-59.
- [14] Kvet, M., Matiako, K. and Kvet, M., Transaction management in fully temporal system. In Computer Modelling and Simulation (UKSim), 2014 UKSim-AMSS 16th International Conference on (pp. 148-153). IEEE.
- [15] Jones S., Mason P., and Stamper, R. A Relational Specification Language for Complex Rules. Information Systems 1979; 4(4):293{305}.
- [16] Ariav, G. A temporally oriented data model. ACM T DATABASE SYST 1986; 11(4), 499-527.
- [17] Sarda, N. L. Algebra and query language for a historical data model. The Computer Journal 1990; 33(1), 11-18.
- [18] Snodgrass, R. The Temporal Query Language TQUEL. ACM T DATABASE SYST 1987; 12(2):247{298, June 1987.
- [19] Steiner, A. A Generalization Approach to Temporal Data Models and their Implementations. A dissertation submitted to the SWISS FEDERAL INSTITUTE OF TECHNOLOGY, ZURICH, 1998.
- [20] Navathe, S. B., and Ahmed, R. A temporal relational model and a query language. INFORM SCIENCES 1989; 49(1), 147-175.
- [21] Noh, S.Y., and Gadia, S. K. *Benchmarking temporal database models with interval-based and temporal element-based timestamping*. J SYST SOFTWARE 2008; 81(11):1931–1943.
- [22] Alromema, N., Rahim, M.S.M. and Albidewi, I., 2016. Performance Evaluation of Attribute and Tuple Timestamping In Temporal Relational Database. International Journal of Computer Science and Information Security, 14(9), p.374.
- [23] Noh, S.Y., Gadia, S.K. and Jang, H. Comparisons of three data storage models in parametric temporal databases. J CENT SOUTH UNIV 2013;20(7), pp.1919-1927.
- [24] Allen, J.F. Maintaining knowledge about temporal intervals. *Communications of the ACM* 1983; 26(11), 832-843.
- [25] Yang, C., Wang, X., Zhang, M., Zheng, R., Wei, W. and Lou, Y. Standardization on bitemporal information representation in BCDM. In Information and Automation, 2015 IEEE International Conference on (pp. 2052-2057). IEEE.
- [26] Snodgrass, R. *Developing Time-Oriented Database Applications in SQL*, 1<sup>st</sup> edition, Morgan Kaufmann Publishers, Inc., San Francisco, 2000.
- [27] McKenzie Jr., and Snodgrass, R. T. Evaluation of relational algebras incorporating the time dimension in databases. ACM Computing Surveys (CSUR) 1991; 23(4), 501-543.
- [28] Snodgrass, R. Temporal Databases. IEEE Computer 1986; 19(9):35{42}.
- [29] Halawani, S. M., Al-romema, N. A. *Memory Storage Issues of Temporal Database Applications on Relational Database Management Systems*, Journal of Computer Science 2010; 6, (3): 296-304
- [30] Ab Rahman Ahmad, N.A., Rahim, M.S.M. and Albidewi, I., 2015. Temporal Database: An Approach for Modeling and Implementation in Relational Data Model. Life Science Journal, 12(3).
- [31] Halawani, S.M., AlBidewi, I., Ahmad, A.R. and Al-Romema, N.A., 2012. Retrieval optimization technique for tuple timestamp historical relation temporal data model. Journal of Computer Science, 8(2), p.243.
- [32] Alromema, N.A., Rahim, M.S.M. and Albidewi, I., 2016. Temporal Database Models Validation and Verification using Mapping Methodology. VFAST Transactions on Software Engineering, 11(2), pp.15-26.

# Semantic based Data Integration in Scientific Workflows

M. Abdul Rehman, Jamil Ahmed, Ahmed Waqas, Ajmal Sawand  
Department of Computer Science, Sukkur IBA University  
Sukkur, Pakistan

**Abstract**—Data Integration has become the most prominent aspect of data management applications, especially in scientific domains like ecology, biology, and geosciences. Today's complex scientific applications and the rise of diverse data generating devices in scientific domains (e.g. sensors) have made data integration a challenging task. In response to these types of challenges, data management applications are providing groundbreaking functionalities which come at the price of high complexity. This paper presents a semantic data integration framework which is based on the exploitation of ontologies. Exploiting a Description Logics formalism and associated reasoning procedures, the framework is able to handle heterogeneous formats and different semantics. Besides an in-depth discussion of the ontology-based integration capability, the paper also discusses a brief overview of the system architecture and its application in a real world scenario taken from ecological research.

**Keywords**—Data integration; scientific workflows; ontology; data semantics; data management

## I. INTRODUCTION

In order to understand complex scientific scenarios such as the world climate or the impact of the decreasing number of one species on others [1] lots of data are required. These data are usually not coming from one institution only but from many heterogeneous sources that need to be integrated [2]. Indeed the problem is not the availability of data but how to relate and interpret data correctly. The rise of data generating devices in scientific applications such as sensor networks has also made integration tasks more challenging. These sensors produce streams of unstructured raw data at different temporal and spatial granularities. Thus, before being used in any climatic application, these data need to undergo several processing steps of transformation and integration since such datasets are highly heterogeneous in terms of format, syntax, structure and semantics.

In such type of scenarios, besides a powerful semantic data integration capability, integration systems must provide comprehensive data management solutions that handle assorted formats and cares about data structures when data flow from a source to a sink. Furthermore besides the detection and resolution of conventional semantic conflicts [3] these systems must deal with issues like:

- *Information transformation* – Based on discovered mappings, information expected by the sink does not correspond to the information provided by the source. For instance, the source stores a *deviceId* while the sink

requires *deviceName*; then these data must be transformed such that the same information is dealt with by both source and sink.

- *Missing and incomplete data* – when the target requires information which is not available in the source. For instance the source contains only a *deviceId* whereas the target also expects information about the location of that device (*locationId*, *locationName*).

The integration of data is often implemented in two ways which are sometimes overlapping. First, the use of standards is facilitated. Examples of such standards are the ABCD schema (Access to Biological Collection Data, [1]), the SEEK observation ontology [4], and biomedical ontologies [5]. A common pitfall of standards lies in the standardization process itself. Many stakeholders try to reflect their interests which can either cause the standard to be very general or very specific. In the first case it is hard to apply standards to specific problems because they are getting too complex (hundreds of elements), in the second case many competing standards arise. But standards – and this is their main benefit – fix semantics and syntax of data to a very high extend in a specification.

A second approach towards the data integration issue is the provision of handwritten wrappers which perform integration tasks on their own. This kind of a solution is still widespread even scientists start to realize that these wrappers are only useful as long as the corresponding developers are still available. Wrappers also fix the syntax and semantics of data, but unlike standards syntax and especially semantics are not directly accessible because they are hard-coded into the application. Looking up the syntax is possible by examining data provided from the wrapper. But the meaning of data is still hidden.

Both presented approaches work and provide some kind of a solution – however they share a common pitfall; the syntax and even more the semantics of data often depend on the application or the users' perception. Thus a method for data integration must provide the freedom of choosing one specific semantic and should not require the user to change standards or adapt wrappers permanently.

It is widely recognized that ontologies play a central role in modern scientific applications [5] [6] since their use is considered a possible solution for semantic based integration [2] [7]. Nevertheless a clear methodology for setting up the data integration task (cf. Section 2.2) also plays a vital role since usability and adaptability are determined by it.

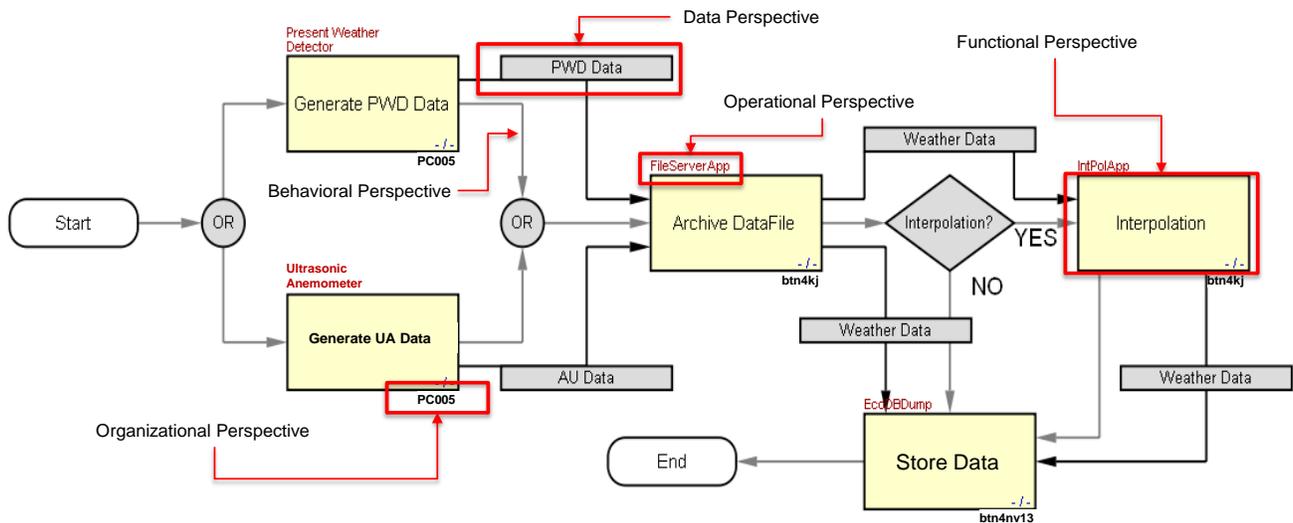


Fig. 1. POPm workflow model for the acquisition of sensor data within BayCEER Lab.

Workflow technology has also contributed a lot towards the integration of scientific data. In recent years many scientific workflow systems like Kepler [8], Taverna [9] and Triana [10] have come up with the capability to integrate scientific data based on ontologies. The lack of a clear method as well as the inability to separate integration steps from domain specific analysis steps (cf. Section 2.1) makes these systems hard to be utilized by a normal domain user. Data integration is implemented in these systems as an atomic step within a workflow – even though it can comprise very complex actions.

The proposed system promises to offer an end-to-end solution (i.e. from data source to data sink) for the data management issues discussed above including powerful and real semantic data integration through the exploitation of ontologies. Furthermore, a well-structured methodology facilitates normal users to manage their specific data integration scenario in a better manner. The main contribution in this paper is to present the semantic integration feature of DaltOn system as well as the methodology for setting up integration tasks.

The remainder of this contribution is structured as follows: Section 2 introduces the overall method which is already applied in several real-world use cases. Section 3 discusses the core algorithm of the semantic integration component of DaltOn together with its foundations. Section 4 relates the work to other approaches and Section 5 finally summarizes and concludes the contribution.

## II. METHOD AND ARCHITECTURE OF DALTON

In the following sections a method for data integration within scientific workflows based on the DaltOn framework is presented.

### A. Overview and Motivating Scenario

**POPm** (Perspective Oriented Process Modelling) [11] is a paradigm for modelling processes and/or workflows. Within the POPm paradigm, each modelling construct comprises several orthogonal building blocks, called perspectives. Thus, a modelling construct can be specified by defining different

perspectives. There are five main perspectives (shown in Fig. 1) that provide a strong foundation for a process modelling language. The *Functional Perspective* determines the existence and purpose of a process step. The *Operational Perspective* is used to specify the application, service or tool which is required during enactment of a work step. The *Behavioral Perspective* is a mean to determine the execution order of the work steps. The *Organizational Perspective* is used to introduce agents who are eligible or responsible for performing certain work steps. The *Data Perspective* identifies data items and their flow in a process. This latter perspective is of special interest since all data management related issues are contained within this perspective.

i>PM [12] is a graphical tool built upon the POPm paradigm. It allows users to develop process / workflow models by specifying the above discussed perspectives

A **Motivating scenario** is taken from meteorological research and is currently carried out by The Data Group of the Bayreuth Center of Ecology and Environmental Research [13]. The main purpose is to retrieve data from various sensor devices, store it at an intermediate place for archival and finally dump the data into a central database at the institute.

Fig. 1 shows a POPm-based workflow model composed by a domain user for this scenario. The first two steps of the workflow describe the generation of sensor data either in the PWD (Present Weather Detector) or UA (Ultrasonic Anemometer) format depending on the actual sensor. Then these data are transmitted over a serial line to the archival step *Archive DataFile* which is describing the process of storing the sensor data in a file. After a successful data archival, data are usually stored in the central database called *EcoDB* within the *Store Data* work step. The work step *Interpolation* is optional and manipulates data whenever needed. For demonstrating the capabilities of the methodology, the main focus would be on the management of data as it occurs in between the two steps *Archive DataFile* and *Store Data* as shown in Fig. 1.

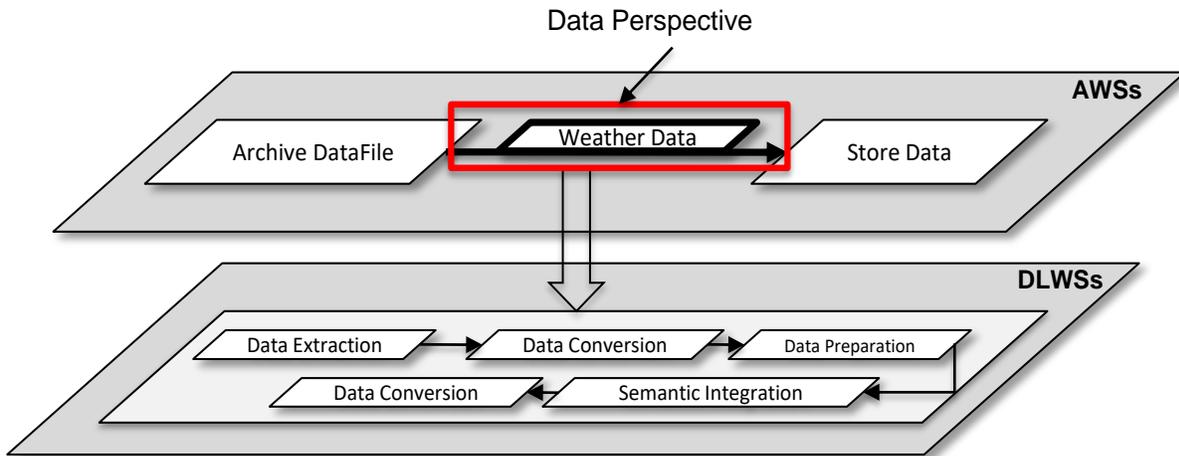


Fig. 2. Classification of work steps in a scientific workflow.

Before storing data in the database, it needs to be passed through some preparation operations since the raw sensor data does not fit in the structure and semantics of the database – sensor devices generate data in their own proprietary formats (here “PWD” or “UA”) and also might use different interpretations. Thus the scenario poses many data management related issues:

- Physical data transport from a source (here: file server) to a sink (here: database)
- Management of data formats (here PWD and AU)
- Validation of data: Sometimes files are truncated due to an interruption in the file transfer from a sensor to the file server.
- Data filtering: This challenge belongs to the process of discarding unwanted values from data files during data transmission.
- Data integration: This is a major challenge since data coming from various devices have completely assorted structures and give different data interpretations.

### Separation of workflow steps

In Fig. 2, the upper layer (denoted as ‘AWSs’) shows an extract of the workflow model of Fig. 1 including only the two work steps ‘Archive DataFile’ and ‘Store Data’.

As stated in the previous section, several steps are needed in order to make the data of ‘Archive DataFile’ compatible to ‘Store Data’. Hence, an application consists of two categories of work steps. First, steps which enact domain operations such as data acquisition, data analysis or data storage. Second, steps which are only used for data integration. Up to now these data related work steps were specified explicitly in the workflow, making it more complex. But separating the steps of a process into two categories assures the productivity of the scientific community since normal domain users (scientists) really do not desire to involve into data specific tasks.

The first layer of the approach is called Application Work Steps (AWSs) layer and contains only those steps of the

scientific workflow which are specific to the application but whose nature is not related to pure data management. Examples are ‘Archive DataFile’ and ‘Store Data’. The Data Logistic Work Steps (DLWSs) layer instead contains solely data preparation tasks (such as data integration and conversion). DLWSs are defined in terms of operations provided by DaltOn and incorporated into the scientific workflow instead of being explicitly modelled into it. Together both layers describe a (executable) scientific application. The AWSs layer can be controlled by a normal workflow management system (WfMS); the DaLo-WFs are – due to their nature – also controlled by a WfMS but the execution of these processes is heavily supported by the DaltOn framework.

The data flow perspective of POPM directly corresponds with the definition of DaLo-WFs. It thus wraps up all data management functionality and provides means to define where data resides, where it must be transported to and how it is transformed within scientific workflows.

### B. A Methodology for Data Integration within Scientific Workflows

The methodology is shown in Fig. 3. It is itself described by a process model. Together, this process describes how a scientific application consisting of AWSs and DLWSs has to be developed. Further, the method assigns clear responsibilities: the application workflow with the scientific analysis is defined by scientists (generally: the domain user) and the DaLo-WFs with data integration tasks are set up by data experts. A short description of each step of the process is given in the following.

*Define Application Process:* The aim of this step is to define the domain application by developing a workflow which includes only domain specific work steps (as AWSs layer). Domain users, e.g. scientists, are responsible for enacting or at least for supervising the execution of this step. In general every workflow modelling tool can be used to describe an application process; however, i>PM [12] has been used in this work. In the example scenario, the output of this process step is the workflow model shown in Fig. 1.

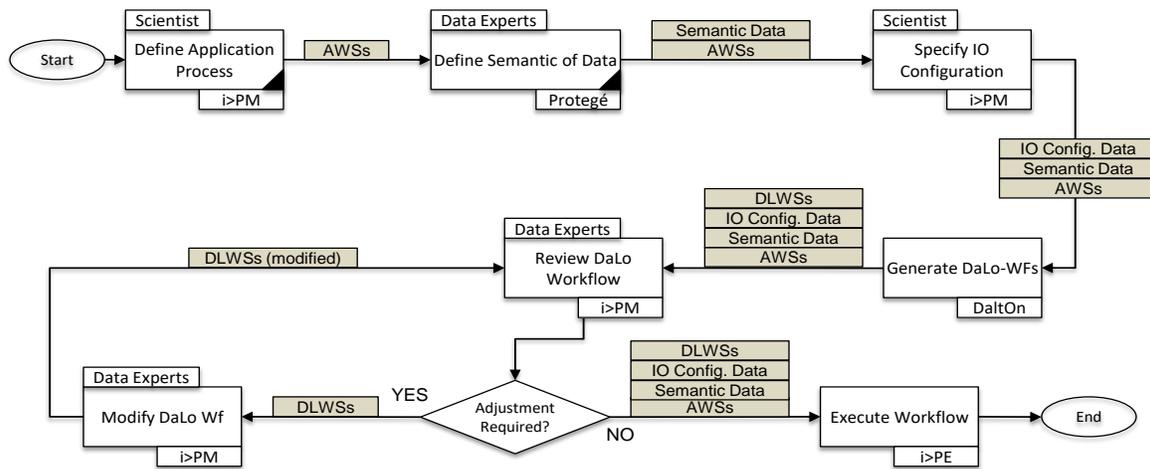


Fig. 3. POPM process of the methodology for the definition of scientific applications.

**Define Semantics of Data:** Data are described semantically by developing ontologies. Ontologies for the use case are shown in Section III. Both, domain users and data experts are responsible for this step. Domain users because only they can describe a common vocabulary and data experts because only they know specific details about applications, data sources and sinks. However, this step is optional since existing ontologies can be re-used.

**Specify IO Configuration:** This step supports the generation of DaLo-WFs taking place in the next step of the methodology. The I/O specification may contain information about the exact location and type of data sources/sinks, a description of data formats for input and output data (for instance data schemata), an assignment of the previously defined local ontologies to steps, applications and agents and finally the definition of criteria applied during extraction and insertion of data. Table 1 shows the specification of the ‘Weather Data’ data container for the example scenario.

**Generate DaLo-WFs:** In this step each data flow of the application process is transformed into single DaLo-WFs that specify single data integration tasks. A first version of a DaLo-WF can automatically be derived by exploiting the specifications provided in the previous step. The DaLo-WF generation is supervised by a data expert and enacted by DaltOn functions. As an output, this step delivers an executable workflow containing both the AWSs and the DLWSs.

**Review DaLo-WFs and Modify DaLo-WFs:** As stated earlier, the generated DaLo-WFs may not be sufficient for

special integration scenarios. Therefore, the DaLo-WFs might have to be adjusted accordingly. Review and modification are again enacted by data experts.

**Execute Workflow:** Finally, the workflow can be executed by a suitable execution environment [14] for a description of such an execution environment). Both types of steps, AWSs and DLWSs, are executed by the same environment. Involving human actors is possible for both types of steps since some applications might require the interaction of process and scientists.

1.1 Architecture of the DaltOn Integration Framework

The architecture of the DaltOn Integration Framework [15][16] follows the approach of separating concerns into single and independent functions. Thus, DaltOn has three major conceptual abstractions, namely Data Provision, Data Operation and Data Integration.

**Data Provision** bundles components which are used for enacting physical data exchange between data sources (data producing steps) and data sinks (data consuming steps). Each of the sub-components of the Data Provision fulfils a specific task: Data Extraction and Selection (DES) cares about the extraction of a (sub-) set of data from a source based on user- and application-specific criteria, Data Transportation (DT) handles physical data transport and Data Insertion (DI) performs insertion of data.

TABLE I. SPECIFICATION OF ‘WEATHER DATA’ DATA CONTAINER

Configuration	Value	
	Source	Sink
Location	<IP address>:<Path to the File>	<DB Connect String>
Format	Pwd	Xml
Criteria	Null	Visibility=<2000
Schema	URI to Eco1.xsd (schema1)	URI to Eco2.xsd (schema2)
Local Ontology	URI to LocOntoEco1.owl (local ontology1)	URI to LocOntoEco2.owl (local ontology2)
Mapping	URI to mapEco1.rdf (mapping1)	URI to mapEco2.rdf (mapping2)
Reference Ontology	URI to RefOntoEco.owl	

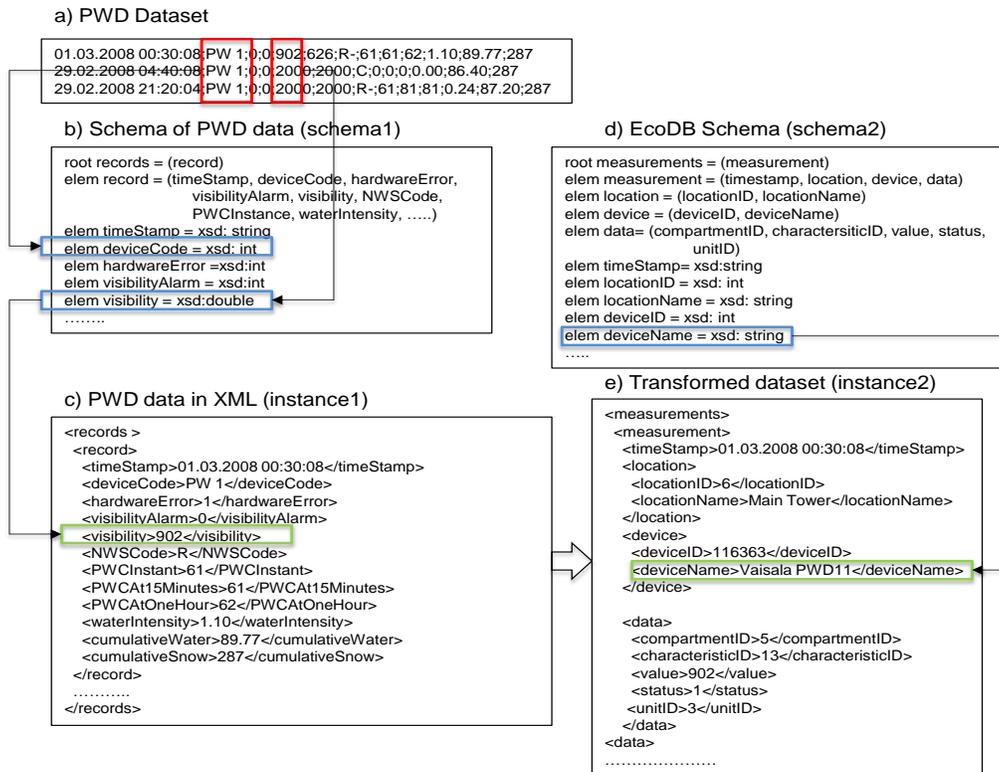


Fig. 4. Excerpt of instance data and schemas for both sides along with actual PWD dataset.

Data Operation encompasses Format Conversion (FC) and Data Preparation (DP); the FC sub-component is carrying out syntactic transformations of data, for instance the conversion of data given in CSV (comma separated values) into a XML representation and back. DP contains functions which can be applied to data such as unit conversions or simple arithmetic operations but is not meant to replace scientific analysis steps.

Data Integration finally is the heart of the DaltOn system [17] that aims at the semantic integration of data. It comprises only one sub-component so far, the Semantic Integration (SI).

Beside these main three abstractions, DaltOn is using wrappers for accessing data sources through a unified interface (but not for executing data integration tasks) and a RDF based (triple) data store as repository.

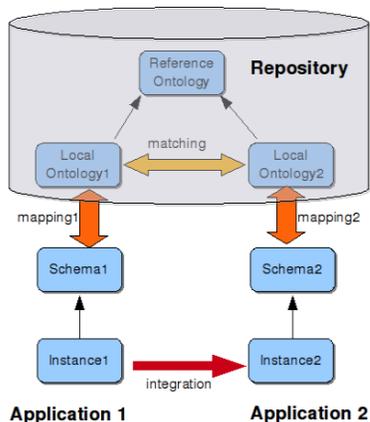


Fig. 5. The architecture of the SI component

### III. SEMANTIC BASED INTEGRATION

Schema mapping is a specification describing how data from a source schema can be mapped to a sink (or target) schema. This is usually considered an essential building block of data exchange and integration solutions.

Schema mappings are usually discovered (semi-) automatically using a match operation that can either corresponds to a structure-based or a semantic-based approach. The feasibility of later approach is supported by the design of DLWSs which require that AWSs share a common domain of discourse. In DaltOn, this aspect is represented with ontologies. A first issue is that several applications may interpret elements of the same domain differently. In such situations, an alignment between the different interpretations needs to be discovered. Based on discovered mappings, DaltOn solves related issues. One of these concerns is information integration. That is, the information expected from the source document does not correspond to the information the source is able to provide. Another important issue deals with missing or incomplete information in the source instance document. Both situations generally prevent the fulfilment of the generation of a source instance document. In order to pursue the integration, the paper proposes a solution based on the exploitation of a repository. The repository is a central place which stores and proposes query facilities to retrieve information related to the domain of discourse.

The example scenario is interesting as it exploits an important set of functionalities available in the SI component, e.g. different kinds of mapping correspondences and repository exploitation.

A. Basic Notions

Basically Ontologies are used to represent the knowledge of a domain in a common way, enabling these to be shared among machines and human beings [18]. Thus, an ontology consists of concepts with relationships between them, which provides a common vocabulary for knowledge to be exchanged between machines and human beings. In order to support reasoning within ontologies, a logical formalism is used, such as Description Logics (DL) [19], as a mean to represent ontologies. This family of formalisms allows the representation and reasoning over domain knowledge in a formally and well-understood way. Central DL notions are concepts (unary predicates) and relationships, also called properties or roles (binary predicates). A key notion in DL is the separation of the terminological (or intensional) knowledge, called TBox, from the assertional (or extensional) knowledge, called ABox. The TBox contains the descriptions of concepts and their relationships in the following form:

$Device \sqsubseteq \forall situatedAt.Location \sqcap$   
 $\exists situatedAt.Location \sqcap$   
 $\forall hasDeviceName.String \sqcap$   
 $\exists hasDeviceName.String$

This description states that the Device concept is defined as being situated in at least one location, and locations only, and has at least one name which must be string of characters (in an OWL serialization this is supported by XML Schema data types). In contrast, ABoxes contain assertions of concepts and their roles in the following form:

`Device(device_116366),`  
`Location(location_3),`  
`hasDeviceName(device_116366,`  
`"Vaisala QLi50 2")`  
`situatedAt(device_116366, location_3)*`

These assertions state that objects with identifiers 'device\_116366' and 'location\_3' are instances of respectively the 'Device' and 'Location' concepts. These two objects are related by the 'situatedAt' object property. Finally, the 'device\_116366' object is related to the value 'Vaisala QLi50 2' via the 'hasDeviceName' data type property.

A TBox and an ABox together denote a Knowledge Base (KB), denoted as  $KB = \langle TBox, ABox \rangle$ .

B. Data On SI Component Architecture

The main objective of SI is to generate a valid input document for the target step of an Application Workflow. The Fig. 5 presents the details of this component's architecture. This architecture is based on the set of documents each DaLo-WF application can access. This set comprises four kinds of documents:

- 1) An instance document which corresponds to the output document of Application1 (produced by the source step), respectively the input document of Application2 (consumed by the sink step).
- 2) A schema associated to each instance document.
- 3) A mapping between elements of the schema to elements of a local ontology.
- 4) A local ontology which supports the particular interpretation of each concept in an application.

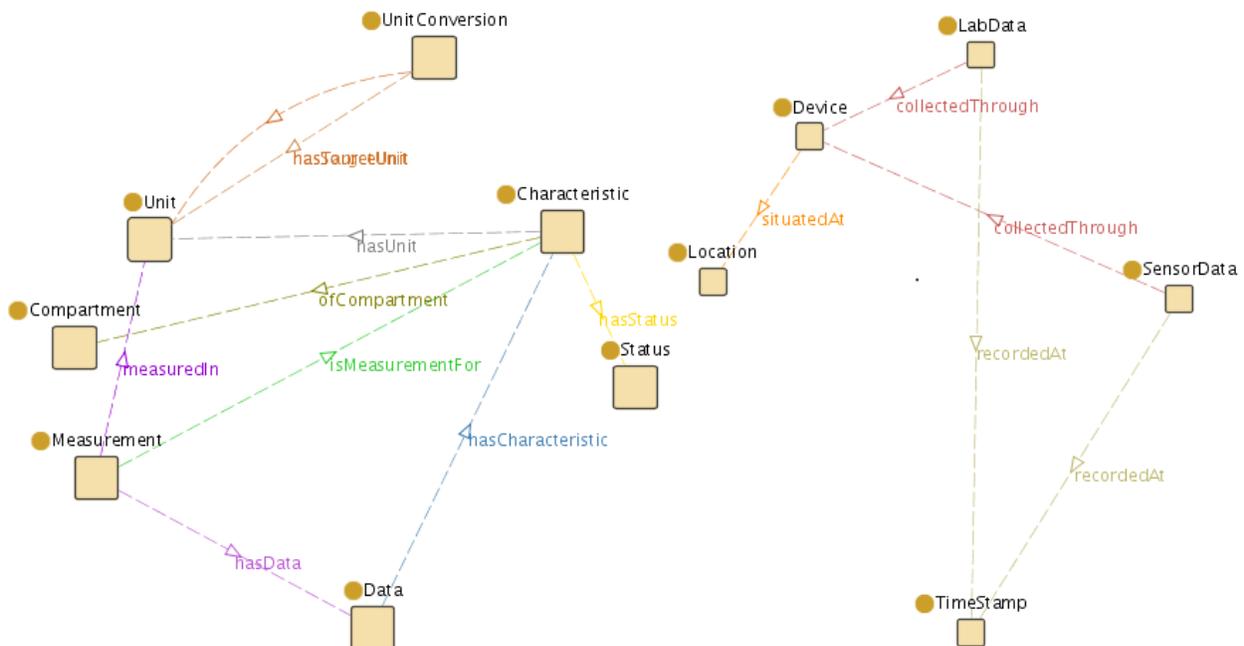


Fig. 6. Concepts and roles of the reference ontology in the meteorological example.

The upper part of Fig. 5 emphasizes two other components:

1) A reference ontology which provides a common vocabulary to the local ontology. This approach makes the local ontologies comparable and enables to process matches between concepts.

2) The repository is responsible for the storage of the knowledge bases associated to the application domain. It also stores the mappings that are being discovered by the matching solution.

The role of an application instance and schema document is obvious in the context of a DaLo-WF. They are usually created by the application developer and come at no extra cost. These documents are expressed in XML and XML Schema (henceforth XSD) respectively. Fig. 4 presents an extract from instance1, instance2, schema1 and schema2.

The mappings, ontologies and ABox assertions contained in the repository impose extra work from the application designers. Nevertheless, the task of developing these documents is limited due to the following: (i) generation of a single reference ontology is generally sufficient, (ii) reuse of local ontologies among several DaLo-WFs is generally possible, (iii) low expressivity of the reference ontology, (iv) use of adapted tools which simplify the creation of these documents.

Concerning aspects (i) and (ii), as per experiences in using DaltOn in medicine, biological and ecological domains emphasize that usually one unique reference ontology is sufficient for all DaLo-WFs of an application. The design of this reference ontology can be facilitated by exploiting existing domain ontologies. The reference ontology does not need all the expressiveness proposed by some well-known ontologies in scientific domains.

Concerning (iii), the expressive power of the local and reference ontologies are not the same. The reference ontology provides a common vocabulary on the domain of discourse. This common vocabulary enables schema mapping to be generated. Fig. 6 presents a graph of the reference ontology in

the meteorological use case, developed using Protégé tool [25]. This graph presents concepts as nodes, roles as labeled edges. For readability reasons, subsumption relationships are not depicted.

A local ontology implements the local interpretation to the concepts of the reference ontology and also provides the way to include new concepts defined with respect to the concepts and roles of the reference ontology. Example 1 presents an extract of the concept definitions of local ontology1 in the meteorological use case.

**Example 1** Concept definitions of local ontology1

```
loc1.HardwareErrorCode ⊆ Status
loc1.TimeStamp ≡ TimeStamp
loc1.PresentWeatherData ⊆ SensorData
loc1.SensorData ⊆ Data ⊓
  ∀recordedAt.TimeStamp ⊓
  ∃recordedAt.TimeStamp
  ⊓ ⊓
  ∀collectedThrough.Sensor ⊓
  ∃collectedThrough.Sensor
  ∀hasStatus.Status ⊓ ∃hasStatus.Status
loc1.CumulativeSnow ⊆ Characteristic
loc1.Status ≡ Status
loc1.NWS ≡ NWS
loc1.Sensor ≡ Sensor
```

Concerning (iv), the design of the different ontologies (reference and local ones) as well as the generation of reference ontology concept and role assertions, stored in the repository, are facilitated by the use of a Protégé plug-in named DBOM [20], [21]. This plug-in eases the creation of knowledge bases expressed in DL from relational databases. For instance, in the meteorological use case, the meteorologists provided us with databases containing domain specific information about location of sensors, devices, units used by these devices, etc. Using the DBOM plug-in the system is able to create a reference ontology serialized in OWL and at the same time to generate a valid ABox which was later integrated in the repository.

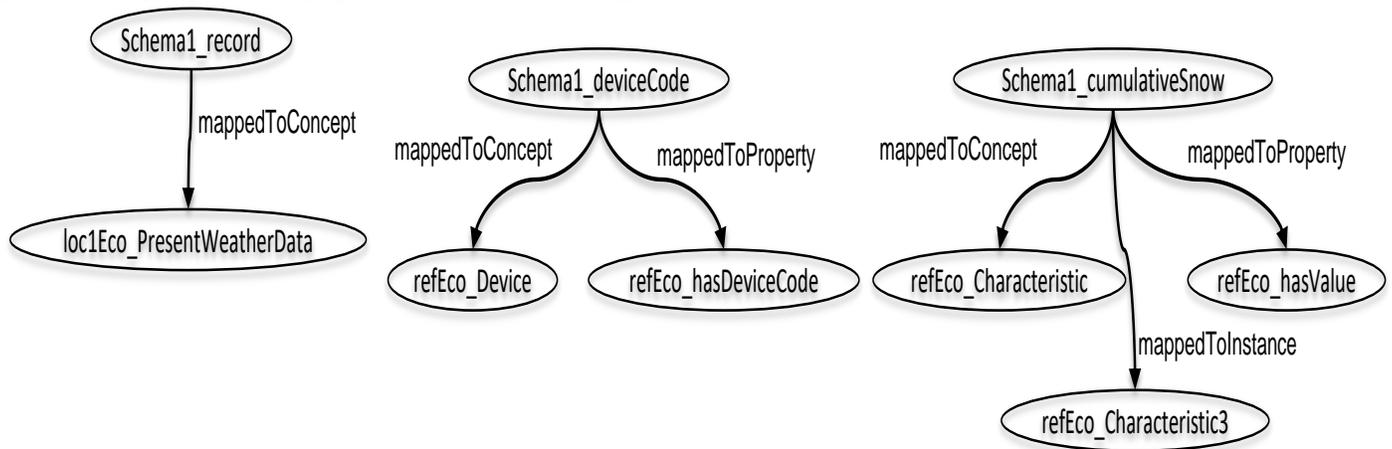


Fig. 7. Excerpt of the graph of mapping1 in use case.

### C. Schema to Ontology Mapping

The mapping relates elements from schema1 (respectively schema2) to concepts and roles of local ontology1 (resp. ontology2). A schema mapping is generally represented as a triple consisting of a schema, a local ontology, and a mapping specifying relationships between them. The system uses this representation and restricts the set of mapping relationships to:

- a mapping to an ontology concept (denoted 'mappedToConcept')
- a mapping to an ontology role (denoted 'mappedToRole')
- a mapping to a concept instance which is stored in the repository (denoted 'mappedToIndividual').

The syntax of the mapping solution is restricted such that not all combinations of the mapping relationships are accepted. The restrictions and their associated semantics are characterized in Table 2.

The simplest abstraction of an XML document is a labeled ordered tree, possibly with data values associated to the leaves. But for the mapping approach, the system takes advantage of the object model view which can also be applied to an XML document. Starting from this view, the system assumes that any XML element is at least mapped to a DL concept or DL individual. This first assumption enables the system to disallow the mapping #1 and #3 which do not inform about an associated DL concept nor DL individual. The purpose of mapping #4 is to inform the system about the absence of mapping for a given XML element. In fact, this most effectively and rapidly performed by users by omitting such a mapping for this element.

In cases where an XML element is not empty, i.e. it contains a data value, it is necessary to map it to a DL property. This is the case of mapping #5, #7 and #8 in Table 2 Mapping #8 is a specialization of mapping #5 where extra information about an associated concept individual is provided. Mapping #7 can be viewed as being equivalent to #8 where the type instance is not specified. In cases of an empty XML element, no DL property needs to be attached to the mapping. Hence, it corresponds to mappings #2 or #6. The latter being a specialization of the former where extra information about a DL concept instance is provided.

Finally mapping #4 is considered as a shortcut of mapping #6 where the DL concept is omitted. This kind of mapping is supported if the processing of the DL realization reasoning procedure, i.e. providing the most specific concept an individual is an instance of, returns a single concept. Thus there cannot be any ambiguities about the type of this individual. Fig. 7 displays an extract of a graph representing the mapping between the schema and the local ontology associated to application1 in the meteorological use case. In this figure, relations start from the XML element of a given schema (pattern is "schemaName\_elementName") and points to an ontology element, i.e. concept, property or individual, of a given ontology (pattern is "ontologyName\_elementName").

This figure emphasizes mappings related to mapping #2, #5 and #8 of Table 2. For instance, the mapping of the 'record' XML element, which is empty and root of the document, is mapped, via 'mappedToConcept' to the 'PresentWeatherData' DL concept. The 'deviceCode' XML element is non-empty and mapped to the 'Device' DL concept and its 'hasDeviceCode' property. Finally the 'cumulativeSnow' element, again a non-empty element, is mapped to a concept ('Characteristic'), a property ('hasValue') and a individual ('characteristic/3'). In order to explain the integration methodology and present the matching issues, it is necessary to present the mapping associated to application2 as well (Fig. 8). Notably, only two elements are mapped to a DL concept: 'measurement', the root element of schema2 and 'data', an empty and nesting element. All other elements are mapped to DL concepts and properties. Finally, the mapping language is the same for the output and input applications of a DaLo-WF.

### D. Methodology and Heuristics

Generating the input document of a DaLo-WF's Application2 is a multi-step process. These steps correspond to (i) matching the local ontologies, (ii) matching the (XML) schemata and (iii) generating the target instance document.

#### Matching local ontologies

This matching step searches for correspondences between the DL concepts of both local ontologies. This operation is supported by the existence of a common vocabulary, the reference ontology. In order to discover as many matches as possible, two techniques are considered to find correspondences: DL-based and navigation-based mappings.

TABLE II. MAPPING POSSIBILITIES IN SI

#	mappedTo Concept	mappedTo Role	mappedTo Individual	Semantics
1				Not accepted
2	X			Empty XML element is mapped to an ontology concept
3		X		Not accepted
4			X	Equivalence to a concept instance
5	X	X		Non empty XML element is mapped to a concept and a role
6	X		X	Empty XML element is mapped to an ontology concept and an individual
7		X	X	Not empty XML element mapped to a role and a concept instance
8	X	X	X	Non empty XML element is mapped to a concept and role as well as a concept instance

The DL-based approach is performed using a DL reasoner and particularly its concept subsumption inference procedure.

In the navigation-based approach, an ontology is taken in terms of a directed acyclic graph where nodes correspond to DL concepts and the edges correspond to DL properties. Basically, it searches for navigation paths between two concepts. This is performed by exploiting the (SPARQL) query facilities of the (triple store) repository. The navigation-based approach also exploits a DL reasoner with its concept subsumption, instance checking and realization inference procedures. This approach is non-deterministic and may return several different paths. So it is important for the algorithm to qualify paths and to select the most appropriate one. This qualification is based on several factors: the length (L) of each path (i.e. the number of properties along a path) and the characteristics of the properties used along a path, i.e. functionality, inverse functionality.

As the implementation formalizes ontologies using decidable species of OWL, i.e. OWL Lite and OWL DL, it is possible to distinguish properties based on their functional characteristics. As a functional property, denoted as 'prop', is defined as:

$$\forall x, y_1, y_2 \mid \exists \text{prop}(x, y_1) \cap \exists \text{prop}(x, y_2) \Rightarrow y_1 = y_2$$

The decidability issue of DL reasoning tasks is a main concern in the solution. For this reason, inverse functional properties are not considered, which are supported in OWL, but are only associated to decidable inferences for object properties. Thus inverse functional properties on data type properties yield an OWL Full ontology which is not decidable.

The system distinguishes between several navigation approaches:

- L=1 and the property is functional: 'functionalNavigation'.
- L=1 and the property is not functional: 'nonFunctionalNavigation'.
- L>1: 'pathNavigation'.

The match operator applied in the DaLo context is able to find several correspondences, usually belonging to the two presented categories, between a given pair of DL concepts. In order to deal with this issue, the system propose a heuristic to select a preferred correspondence. This heuristic is based on a total order of the DL-based and navigation-based categories.

**Definition:** For a given pair of DL concepts C1 and C2, respectively from local ontologies 1 and 2, if a set of correspondences are found between these two concepts: the system knows that there must be at most one DL-based correspondence between C1 and C2 but several navigation-based mappings can coexist with it. For this reason, the system ranks the navigation-based correspondences according to a preference total order: functional Navigation > non-functional Navigation > path Navigation.

Concerning navigation-based relationships, setting a property to be functional is an important commitment for the knowledge engineer. The system thus considers that a functionalNavigation is preferred to a nonFunctionalNavigation. Finally the system considers that navigation with a single edge is more trustable than a path made of several edges.

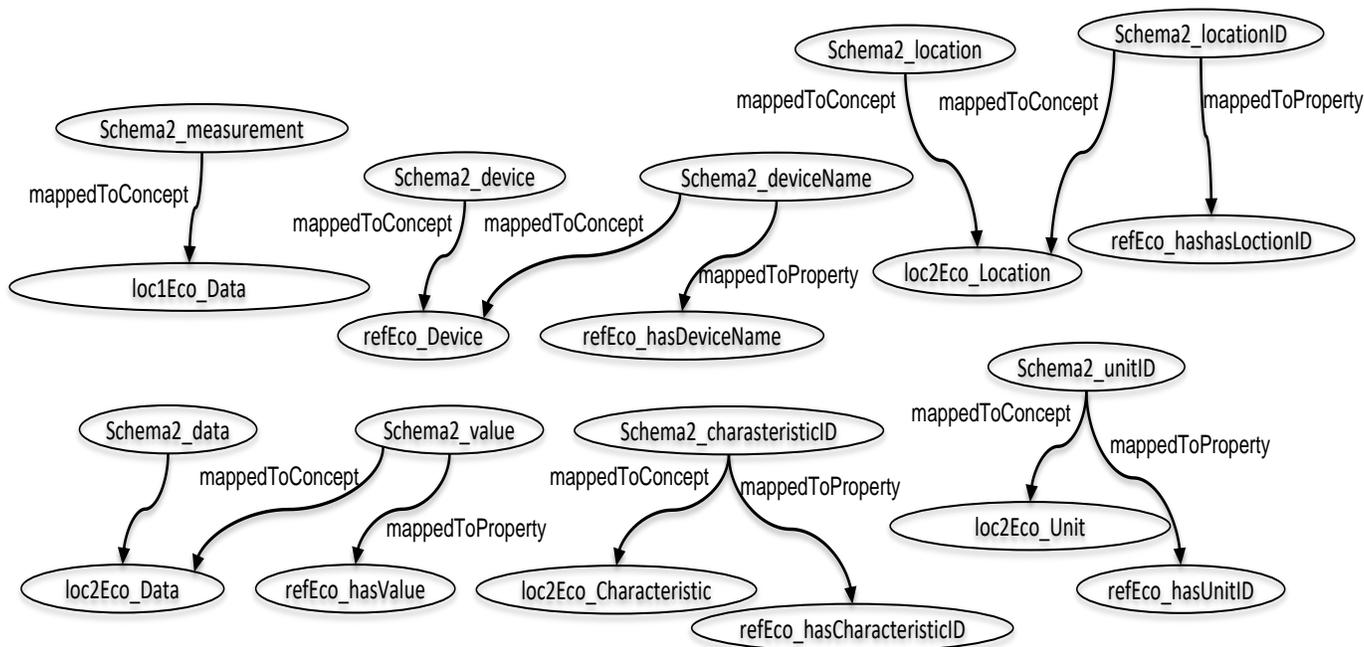


Fig. 8. Excerpt of the graph of mapping2 in the use case.

Thus the system obtains a partial order on the total set of discovered correspondences. On the use cases which are implemented with DaltOn so far, a heuristic has been added, stating that functionalNavigation is preferred to concept generalization which is preferred to nonFunctionalNavigation, thus obtaining a total order on correspondence preferences: concept equivalence > concept specialization > functionalNavigation > concept generalization > nonFunctionalNavigation > pathNavigation.

Other heuristics could also be applied, e.g. generalization > functionalNavigation, and SI supports the definition of specific preference orders.

### Matching schemata

The purpose of this step is to discover mappings between Schema1 and Schema2 from the mappings discovered in the previous step, i.e. between local ontology1 and Local ontology2 [Fig. 5]. This step can be easily performed using the schema to ontology mapping, i.e. from schema 1 to local ontology 1, respectively schema 2 and local ontology 2.

Finally, due to the dual matching solution (logic-based and navigational-based), the accuracy of data stored in the repository and the possibilities to adjust heuristics. These false matches generally occur when local ontologies are modified due to replacement or configuration modifications at the sensors. In these cases, the adjustments need to be performed on the local ontologies and, possibly in non-monotonicity situations of the local ontologies, to the reference ontology [Fig. 6].

### Target instance generation

Starting from these mappings, it is possible to consider the generation of data values for (non-empty) target elements. For navigation-based correspondences, the processing is relatively obvious as it is sufficient to follow the selected paths between two concepts. This navigation is performed starting from a specific node of the ontology graph. For instance, in the case of the location sink element, the mapped source element is 'deviceCode' and the 'hasDeviceCode' DL property (Fig. 7). For a given source instance which has 'PW1' as a value for 'deviceCode', SI will use methods of reference reconciliation [22] to identify the associated graph node. Starting from this node, it is possible to follow the path to the searched value.

For DL-based correspondences, it is required to inspect the DL properties associated to each mapping in order to detect possible transformations. For instance in Table 3, the sink element 'deviceName' is related, via Concept equivalence, to the 'deviceCode'. But the 'deviceCode' element is mapped to the property 'hasDeviceCode' (Fig. 7) while 'deviceName' is mapped to 'hasDeviceName'. Thus a transformation needs to be performed.

A final step consists in enabling the integration of data from application1's instance onto application2's instance

document. Different forms of mappings are available, e.g. relational queries, relational view definitions, XQuery queries or XSLT transformations, to perform this task. The system opted for XSLT transformation since it does not need the expressiveness and complexity of relation queries and views.

By selecting XSLT, the system also benefits from procedural attachment possibilities when performing transformations. That is SI includes a set of procedures, developed in the Java language, to enable the retrieval of values stored in the repository at runtime. Most of these procedures generate, from predefined templates, SPARQL queries and execute them on the repository's ABox.

## IV. RELATED WORK

Kepler [8] is an open-source scientific workflow system which is evolved from Ptolemy system [23]. Kepler's data integration approach is based on a semantic mediation system and utilizes the automated integration services from a middleware called SEEK [4]. SEEK exploits ontological information to support structural data transformation for scientific workflow composition. The prerequisite of the system is to define the structural and semantic type of input and output ports of actors and services they represent. A user then defines registration mappings to associate contextual paths on the ontologies to data objects generated in response to the queries on the service input/output, named ports. Then these input and output registration mapping rules are composed to construct correspondence mappings between structural types of the source and those of the target; DaltOn's SI implementation also generates correspondence mappings and stores them for future use as well. The approaches of Kepler and DaltOn are quite similar as they both aim at transforming data semantically based on a semantic mediation system. Although the objectives are almost identical, the design of the solution is different as Kepler does not use the semantic of the ontologies to generate mappings. Another notable difference between Kepler and DaltOn is that SEEK does not consider format conversion (syntax incompatibility) and data transportation implicitly – which is beneficial for the normal scientific user.

Triana [10][24] is a workflow-based, graphical problem solving environment. Like Kepler, Triana also provides a rich library of pre-configured and built-in tools. As far as semantic integration is concerned, unlike Kepler, Triana does not support semantic data integration. In Triana, domain users need some pre-developed tools which can perform schema transformation, generate mappings and correspondences and finally integrate data. Also the users need to know how such tools are developed, how these tools are used and in which sequence they must be applied. With the approach, a standard schema (DaLo-WF) is provided must fits most use cases but which can be adjusted in case it does not fit the requirements of an application.

TABLE III. MAPPING POSSIBILITIES IN SI

Elements of XSD1	Elements of XSD2	Preferred correspondence
record	measurement	Concept specialization
deviceCode	device	Concept equivalence
	deviceName	Concept equivalence
	location	nonFunctionalNavigation
	locationID	nonFunctionalNavigation
cumulativeSnow	data	Concept generalization
	value	Concept generalization
	characteristicID	Concept generalization
	unitID	nonFunctionalNavigation

Taverna [9], a scientific workflow management system, is part of the myGrid project. In order to convert data formats, Taverna provides “Shims”, which are used as web services. DaltOn differs from Taverna in the way that it handles format conversions (syntactical conversions) dynamically. In Taverna, the domain user required some sort of specialized services that convert schemas and performs mappings as well. DaltOn instead provides a transparent way to deal with semantic integration issues.

### V. CONCLUSION

This contribution discussed in detail a method for developing scientific applications. One of the main messages is that separation of concerns can help to ease handling complex application scenarios as they often occur in scientific domains. This is achieved by applying POPM which already introduces a separation of concerns and by further separating data integration tasks from domain related tasks. Thus the readability of a process is increased and domain users can focus on their expertise – the scientific analysis.

The other main contribution is an ontology based data integration framework called DaltOn; instead of fixing transformation semantics in code, it is specified as a mapping between ontologies. Since data transformation is specified on a conceptual level, changing and adjusting these transformations whenever schemas or ontologies evolve is rather easy.

### REFERENCES

[1] ABCD Schema – Task Group on Access to Biological Collection, Website: <http://www.bgbm.org/TDWG/CODATA/>, [May 2017].

[2] Cheatham, Michelle, and Catia Pesquita. "Semantic Data Integration." In Handbook of Big Data Technologies, pp. 263-305. Springer International Publishing, 2017.

[3] Ning, Z. H. A. O. "Semantic Conflict Resolution Scheme Based on Ontology." DEStech Transactions on Engineering and Technology Research sste (2016).

[4] SEEK Observation Ontology, draft design document, Website: <http://lists.admb-project.org/seek.ecoinformatics.org/>, [May 2017].

[5] Groß, Anika, Cédric Pruski, and Erhard Rahm. "Evolution of biomedical ontologies and mappings: Overview of recent approaches." Computational and structural biotechnology journal 14 : 333-340, 2016

[6] Liu, Jia, et al. "Grid workflow validation using ontology-based tacit knowledge: A case study for quantitative remote sensing applications." Computers & Geosciences, 98 : 46-54, 2017

[7] Hitzler, P. "Ontology Design Patterns for Data Integration: The GeoLink Experience." Ontology Engineering with Ontology Design Patterns: Foundations and Applications, 25 : 267, 2016

[8] Crawl, Daniel, Alok Singh, and Ilkay Altintas. "Kepler WebView: A Lightweight, Portable Framework for Constructing Real-time Web Interfaces of Scientific Workflows." Procedia Computer Science, 80 : 673-679, 2016

[9] Katherine Wolstencroft, Robert Haines, Donal Fellows, Alan Williams, David Withers, Stuart Owen, Stian Soiland-Reyes, Ian Dunlop, Aleksandra Nenadic, Paul Fisher, Jiten Bhagat, Khalid Belhajjame, Finn Bacall, Alex Hardisty, Abraham Nieva de la Hidalga, Maria P. Balcazar Vargas, Shoab Sufi, and Carole Goble: "The Taverna workflow suite: designing and executing workflows of Web Services on the desktop, web or in the cloud", Nucleic Acids Research, 41(W1): W557-W561, 2013.

[10] Vahi, Karan, et al. "A case study into using common real-time workflow monitoring infrastructure for scientific workflows." Journal of grid computing 11.3 : 381-406, 2013

[11] Jablonski, S.; Bussler, C.: Workflow Management – Modeling Concepts, Architecture and Implementation. London: Int. Thomson Computer Press, 1996

[12] ProDatO Integration Technology GmbH: Handbuch iPM Integrated Process Manager. Softwaredocumentation (in German), Erlangen, Germany, [www.prodato.de](http://www.prodato.de) [May 2017]

[13] Website BayCEER, <http://www.bayceer.uni-bayreuth.de/>, [May 2017].

[14] Jablonski, S.; Faerber, M.; Götz, M.; Volz, B.; Dornstauder, S.; Müller, S.: Configurable Execution Environments for Medical Processes, 4th Int'l Conference on Business Process Management (BPM), Vienna Austria, 2006.

[15] Jablonski, S.; Rehman, M.A.; Volz, B.; Curé, O.: Architecture of the DaltOn Data Integration System for Scientific Applications. 3rd International Workshop on Workflow Systems in e-Science (WSES 08), Lyon, France, 2008

[16] Jablonski, S.; Rehman, M.A.; Volz, B.; Curé, O.: DaltOn: An Infrastructure for Scientific Data Management. Proc. of the Int'l Workshop on Applications of Workflows in Computational Science (AWCS), pp.520-529, Krakow, Poland, 2008.

- [17] Jablonski S.; Curé O.; Jochaud F.; Rehman M. A.; Volz B. : Semantic Data Integration in the DaltOn System. Workshop Information Integration Methods, Architectures and Systems (IIMAS) at 24th Int'l Conference on Data Engineering, Cancún, México, 2008.
- [18] Staab, S.; Studer, R.: Handbook of Ontologies. Springer, Germany. 2004.
- [19] Baader, F.; Calvanese, D.; McGuinness, D.L.; Nardi, D.; Patel-Schneider, P.F.: The Description Logic Handbook: Theory, Implementation, and Applications, New York, USA: Cambridge University Press, 2003.
- [20] Cure, O.; Jochaud, F.: Preference-Based Integration of Relational Databases into Description Logic. in Proc.DEXA'07, 2007, pp. 854-863.
- [21] Cure, O.; Bensaid, J.D.: Integration of relational databases into OWL knowledge bases: demonstration of the DBOM system in Proc. ICDE Workshops pp 230-233, 2008
- [22] Saïs, Fatiha, and Rallou Thomopoulos. "Ontology-aware prediction from rules: A reconciliation-based approach." Knowledge-Based Systems 67 : 117-130, 2014.
- [23] Ptolemaeus, Claudius: System design, modeling, and simulation: using Ptolemy II, Vol. 1. Berkeley: Ptolemy.org, 2014.
- [24] Website myGrid, <http://www.mygrid.org.uk>, [May 2017]
- [25] Website Protégé, <http://protege.stanford.edu/>, [May 2017]

# Comparative Analysis of Online Rating Systems

Mohammad Azzeh

Faculty of Information Technology  
Applied Science Private University  
Amman, Jordan POBOX 166

**Abstract**—Online rating systems serve as decision support tool for choosing the right transactions on the internet. Consumers usually rely on others' experiences when do transaction on the internet, therefore their feedbacks are helpful in succeeding such transactions. One important form of such feedbacks is the product ratings. Most online rating systems have been proposed either by researchers or industry. But there is much debate about their accuracies and stability. This paper looks at the accuracy and stability of set of common online rating systems over dense and sparse datasets. To accomplish that we used three evaluation measures namely, Mean Absolute Errors (MAE), Mean Balanced Relative Error (MBRE) and Mean Inverse Balanced Relative Error (MIBRE), in addition to Borda count to assess the stability of ranking among various rating systems. The results showed that both median and Dirichlet are the most accurate models for both sparse and dense datasets, whereas the BetaDR model is the most stable model across different evaluation measures. Therefore we recommend using Dirichlet or BetaDR for the products with few number of ratings and using the median model with products of large number of ratings.

**Keywords**—Online rating systems; reputation models; comparative analysis; decision making; e-commerce

## I. INTRODUCTION

Online rating systems play a vital role in most ecommerce applications. They help users to facilitate their decisions while they perform internet transactions [1], [4]. The online rating system is responsible for collecting, processing and aggregating ratings given for a specific product. The main challenge that faces the online rating systems is how to aggregate the collected ratings for a specific product in way that can reflect its real quality [13]. In practice, most of the well-known ecommerce portals such as eBay, Amazon, etc. use their own methods to compute the quality of product. But some other portals use the simplest aggregation method which is the Naïve average methods (i.e. mean, median and mode). In contrast, many authors proposed different method to compute product score based on statistical and machine learning methods. The accuracy of such methods depends mainly on the user satisfaction about the results achieved [14]. This satisfaction is difficult to be measured because most ecommerce application don't provide a tool to evaluate the user satisfaction, and whether the given aggregate rating help them in performing the successful transaction. The rating aggregation methods in literature can be divided into four groups, Naïve models, weighted average models, Fuzzy models and probabilistic models. The weighted average models are the widely used among researchers, where the weights are derived from historical user data or time factor. These weight

values work as discount factors to reflect different aspects of users' behavior such as their reliability, trustworthiness and credibility in providing rating. One of the common problem that faces rating systems is unfair ratings that biases aggregate scores for some products.

This paper attempts to look at the accuracy and stability of the most common online rating systems over dense and sparse datasets. Practically, not all methods perform well over dense or sparse datasets. This fact has been confirmed by almost all previous studies because each model attempts to treat a specific limitation in previous rating systems. To best of our knowledge, there is no systematic procedure has been conducted to compare and evaluate different online rating systems in terms of accuracy and stability. The proposed research questions are:

**RQ1:** Is there any one method that can perform stably well under all conditions?

**RQ2:** Which group of methods is more appropriate for dense datasets?

**RQ3:** Which group of methods is more appropriate for sparse datasets?

This paper is structured as follows: section 2 presents the literature and overview of existing online rating systems. Section 3 introduces the experimental methodology and comparison procedure. Section 4 presents the obtained results, finally we end up with the conclusions in section 5.

## II. OVERVIEW

Online rating system receives ratings from users as input to compute the aggregate score of product. Given a set of users  $U = \{u_1, u_2, u_3, \dots, u_n\}$  where each user rated at least one product, also given a set of products  $P = \{p_1, p_2, p_3, \dots, p_m\}$  where each product received at least one rating, the intersection between user  $u_i$  and the product  $p_j$  is the rating  $r_{ij}$  such that  $1 < r_{ij} \leq k$ .  $k$  is the maximum rating level for rating system.  $\bar{p}_j$  is the ratings average of product  $p_j$ , and  $\bar{P}$  is the average of all ratings in the dataset. Indeed, Naïve methods such as arithmetic mean (see Equation 1) and median are the most common used methods. Garcin et al. [15] compared between Naïve methods and other rating systems. They revealed that the median is the most accurate method. In contrast, other studies [8], [9] showed that the naïve methods are ineffective because they are easily influenced by unfair and malicious ratings and cannot discover trend emerging from recent ratings.

IMDb is another famous online rating system that uses true Bayesian estimation to calculate the aggregate product score as

shown in (2). The exact implementation of this model is still unpublished in order to keep the policy effective.

$$\bar{p}_j = \frac{1}{n} \sum_{i=1}^n r_{ij} \quad (1)$$

$$IMDb\_score_j = \frac{n}{(n + MinR)} \times \bar{p}_j + \frac{MinR}{(n + MinR)} \times \bar{P} \quad (2)$$

Where  $n$  is the number of ratings received for product  $p_j$ .  $MinR$  is the minimum number of rating count required to appear on the top 250. IMDb usually uses  $MinR=2500$ .

In literature, the weighted average models are the widely used models, where the weights are computed based on either time or user data. Josang and Haller [5] introduced the age of rating as discount factor in computing and aggregating rating, where old ratings receive less weight than recent ratings because they are not informative. The main problem with this model is which time unite (i.e. day, week, month, year) should be considered with this function. Another time discount function used is the number of past transactions instead of using the ratings age [10]. Leberknight et al. [8] stated that the naïve methods are good when there is clear trend of ratings over time, but when the ratings do not have that trend one should involve the volatility of ratings as discount factor to compute the product score. They proposed discount function based on rating volatility, but they ignored the importance of other factors such as trustworthiness and credibility of users. On the other hand, many online rating systems use users' data to measure their reliability, credibility and trustworthiness and reflect that as weight during aggregation process [12]. In this direction, Riggs et al. [11] defined the reliability of a user by his ability to provide rating that is very close to the current ratings average. They defined a measure to calculate that closeness and use it with their weight average model. Lauw et al. [7] studied the leniency of user while they rate products. They proposed a function that can calculate the leniency and strictness of user and reflect that as weight. They classified users into two classes (lenient and strict) based on leniency variable  $l_i$  as shown in Equation 3, such that if  $l_i < 0$  then reviewer is strict, otherwise reviewer is lenient. This model is called LQ.

$$l_i = \frac{1}{|u_i|} \sum_{j=1}^{|u_i|} \left( \frac{r_{ij} - q_j}{r_{ij}} \right) \quad (3)$$

$$q_j = \frac{1}{|e_j|} \sum_{i=1}^{|e_j|} r_{ij} \times (1 - \alpha \cdot l_i) \quad (4)$$

Where  $q_j$  is the initial quality of the item  $j$  which is usually the average of ratings.  $l_i$  is the leniency of the reviewer.  $\alpha \in [0, 1]$  is a compensation factor determined by expert. Abdel-Hafez et al. [1], [16] used Beta distribution function to compute ratings weights. Their model is called BetaDR. The product ratings should be first sorted from smallest to largest and scaled as shown in (5). The beta distribution function has

advantage such that it can change its shape based on the rating distribution. Therefore they controlled the shape of the function by two variables  $\alpha$  and  $\beta$  as shown in (6). Finally, the product score is measured as shown in (7).

$$x_i = \frac{0.98 \times i}{n - 1} + 0.01 \quad (5)$$

$$Beta(x_i) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1} (1 - x_i)^{\beta-1} \quad (6)$$

$$BetaDR = \sum_{l=1}^k (L \times w_l) \quad (7)$$

Where  $\Gamma$  is the gamma function, and  $\alpha$  and  $\beta$  are Beta distribution parameters that are determined based on mean and distribution of ratings.  $L$  is rating level (i.e. 1, 2, ...  $k$ ).  $w_l$  is the summation of normalized Beta weight for the target level. Jøsang et al. [6] introduced a reputation model based on Dirichlet probability distribution as shown in Equations 8 and 9. This model is a generalized form to their previous model and takes the rating counts in calculation. The model works well with good accuracy over sparse datasets because it involves factors that can treat uncertainty in the data.

$$\bar{S}_y: \left( S_y(i) = \frac{R_y(i) + Ca(i)}{C + \sum_{j=1}^k R_y(j)}; |i = 1 \dots k \right) \quad (8)$$

$$score = \sum_{i=1}^k v(i)S(i); \text{ where: } v(i) = \frac{i - 1}{k - 1} \quad (9)$$

where  $\bar{S}_y$  represents the score vector of each rating level,  $S_y(i)$  represents the probability that one agent gives rating  $i$  to agent  $y$ .  $C$  is a constant value, and  $(i)$  is the base rate, which equals to  $1/k$ .  $R_y(i)$  is the number of ratings of the level  $i$ .

Bharadwaj et al. [2] used the ordered weighted averaging method with fuzzy computation as part of their trust model to aggregate rating as shown in Equations 10 to 12. According to them, the reputation of a reviewer is defined as the accuracy of his prediction to other reviewer's ratings towards different items. Recently, Liu et al. [9] proposed several factors to identify unfair ratings. These factors are combined together using Fuzzy Logic System based on human predefined rules. The output of Fuzzy logic system is the discount weight of rating.

$$score_j = \sum_{i=1}^n W_i \times r_{ij} \quad (10)$$

$$W_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \quad (11)$$

$$Q(r) = \begin{cases} 0 & 0 \leq r \leq 0.3 \\ 2 \times r - 0.6 & 0.3 < r \leq 0.8 \\ 1 & 0.8 < r \leq 1 \end{cases} \quad (12)$$

### III. EXPERIMENTAL SETUP

#### A. Datasets

Most authors used public datasets to validate their models which allow them to generalize the extracted knowledge. In this paper we continue that approach to facilitate the replication studies in future. Two stable versions of MovLens datasets have been used namely, 100K and 1M [3]. Both datasets have large number of ratings which are considered dense datasets as shown in Table 1. To compare online rating systems over sparse datasets, we extracted new three datasets from the original 1M dataset, where each new dataset contains randomly selected 4, 6 and 8 user ratings respectively. These datasets are called 1M4, 1M6 and 1M8. The characteristics of all datasets are shown in Table 1.

TABLE I. DATASETS CHARACTERISTICS

Dataset	#User	#Movies	#ratings
100K	943	1682	100,000
1M	6040	3706	1,000,209
1M4	6040	920	24,160
1M6	6040	1286	36,240
1M8	6040	1625	48,320

#### B. Evaluation measures

Evaluation measures are used to assess the accuracy and stability of online rating systems. To measure the accuracy of a model we used three measures, Mean Absolute Errors (MAE), Mean Balanced Relative Error (MBRE) and Mean Inverse Balanced Relative Error (MIBRE). These measures have been selected as they are not biased. The MAE assesses, for each product, the closeness of the generated score to the actual ratings for a product as shown in Equation 13. Both MBRE and MIBRE compute the relative accuracy of the generated scores as shown in Equations 14 and 15.

$$MAE = \frac{1}{m} \sum_{j=1}^m \frac{\sum_{i=1}^n |r_{ij} - score_j|}{n} \quad (13)$$

$$MBRE = \frac{1}{m} \sum_{j=1}^m \left( \frac{1}{n} \sum_{i=1}^n \frac{|r_{ij} - score_j|}{\min(r_{ij}, score_j)} \right) \quad (14)$$

$$MIBRE = \frac{1}{m} \sum_{j=1}^m \left( \frac{1}{n} \sum_{i=1}^n \frac{|r_{ij} - score_j|}{\max(r_{ij}, score_j)} \right) \quad (15)$$

Where  $score_j$  is the aggregated score for product  $p_j$ .  $m$  is the number of products in the testing data.

#### C. Experimental procedure

As mentioned in the literature, there are many models have been proposed to aggregate online ratings. In this study we used eight state-of-art models are: Mean, Median, BetaDR [1], Bayesian [6], Dirichlet [5], IMDb, Fuzzy rating [2] and LQ [7]. For comparison purpose we used 10-Fold cross validation. This procedure divides the dataset into 10 groups of training and testing data. Each group has 90% of the data as training data

and 10% as testing data. The training data is used to build the online rating system, while the testing data is used to evaluate the model. The validation is running 10 times, one time for each group. In each run we record the MAE, MBRE and MIBRE for test ratings. The fundamental idea of using this validation technique is that a reputation score that is produced from training dataset is considered accurate if it is very close to actual ratings in the testing dataset. To measure the stability for each model across different evaluation measures, we rank all models according to their accuracy in terms of MAE, MBRE and MIBRE over all datasets. Then we run Borda count method over all datasets, dense datasets and sparse datasets respectively. Borda count is voting ranked method used to rank various candidates based on the ranks provided by voters. This method is simple and very common in decision making area. First we evaluate the stability of all models over all datasets across all evaluation measures. Then in the second round we evaluate the stability over only dense datasets, then finally over sparse datasets. In all cases the evaluation measure work as voters.

### IV. RESULTS AND DISCUSSION

This section presents the results of comparisons among different online rating systems. Table 2 shows the MAE results over all datasets. From the results we can notice that the differences between all models are nearly negligible, except for LQ model where it is extreme over both dense and sparse datasets. It is interesting to know that Naïve models produce accurate results in comparison to more sophisticated models such as Bayesian and LQ. For the dense datasets (i.e. 100K and 1M) the median model produces the more accurate results, while for sparse datasets the Dirichlet and BetaDR are more accurate. This results confirmed previous findings that confirm that both Dirichlet and BetaDR were originally proposed to handle sparse datasets that contain very few ratings. In spite of that, the median model still produces comparable accuracy to Dirichlet model over all sparse datasets.

TABLE II. MEAN ABSOLUTE ERROR RESULTS

Dataset	Mean	Median	BetaDR	Bayesia n	Dirichle t	IMDb	Fuzzy	LQ
100K	0.905	0.886	0.892	0.902	0.892	0.906	0.919	1.021
1M	0.841	0.810	0.832	0.844	0.841	0.855	0.848	0.962
1M4	0.877	0.876	0.872	0.882	0.883	0.909	0.887	0.982
1M6	0.911	0.907	0.906	0.926	0.886	0.908	0.916	1.023
1M8	0.907	0.897	0.901	0.902	0.883	0.909	0.921	1.007

To perform further investigations, we run the analysis using MBRE and MIBRE evaluation measures. Table 3 shows the results of MBRE over all datasets. Similar to Table 2, the accuracy results are close. Generally, we can observe that the Dirichlet model is the most accurate model over both dense and sparse datasets. Table 4 suggests that the median model is the most accurate model over all datasets. This variation in the results confirm that both median and Dirichlet models are the most accurate models for both sparse and dense datasets. Based on above analysis we can recommend using the median model because it has simple implementation than Dirichlet and can

produce comparable to Dirichlet and better than many sophisticated models.

TABLE III. MBRE RESULTS

Dataset	Mean	Median	BetaDR	Bayesian	Dirichlet	IMDb	Fuzzy	LQ
100K	0.477	0.491	0.476	0.480	0.464	0.478	0.495	0.549
1M	0.418	0.422	0.419	0.429	0.416	0.430	0.431	0.463
1M4	0.409	0.425	0.414	0.418	0.395	0.416	0.426	0.457
1M6	0.428	0.439	0.428	0.445	0.395	0.410	0.434	0.487
1M8	0.421	0.430	0.425	0.425	0.394	0.411	0.438	0.506

TABLE IV. MIBRE RESULTS

Dataset	Mean	Median	BetaDR	Bayesian	Dirichlet	IMDb	Fuzzy	LQ
100K	0.251	0.240	0.246	0.248	0.248	0.249	0.250	0.288
1M	0.233	0.218	0.229	0.231	0.233	0.235	0.231	0.268
1M4	0.221	0.216	0.218	0.218	0.225	0.228	0.220	0.250
1M6	0.228	0.223	0.225	0.229	0.226	0.228	0.226	0.261
1M8	0.228	0.220	0.225	0.225	0.224	0.229	0.228	0.258

To analyze the stability of all models over all datasets and both sparse and dense datasets, we first rank all models over each dataset individually and over each evaluation measure. Then we apply the Borda count method. Table 5 presents the ranking stability of all models over dense and sparse datasets. From the results of ranking we can notice that the BetaDR is the most stable model over all datasets and especially over dense datasets across different evaluation measures, whereas the Dirichlet model is the most accurate model over sparse datasets. Generally, we can notice that the top three models in the table (i.e. BetaDR, Dirichlet and median) are the most stable models. The results obtained surprisingly suggest that the BetaDR is better than both Dirichlet and median over all datasets. In contrast, we can observe that the sophisticated models such as Fuzzy and LQ are not accurate as they occupy the last position over all datasets and across all evaluation measures. Also the commonly used mean model occupies mid positions with unstable ranking across all evaluation measures.

TABLE V. RANKING STABILITY

Rank	All datasets	Dense datasets	Sparse Datasets
1	BetaDR	BetaDR	Dirichlet
2	Dirichlet	median	BetaDR
3	median	Dirichlet	median
4	mean	Bayesian	mean
5	Bayesian	mean	IMDB
6	IMDB	IMDB	Bayesian
7	Fuzzy	Fuzzy	Fuzzy
8	LQ	LQ	LQ

Finally we revisit the proposed research questions in this study:

**RQ1:** Is there any one method that can perform stably well under all conditions?

*Ans.* Actually, there is no accurate answer because the

difference among all models are negligible, but we can say that median and Dirichlet models produce the most accurate results as shown in Tables 2, 3 and 4.

**RQ2:** Which group of methods is more appropriate for dense datasets?

*Ans.* From Table 5 we can see that both BetaDR and median models are the most stable and accurate models over dense datasets.

**RQ3:** Which group of methods is more appropriate for sparse datasets?

*Ans.* Similar to previous answer, we can observe that Dirichlet and BetaDR are the most accurate and stable models over sparse datasets. This is not surprising results because the purpose of construction of both models was to treat the sparse datasets. Also both models are good for new rating system that has few numbers of ratings.

## V. CONCLUSIONS

Online rating system is a helpful tool to facilitate user decision in conducting online transactions. However, the accurate rating system can let user choose the correct product which leads to better user satisfaction. Many models have been proposed in literature, but their accuracy are subject to the degree of helpfulness. In this paper we conducted a comparative analysis for the widely used online rating systems to investigate their accuracies and stability over dense and sparse datasets. Three evaluation measures in addition to Borda count method have been used to assess the stability and accuracy of the employed models. From the obtained results we found that both median and Dirichlet are the most accurate models over dense and sparse datasets respectively. Also we found that the BetaDR are most stable model across all evaluation measures. Finally, the Fuzzy and LQ were the worst models. From these results we can figure out that while the top three ranked models: median, BetaDR and Dirichlet produce relatively accurate and stable results we recommend using median because it has the simplest implementation among three models, and does not consume cost when running. On the other hand, we recommend to use the median model for products with many ratings and using the Dirichlet model when the products have few number of ratings.

## ACKNOWLEDGMENT

The author is grateful to the Applied Science Private University, Amman, Jordan, for the financial support granted to cover the publication fee of this research article.

## REFERENCES

- [1] Abdel-Hafez, A., Xu, Y. Exploiting the beta distribution-based reputation model in recommender system. Paper presented at the 28th Australasian Joint Conference on Artificial Intelligence, 2015, pp. 1-13.
- [2] Bharadwaj, K. K., Al-Shamri, M. Y. H. Fuzzy computational models for trust and reputation systems. *Electronic Commerce Research and Applications*, 2009, 8(1), 37-47.
- [3] F. Maxwell Harper, Joseph A. Konstan. The MovieLens Datasets: History and Context. *ACM Transactions on Interactive Intelligent Systems* 2015, 5, 4, Article 19.
- [4] Jøsang, A., Ismail, R., & Boyd, C. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 2007, 43(2), 618-644.

- [5] Jøsang, A., Haller, J. Dirichlet reputation systems. In The 2nd International Conference on Availability, Reliability and Security (ARES), 2007, 112-119.
- [6] Jøsang, A., Ismail, R. The beta reputation system. In Proceedings of the 15th Bled Electronic Commerce Conference, 2002, 41-55.
- [7] Lauw, H. W., Lim, E.-P., Wang, K. Quality and Leniency in Online Collaborative Rating Systems. ACM Transactions on the Web (TWEB), 2012, 6(1).
- [8] Leberknight, C. S., Sen, S., Chiang, M. On the Volatility of Online Ratings: An Empirical Study E-Life: Web-Enabled Convergence of Commerce, Work, and Social Life 2012, Vol. 108, pp. 77-86.
- [9] Liu, S., Yu, H., Miao, C., Kot, A. C. A fuzzy logic based reputation model against unfair ratings. In Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems, 2013, 821-828.
- [10] Malik, Z., Bouguettaya, A. Rateweb: Reputation assessment for trust establishment among web services. The International Journal on Very Large Data Bases, 2009, 18(4), 885-911.
- [11] Riggs, T., Wilensky, R. An algorithm for automated rating of reviewers. In Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries, Roanoke, Virginia, USA, 2001, 381-387, ACM.
- [12] Cho, J., Kwiseok K., and Yongtae P. "Q-rater: A collaborative reputation system based on source credibility theory." Expert Systems with Applications, 2009, 36(2), 3751-3760.
- [13] Hermoso, R., Roberto, C., Maria ,F. From blurry numbers to clear preferences: A mechanism to extract reputation in social networks. Expert Systems with Applications, 2014, 41(5): 2269-2285.
- [14] Chiu, D. K. W., Ho-Fung, L., and Ka-Man, L. On the making of service recommendations: An action theory based on utility, reputation, and risk attitude. Expert Systems with Applications, 2009, 36(2): 3293-3301.
- [15] Garcin, F., Faltings, B., Jurca, R. Aggregating reputation feedback. In Proceedings of the 1st International Conference on Reputation: Theory and Technology, 2009, 62-67.
- [16] Abdel-Hafez, A., Xu, Y., Jøsang, A. A normal-distribution based rating aggregation method for generating product reputations. In Web Intelligence, 2015, 13(1), pp. 43-51.

# Method for System Requirements Approval

Lindita Nebiu Hyseni

Faculty of Contemporary Sciences and Technologies  
South East European University  
Tetovo, Macedonia

Zamir Dika

Faculty of Contemporary Sciences and Technologies  
South East European University  
Tetovo, Macedonia

**Abstract**—The requirements approval method is necessary to ensure that the system requirements have been identified in right way and the understanding between the contractor and the client exist. During research conducted is identified that most of the scholars have been working for the requirements definition during the meeting with the client, even they started to initiate the validation by checking whether the requirements captures the needs of client but not the approval of the requirements. Therefore, it is proposed the Joint Approval Requirements (JAR) method based on identified gaps through literature review and work experience. In this paper, this theoretical JAR method has been developed further on, through the presentation of its details about approval of the final version of the functional and non-functional requirements document and the integrated conceptual model of the IS. The presented method is ready for the research community in order to implement in different industries to measure the effect of the JAR method in the system requirements.

**Keywords**—Approval method; approve requirements; system requirements; functional and non-functional requirements; joint approval requirements

## I. INTRODUCTION

The information systems (IS) enable individuals, businesses and society to achieve their goals through information and communication technology components [5]. Considering this, it is very important to review and approve the system requirements by the client in meeting with the contractor. Therefore, is asked the research question, what the scholars worked in this direction? Based on this, the focus is on used methods and techniques during the system analysis and design stage because they impact the successful rate of the information system through the clarification of the incompleteness and inconsistencies that may turn out [1], [2]. During the research conducted is identified that most of the scholars have been working on requirements definition during the meetings with the client, even if they have started to initiate the issue about checking whether the requirements capture the client's need (requirements validations) but not approving the final requirements in these meetings [3], [4]. Consequently, it is proposed the Joint Approval Requirements (JAR) method to be applied in the JAR meetings with the client for reviewing and approving the created functional and non-functional requirements document and the integrated conceptual model based on proposed guides of the Integrated Framework for Conceptual Modeling (IFCMod) [5]. As well, the conceptual modeling plays the crucial role in the success of the information system [7]-[9]. The focus of this paper is to present

the details of reason behind the proposed Joint Approval Requirements (JAR) method [1], through a state of the art review for the identification of the research gap, as well as presentation of details of the JAR method. This paper is organized as follows: in Section 2 is state of the art; in the Section 3 are presented identified research gaps; in Section 4 is presented the Joint Approval Requirements (JAR) method; in Section 5 is conclusion and future work; at the end, in Section 6 are presented the references used in this paper.

## II. STATE OF THE ART

During the research conducted is identified the small number of scholars who have been working for the requirements definition during the meeting with the client, even if they have started to initiate the issue about validation requirements by checking whether the requirements captures the client's need but they do not treated the approval of the requirements. The technique which started to initiate the validation requirements as optional during the session with client is Joint Requirements Planning (JRP) presented by Bentley, L. D., & Jeffrey, L. (2007). This technique is for problem identification, analysis and system requirements definition during the JRP session. It includes also planning for the JRP session, the way how to conduct the JRP Session, the list of benefits from using this technique, the product of the JRP Session [3].

As well, the Joint Application Development (JAD) technique presented by Dennis, A., Wixom, B. H., & Roth, R. M. (2012) is used for the requirements definition but in this technique, is not mention the validation or approval of the requirements, even if the client made a decision about which type of requirements seems to be same [4]. This technique allows project team, users, and management to work together to identify requirements for the information system during the JAD Session. It includes the way how selecting participants and location for the JAD session, how to design the JAD session, how to be prepared and conduct the JAD session, and what will be the JAD post-session report. In the following Table 1 is presented comparison of the summary of the JRP and JAD technique.

Considering the identified gaps through literature review, work experience and this comparison presented in Table 1, the proposed Joint Approval Requirements (JAR) method has been developed further on, through the presentation of its details about reviewing and approving the created functional and non-functional requirements document and integrated conceptual model based on the IFCMod guides [5], [6].

TABLE I. COMPARASION OF THE SUMMARY OF THE JRP AND JAD TECHNIQUE

	Joint Requirements Planning (JRP)	Joint Application Development (JAD)
<b>Technique usage</b>	The JRP technique is used for the problem identification, analysis and system requirements definition.	The JAD technique is used for the requirements definition.
<b>Session planning</b>	Planning for the JRP session contains following steps: Selecting the location for the JRP session, Selecting JRP participants (sponsor, facilitator, users and managers, scribes and IT staff), Preparing the agenda to be followed during JRP session.	Selecting participants for the JAD session is done by selecting very best people in business units by providing broad mix of organizational levels. While the participant which will be JAD facilitator, will be a consultant external to the organization, because as the author explains the organization may not have a regular day-to-day need for JAD expertise.
<b>Session conducting</b>	The JRP technique actively involves management and users in the development of the project; it decreases the time spending for finding facts in planning and analysis phase; it confirms requirements and approve prototypes if have, because prototypes were optional in this technique only for the functional requirements that needed to be validated	The JAD group meets for several hours, several days, or several weeks until all of the issues have been discussed and the needed information is collected. This technique is used in analysis and design phase.
<b>Session product</b>	The product of the JRP session is a formal written document created usually by JRP facilitator and scribes. This document is for confirmation of the specifications agreed on during the JRP session by users and managers. The content and organization of the specifications depended on the objectives of JRP session that were, the objective of the session is not determined. This document is published immediately following the meeting in order to maintain the momentum of the JRP session.	The product of the JAD session shall be the post-session report. This report is prepared and circulated among participants of the session, it usually takes two or three weeks after JAD session to be prepared.
<b>Session Layout</b>	Selected location shall be the Room Layout in the “U” shape with additional tables for the IT professionals, other observers and scribe; It presents which position is placed in which part of the room; Also, the room layout includes the dimension of the room and food & refreshments inside the meeting room.).  The room layout presents the way of the organization of the meeting	This technique specifies: The location for the JAD session by presenting the JAD Meeting Room in “U” shape. It doesn’t present which position is placed in which part of the room; How to design the JAD session. Most of the JAD sessions are used to collect information from users, so it is required to develop set of question before starting the meeting.  The room layout does not present the way of the organization of the meeting

### III. JOINT APPROVAL REQUIREMENTS (JAR) METHOD

The proposed method called Joint Approval Requirements (JAR) used in the JAR meetings with clients allows reviewing and approving the document of the functional requirements and the non-functional requirements (FRs & NFRs DOC), and the integrated conceptual model (ICM) which are created by using guides of the Integrated Framework for Conceptual Modeling (IFCMod) [5], [6]. In the following are presented two types of JAR meetings including structure and facilitation.

*Unit JAR meetings* – in the Unit JAR meeting, the participants from the client side shall be: Top Management, Unit Manager, Unit users and IT Team or System Analyst. If it does not have IT Team on the client side, it’s recommended to employ the system analyst for this process. While, the participants from the contractor side shall be: Top Management, Managers and IT Team (System Analyst, System Developer, Database Developer, System Administrator, and System Network Engineer), JAR Scribes which shall be the persons who written the functional and non-functional requirements document (FRs & NFRs DOC) and modeled the graphical representation of the integrated conceptual model (ICM) and JAR Presenter (The system analyst) who present the FRs & NFRs DOC and ICM. The number of Unit JAR meetings depending on the number of units on the client side. If the Unit JAR meeting is holding for the information system which shall be developed in-house than the Top Management

and IT team shall be in the role of the contractor and the client, whereas the Unit Managers and Unit users shall be in the role of the client. Different from the other position of the IT Team, the system analyst of the institution shall be in the role of the JAR Facilitator, JAR Presenter, and JAR Scribe. In Fig. 1 is presented the Unit JAR Meeting Room Layout and the way of organization.

*Final JAR meeting* – in the Final JAR meeting, the participants from the client side shall be: Top Management, all Units` Managers, and IT Team or System Analyst. While, the participants from the contractor side shall be: Top Management, Managers and IT Team, JAR Scribes and JAR Presenter. If the Final JAR meeting is holding for the information system which shall be developed in-house than the Top Management and IT team shall be in the role of the contractor and the client, whereas the Unit Managers shall be in the role of the client. Different from the other position of the IT Team, the system analyst of the institution shall be in the role of the contractor who represents JAR Facilitator, JAR Presenter, and JAR Scribe.

In principle, shall be only one Final JAR meeting, but if the number of Units in the client side is too big, then the Final JAR meeting shall be in groups until all units shall be included. Units shall be included in a group based on work dependency on each other. In Fig. 2 is presented the Final JAR Meeting Room Layout and the way of organization.

The preparation of the Unit and Final JAR meeting shall be done by JAR Facilitator based on the following duties:

*Defining the objectives of the JAR meetings by explaining the purpose which shall be the approval of the FRs & NFRs DOC and ICM after reviewing by the client in JAR meeting with the contractor: If the requirements document and the ICM*

are for the information system which shall be developed in-house than the client shall be the Top Management, Unit Managers, Unit Users and IT Team except for system analyst because s/he shall be in the role of the contractor. Also, the Top Management shall be in the role of the contractor if it is needed. Additional, the FRs & NFRs DOC and ICM shall be provided from a system analyst on the contractor side.

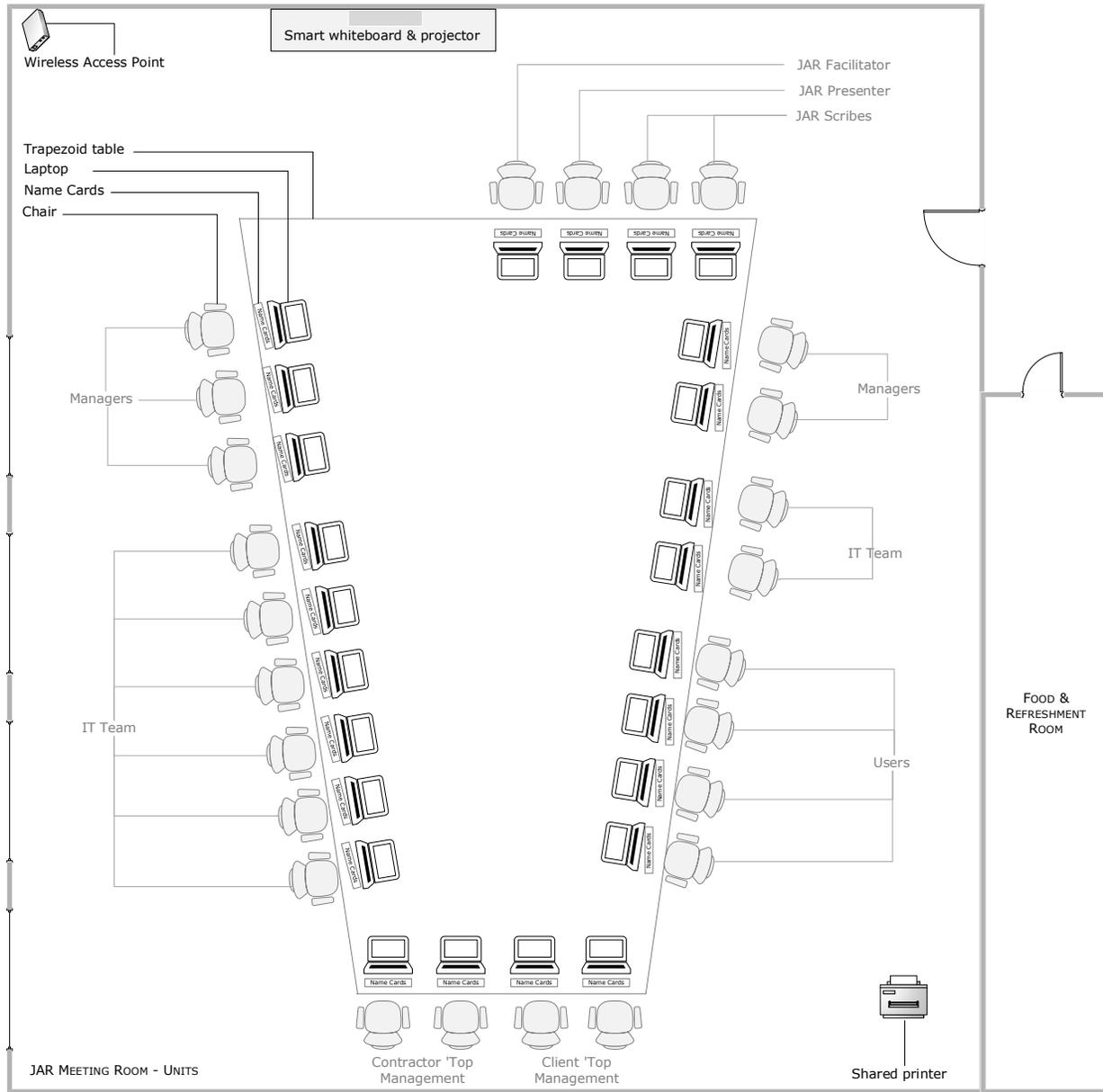


Fig. 1. The Unit JAR Meeting Room Layout.

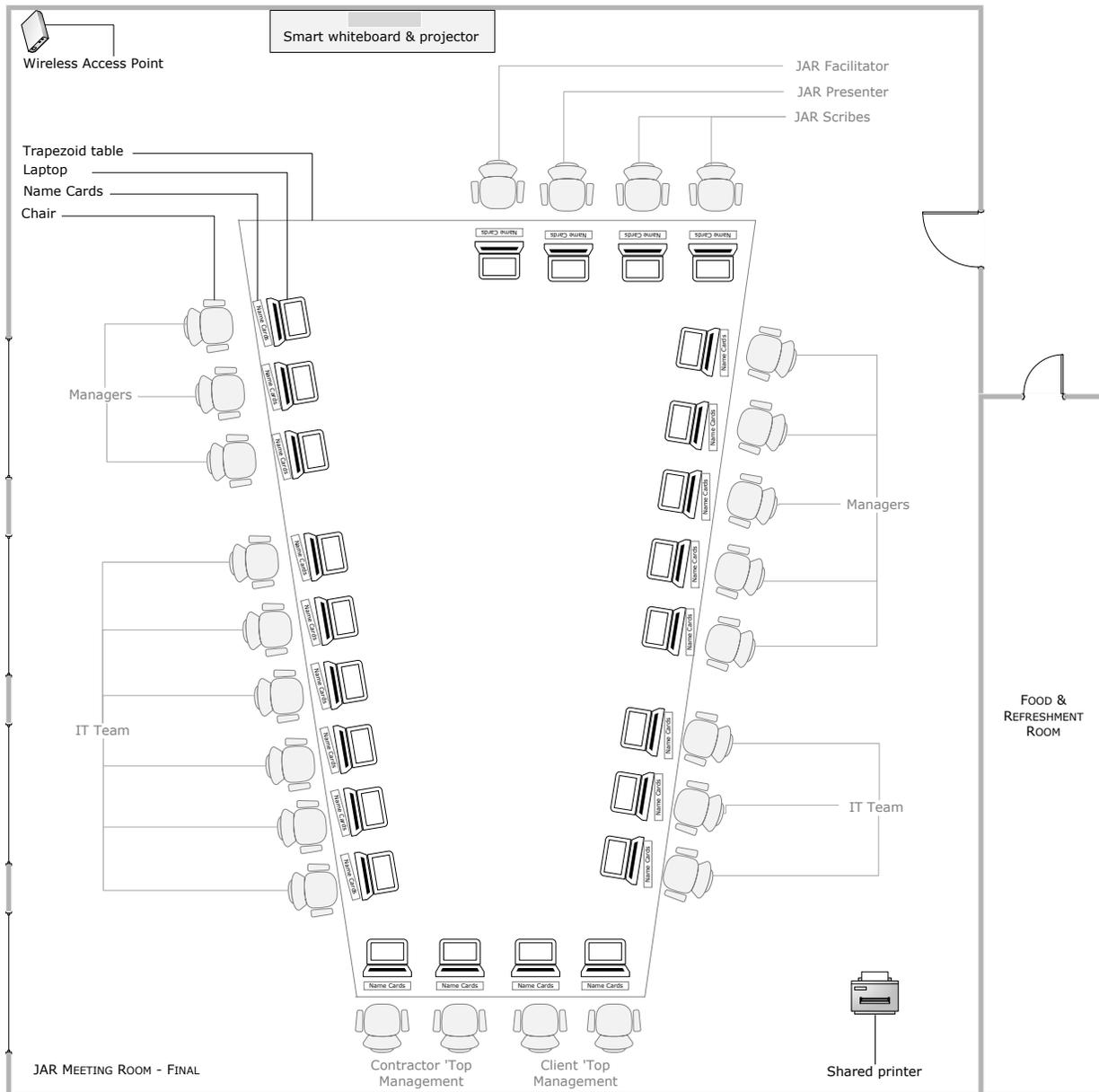


Fig. 2. Final JAR meeting room layout.

Defining the list of the participants with the following information:

- Type of participants – the participants from client side and contractor side.
- Participant unit – the unit which the participant works.
- Participant position – From the client side, shall be selected one person for each position. All participants from the client and contractor side shall be the representative people in their position.
- Participant name and surname.
- Participant email.

- Participant phone number.

Defining the location by finding a place where the JAR meeting shall be conducted. The JAR meeting should be conducted outside of the workplace, while the launch time and the coffee break should be in the same place with JAR meeting but different rooms.

*Defining the duration of the JAR meetings by setting date and time:* A Unit JAR meeting can be of more than one day depending on the complexity of the information system. Also, the Final JAR meeting can be more than one day per all units; it depends on the complexity of the information system. Both types of JAR meetings, Unit and Final, could have maximum 8 hours meeting per day, including one hour lunchtime and four coffee breaks, five minutes per coffee break.

*Informing the participants by email at least ten days before the JAR meeting:* Inform all participants of the JAR meeting and require confirmation from all of them at least five days before the JAR meeting starts. Also, inform that the meeting shall be held only if all invited participants shall confirm JAR meeting. In this information email, should be attached the JAR Meetings Agenda, the materials of the meeting which should be FRs & NFRs DOC and ICM; In the JAR Meetings Agenda, should be time for Welcome and Remembrance (meeting objectives), Participant's Check-in, Activity based on meeting objectives, lunch time, coffee break time and closure.

*The Role of the participants in the JAR meetings:*

The Client participants shall review and approve the functional and non-functional requirements document and the integrated conceptual model during a presentation by JAR Presenter. First, should be reviewed and approve the FRs & NFRs DOC than the ICM.

The Contractor participants' role is shown in the following:

- The JAR Facilitator – shall open the JAR meetings by explaining the objective of the meeting, present and approve the JAR Meetings Agenda by participants and share the materials of the meeting in the paper if it is necessary otherwise all participants have materials in an email when they are invited for JAR Meetings. The JAR Facilitator also takes care for respecting the agenda by all participants in the JAR meetings and close the meetings. If the JAR meeting is holding for the information system which shall be developed in-house than the JAR Facilitator shall be the system analyst of the institution (JAR Presenter).
- The JAR Presenter (The system analyst) – shall present the prepared functional and non-functional requirements document and the integrated conceptual model of the information system in the JAR meetings. If the JAR meeting is holding for the information system which shall be developed in-house than the JAR Presenter shall be the system analyst of the institution who prepared the requirements document and the graphical representation (FRs & NFRs DOC and ICM). The JAR Presenter shall also explain what type of source (gathered requirements) is used to prepare the requirements document and the integrated conceptual model. During the gathered requirements, should have in consideration what type of information the guides of this Integrated Framework for Conceptual Modeling require to create a requirements document (FRs & NFRs DOC) and the graphical representation (ICM) of the information system.
- The other contractor `IT team, IT Managers and Top Management – shall intervene during the presentation if it is necessary. If the JAR meeting is holding for the information system which shall be developed in-house than the Top Management, IT team and IT Manager shall be in the role of the contractor, whereas the Unit Managers shall be in the role of the client.

- The JAR Scribes – shall make changes in the functional and non-functional document and the integrated conceptual model during the JAR meetings based on the client participants' comments, after analyzing them directly in the JAR meeting by the JAR Presenter (the system analyst) and the other IT team members if it is necessary and getting the solution to the client during the JAR meeting, in order all client participants to approve changes directly in the JAR meeting. If the JAR meeting is holding for the information system which shall be developed in-house than the JAR Scriber shall be the system analyst of the institution (JAR Presenter). Also, the JAR Scriber shall be responsible for writing the Minutes of JAR Meetings.

*Conducted JAR meetings:*

Initially shall hold the Units JAR meetings, then the Final JAR meeting. The Units JAR meetings shall be held one by one, based on dependency they have in daily work showing which is the first unit that does the job, which is second and so on. Each Unit JAR meeting shall be held in a separate time interval because Top Management, Managers, IT team and IT Managers of the client and the contractor should participate in each unit JAR meeting.

In the Unit JAR meeting shall be reviewed and approved the functional and non-functional requirements document (FRs & NFRs DOC) and the integrated conceptual model (ICM) by the client participants. During the review process shall be made changes in the FRs & NFRs DOC and ICM, while the approval shall be made after final changes then sending it to the participants at the end of Unit JAR meeting. After finishing each Unit JAR meeting, the Final JAR meeting with the client and the contractor participants shall be held to make the final review and approval from the client participants.

*The Product of the JAR meetings:*

Product of the Unit JAR meetings is the Final version of the functional and non-functional requirements document and the integrated conceptual model of the units which should consider during the Final JAR meeting. During this meeting shall be created also the Minutes of JAR Meeting per each unit with those information's: date, time, main discussions, participants, location, conclusions, and approvals. While, the product of the Final JAR meeting is the Final version of the functional and non-functional requirements document and the integrated conceptual model which should consider during the development of the information system from contractor side. During this meeting shall be created also the Minutes of JAR Meeting per Final JAR meeting with this information's: date, time, main discussions, participants, location, conclusions, and approvals.

*A. Joint Approval Requirements Meeting Room Layout*

The JAR Meeting Room Layout presented in Fig. 1 and 2, is used for Unit and Final JAR Meetings and it contains the following components: A trapezoid table and chairs, Laptops for every participant with adequate software's for the JAR meeting (doc software and Microsoft Visio Professional 2016),

Name cards for every participant, Smart whiteboard and projector in order to create a collaborative environment in JAR meeting. A shared printer in network which is connected to every participant in the JAR meetings, Wireless Access Point.

The arrangement of the trapezoid table based on the JAR Method is presented in the following:

- In the front of the trapezoid table shall be placed, smart whiteboard and projector.
- In the top-left side of the JAR meeting room shall be placed a wireless access point.
- In the bottom-right side of the JAR meeting room shall be placed a shared printer.
- In the beginning part of the trapezoid table shall be placed a laptop for the JAR Facilitator, a laptop for the JAR Presenter (the system analyst) which shall present the Information System based on the functional and non-functional requirements document and the Integrated Conceptual Model, and two laptops for the JAR scribes.
- In the end part of the trapezoid table, in the right side shall be placed two laptops for the client `Top Management, while in the left side shall be placed two laptops for the contractor `Top Management.
- The right side of the trapezoid table is different for Units JAR meetings and Final JAR meeting: In the right side of the trapezoid table in Units JAR meetings shall be placed laptops for the participants from the client side; near of the top management shall be placed Users than the IT team and in the end the Managers. If the Unit JAR meeting is holding for the information system which shall be developed in-house than in the end part of the trapezoid table shall be placed two laptops for the `Top Management who shall be in the role of the contractor and client; near of the Top Management shall be the Users than the Unit Managers whereas, the IT Team and IT Managers shall be in the role of the contractor. Whereas, in the right side of the trapezoid table in Final JAR meetings shall be placed laptops for the participants from the client side; near of the top management shall be placed the IT team and IT Managers shall be in the role of the contractor.
- In the left side of the trapezoid table shall be placed laptops for the participants from the contractor side; near of the top management shall be placed IT team than Managers of the IT team.
- Food and Refreshment room shall be near of the JAR meeting room, approximately 5 minutes by foot.
- Both type of JAR meetings shall be hold in same JAR meeting room as it is presented in Fig. 1 and 2 but only participants in the right side of trapezoid table shall be different for those two types of JAR meetings as it is presented.

#### IV. CONCLUSIONS AND FUTURE WORK

In this paper is proposed a theoretical method Joint Approval Requirements (JAR), which is for review and approve the system requirements in the JAR meetings with client. This JAR method is ready for scholars to validate during the system analysis and design phase through Mixed Method Case Study by using the Sequential Exploratory Design (QUAL→ quan) [10], [11]. This design stand for qualitative data collection and analyzing which are followed by the quantitative data collection and analysis with the intention of increasing the findings generalization (Peng, G. C., Nunes, J. M. B., & Annansingh, F., 2011; Cameron, R., 2009) [11]-[13].

In continues of this study is working on the validation of the JAR method used during the improvement of the e-Schedule system at South East European University, In the end of the JAR Meetings shall be conducted the Semi-Structured Interviews with participants to assess their experience related to this method and the Integrated Framework for Conceptual Modeling (IFCMod).

#### REFERENCES

- [1] Lombriser, Philipp, Fabiano Dalpiaz, Garm Lucassen, and Sjaak Brinkkemper. "Gamified requirements engineering: model and experimentation." In International Working Conference on Requirements Engineering: Foundation for Software Quality, pp. 171-187. Springer International Publishing, 2016.
- [2] Wagner, Dirk Nicolas. "Breakin' the Project Wave: Understanding and avoiding failure in project management.", 2016.
- [3] Bentley, Lonnie D., and Jeffrey L. Whitten. "Systems analysis and design methods". pp. 229-235. Irwin/McGraw Hill, 2007.
- [4] Dennis, Alan, Barbara Haley Wixom, and Roberta M. Roth. "Systems analysis and design", pp. 119-143. John wiley & sons, 2012.
- [5] Hyseni, Lindita Nebiu, and Zamir Dika. "Integrated Approach To Conceptual Modeling." International Journal of Advanced Computer Science and Applications 7, No. 12, pp. 213-219. 2016.
- [6] Hyseni, Lindita Nebiu, and Zamir Dika, "An integrated framework of conceptual modeling for performance improvement of the information systems" INTECH2017: Seventh International Conference on Innovative Computing Technology, 2017, in press.
- [7] Bera, Palash, Anna Krasnoperova, and Yair Wand. "Using ontology languages for conceptual modeling." Cross-Disciplinary Models and Applications of Database Management: Advancing Approaches: Advancing Approaches 1, 2011.
- [8] Braga, Bernardo FB, João Paulo Andrade Almeida, Giancarlo Guizzardi, and Alessander B. Benevides. "Transforming OntoUML into Alloy: towards conceptual model validation using a lightweight formal method." Innovations in Systems and Software Engineering 6, No. 1, pp. 55-63. 2011
- [9] Mehmood, Kashif, Samira Si-Said Cherfi, and Isabelle Comyn-Wattiau. "CM-Quality: A Pattern-Based Method and Tool for Conceptual Modeling Evaluation and Improvement." In ADBIS, pp. 406-420. 2010.
- [10] Morse, Janice M. "Mixed method design: Principles and procedures." Vol. 4. Routledge, 2016.
- [11] Aramo-Immonen, Heli. "Mixed methods research design." In World Summit on Knowledge Society, pp. 32-43. Springer, Berlin, Heidelberg, 2011.
- [12] Cameron, Roslyn. "A sequential mixed model research design: Design, analytical and display issues." International Journal of Multiple Research Approaches 3 No. 2, pp. 140-152. 2009
- [13] Peng, Guo Chao, J. M. B. Nunes, and Fenio Annansingh. "Investigating information systems with mixed-methods research." In Proceedings of the IADIS International Workshop on Information Systems Research Trends, Approaches and Methodologies. Sheffield, 2011.

# A Novel Modeling based Agent Cellular Automata for Advanced Residential Mobility Applications

Elarbi Elalaouy\*, Khadija Rhoulemi & Moulay Driss Rahmani

LRLT associated unit to CNRST (URAC°29), Faculty of Science, Mohammed V University in Rabat  
4 Av.Ibn Battouta B.P. 1014 RP, 10006 Rabat, Morocco

**Abstract**—Nowadays, residential mobility (RM) is usually interconnected with other urban phenomena to give more realistic and effective to the simulation models in order to support urban planners and decision makers. Recent RM research works to describe models from a functional view; however researchers do less focus in providing software modeling of their RM applications. Based on this note, the article presents an agent cellular automata based modeling for advanced RM applications. The proposed modeling contains six models based on UML 2.0 diagrams which models parts of the system from different views. The work could be of interest for specialists (researchers, designers and developers) when modeling advanced RM applications.

**Keywords**—Residential mobility, multi agent systems, cellular automata; urban modeling

## I. INTRODUCTION

Residential mobility (RM) is a very complex phenomenon that had first been studied as an independent system. This tendency is the classical point of view. However, the present tendency is focusing on to interconnect residential mobility with other urban phenomena with which it could give more realistic and effectiveness to the based simulation computer models in order to supporting urban planning and decision making. For example, authors of [1]-[4] had developed simulation's models integrating residential mobility, housing choice, population growth and land use change in order to simulate residential mobility for different duration of years.

Such research works of residential mobility do describe models from a functional view. They describe equations, functions, algorithms that run and simulate models. However, what could be noticed is that scholars do less focus on providing software modeling of their residential mobility' applications. Departing from this notice, this article propose a novel agent cellular automata based modeling for advanced residential mobility applications. This work gets benefit from development and coding experiments and apprehension of existent simulation models [1], [2], [5], [6].

Residential mobility, housing choice, population growth and land use dynamics are urban processes to be modeled; modeling such urban processes is not a simple task. Author of [7] had reported that a great part of challenge is the modeling of interconnections of urban processes that are resulting in a complex spatial and dynamic behavior. To contribute in overcoming this challenge, a novel modeling is proposed which answers to how to model land use, population of an urban area and their dynamics and how to interconnect the four

mentioned urban processes in order to keep track of the outputs of the simulation over a calendar of time.

The paper is structured as follows: the work is put firstly into its context by presenting some related conventional and advanced RM models; secondly, it outlines specifications of advanced RM models, justifies the use of agent cellular automata approach and then presents the proposed novel modeling with its sub-models. Finally, a conclusion is given to highlight usefulness of this model for boosting modeling, coding and development of residential mobility applications.

## II. CLASSICAL AND ADVANCED RM MODELS

Residential mobility is a topic of large concern to urban researchers [8]-[14]. It concentrates on causes, effects and statistic rates of households' relocation decision of an urban area. A good understanding of causes, effects and rates would be of great help for urban planners and decisions makers.

Residential mobility models could be classified on two categories conventional and advanced models. Conventional models [9], [15]-[17] are concentrating on rates of mobility for different households categories. It uses generally cross sectional census data and tries to answer to questions such as why an urban area is representing a big rate of mobility, why a member or household category move frequently. These studies are limited in goals, and focus only on understanding drivers and effects of residential mobility for a given urban area.

With advancement in computer algorithms and GIS spatial analysis, advanced residential mobility models are gradually replacing conventional models. These models combine in addition to residential mobility other urban processes such as housing choice, urban growth, transport system, etc. [18], [19] [3], [4]. Models of this category use series of socioeconomic census data, households' census data, demographic data and spatial data. Models of these complex systems could be constructed only if urban systems are considered as spatial, dynamic, self-organizing and computational systems.

## III. SPECIFICATION OF ADVANCED RM MODELS

Residential mobility is a phenomenon that takes place in an urban city over a calendar of years. Households could decide yearly to move from one location to another. When they decide to move, then subsequently they decide to choose a new housing. Decision of relocation is an algorithm that uses households' learning census data and decides if a given household desire to move to a new housing or not. After decision of relocation comes Housing Choice which is also an

algorithm that uses socioeconomic census data and urban planning data such as properties and quality of housing and neighborhood and give each household the capability to choose from a selection of housings a better one. Structure of land use of the City could change yearly e.g., new built-up area, new industrial area etc. Change of infrastructure between two successive years could be introduced manually using planned infrastructure data or automatically using land use prediction algorithms. Planned infrastructure data are generally plans of new constructions of a City for a horizon of years e.g., 10 years. Land use prediction is an algorithm that uses series of past aerial or satellite imagery and then predicts land use change of next years. Population growth could be seen as an engine for residential mobility. Population grows, in fact households, for one reason or another (birth of a member, death of a member, employment, marriage, departure etc.). Population growth is also an algorithm that uses learning statistic and demographic data of households of the City and project next households' generations. Based on this description, the oncoming model has to give answers on how to model the urban space, population and their dynamics over a calendar of time. More specifically, the model has to keep track of households' movements, population growth, land use dynamics. So it could for each given year computes statistics data about mobility, population growth and urban expansion etc. The City should be considered as a result of interaction of individual smallest units i.e., households and spatial units particularly housings. In fact, the oncoming model will be a virtualization of the City where households interact with the environment e.g., districts and make residential actions such as relocation or housing choice. The interaction could not be done with a top-down approach; the need for a bottom-up approach is arisen. Next section will be devoted to discuss the utility of Agent Based Modeling and Cellular Automata bottom-up approaches.

#### IV. WHY AGENT CELLULAR AUTOMATA APPROACH

The City where urban processes take place, could be represented by two layers, one for individual households which move and choose housings, and the other for infrastructure units such as housings, transport segment, green space etc. to model urban processes, researchers have used a number of approaches which can be generally classified in two categories: top-down and bottom-up models as mentioned in [12].

With the continuous development of computer model, the bottom-up models have gradually replaced the top-down models in modeling urban dynamics. In a bottom-up approach, systems are considered as result of all smallest units' actions. Among these bottom-up approaches, cellular automata (CA) have been widely used for modeling and simulating infrastructure layer of urban areas [5], [20]-[23]. CA to project spatial forms of an urban area, it abstracts land field using a lattice of discrete cells, and presents the overall behavior from simple local rules. However, "bottom-up" CA models cannot incorporate decision of urban actors (e.g., households) and CA cells cannot move in space. This is why Multi-Agent Systems (MAS) are receiving currently much attention in urban modeling [24]-[27]. MAS could analyze potential actors of urban systems in the real world. MAS consist of a number of "agents" able to model interaction between humans and their

environment, and can make choices and decisions in response to this interaction. Individual behavior of these agents determines the behavior of the entire system. None of these approaches could solely represent the two layers (infrastructure layer and layer of urban actors e.g., households). A good modeling approach is a mixture of CA and MAS approaches. CA will represent urban space and its smallest units e.g., housings, transport segment, green space etc. MAS will represent actors of the urban space such as individual households or organizations.

#### V. THE PROPOSED MODELLING

The proposed modeling contains six models based UML 2.0 diagrams which model parts of the system from different views. All models have designed using Enterprise architect [28]. The first four models describe land use, population and their dynamics. The fifth model describe how interconnection of residential mobility, housing choice, land use change and population growth processes is done. The last one describe the structure of the main objects of the system which are the composite' objects responsible of tracking snapshots of the system over time.

##### A. Land use representation

To represent land use, a Cellular automaton based model is designed as in Fig. 1. Its basics elements are cells, states and 2D grid. Each cell will represent a spatial unit of the studied land use. Each cell has a state that represents soil occupancy. CA states number is finite and it is determined based on the classification of the studied land area. Each cell which is identified by geographic position (x and y) has neighbors depending on the neighborhood's type (e.g. Van Neumann, Moore, etc.) or neighborhood's range (e.g., range = 400 meter). The method `transitTo (States)` describe how a cell transit from a current state to another. The `Grid` class is a composition of `Cell` class. After code generation, the `Grid` will contain a `LinkedList` collection `<Cell>` to store all cells representing the land use. The `State` Collection is a composite of CA states. After code generation, `State` Collection will contain a `LinkedList` collection `<State>` to store states of the studied system. The CA model is generic and extensible, new derived models could be created depending on specification of newly studied land use in terms of CA states number, CA transition rules and land use map format etc.

##### B. Population representation

To represent population of the studied, an agent based model is designed as in Fig. 2. The central element of the MAS model is the Household (HH). This later is composed of one householder and many members. The HH lives in a Housing that is located in a cell of the Grid. The MAS model reuses an existent MAS framework (e.g., JADE). Thus, the HH extends a MAS agent of an extern framework. As result, each HH of the population is represented by a MAS agent which is interconnected with its correspondent housing.

##### C. Population dynamics

To represent population dynamics, the following model in Fig. 3 is proposed. The model describes how to track population generations of a studied area over a calendar of years in order to establish an observatory of population growth

but also an observatory of population movements. Each Household is stored with a unique identifier. Population class is a composition of Household class. After code generation, population class will contain a map collection Map<id,Household> to store all Households of a given time step.

Each Population is stored with a given iteration. PopulationArchive is a composition of Population class which

means in other words that the PopulationArchive contains all population generations. After code generation, PopulationArchive will contain a collection Map<iteration,Population> to store each population generation with its specific iteration. Depending on models, first population at time  $t_0$  comes from census data. Next generations of populations are projected using projection algorithms and learning demographic data.

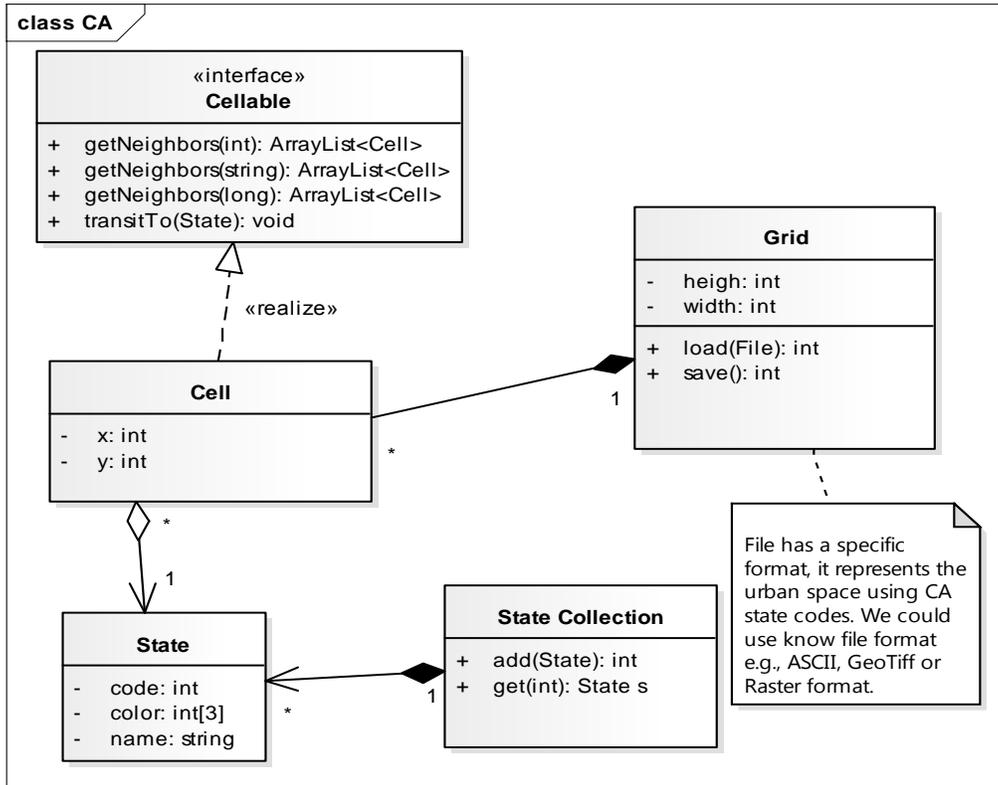


Fig. 1. CA model.

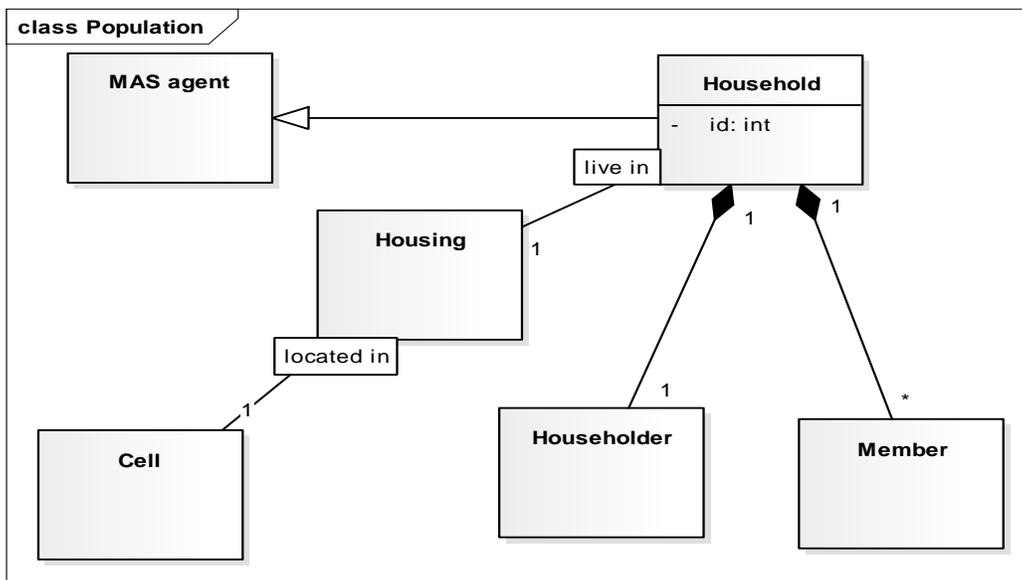


Fig. 2. MAS model.

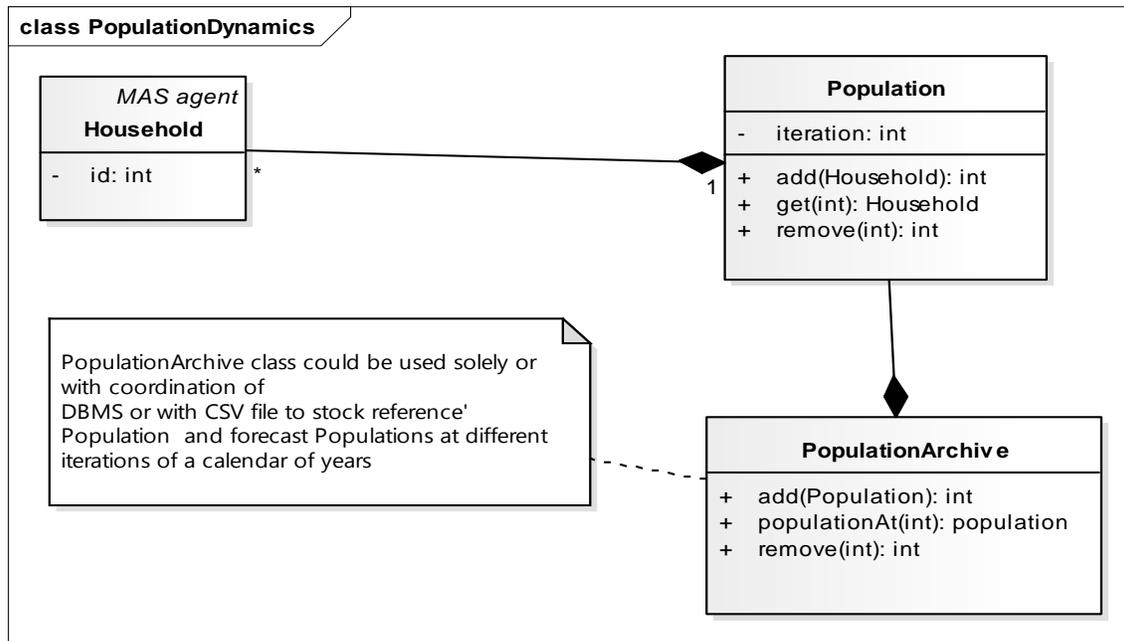


Fig. 3. Population growth model.

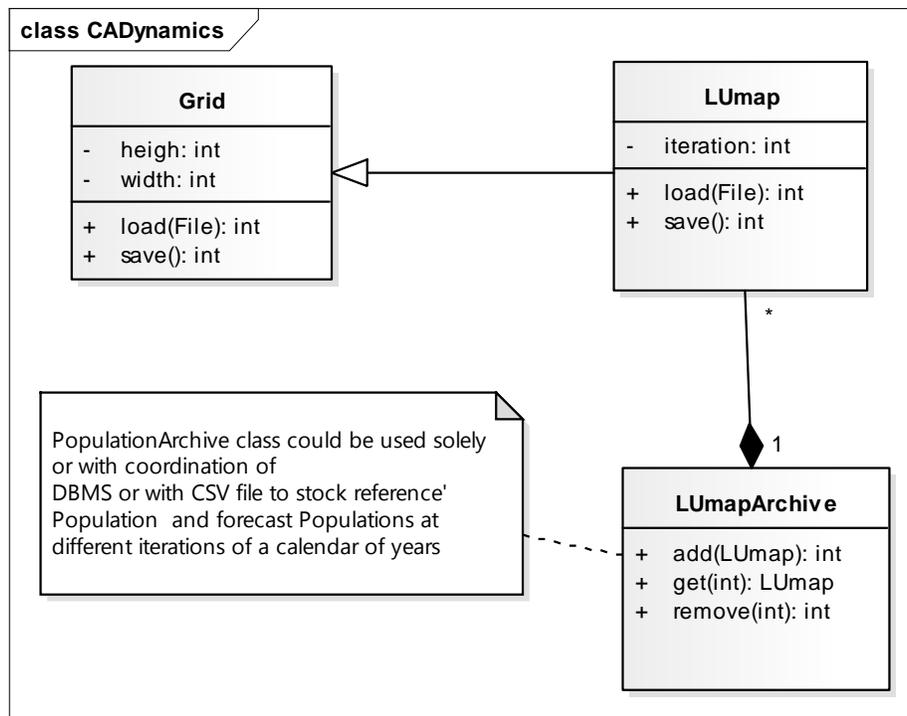


Fig. 4. Land use dynamics model.

#### D. Land use dynamics

To represent land use dynamics, the model in Fig. 4 is proposed. The model describes how to track land use dynamics of a studied area over a calendar of years. LUmapi class extends Grid class. Thus LUmapi will represent soil occupancy of the studied area in a given iteration. LUmapiCollection is a composite of many LUmapi. A Land use map at time  $t_i$  will change in  $t_{i+1}$ . So to track evolution of land use for a calendar

of years, LUmapiCollection will use a map collection of LUmapi instances  $\text{Map}\langle \text{iteration}, \text{LUmapi} \rangle$  to store LUmapi, each one with its specific iteration. As result, dynamics of land use is archived from its first year of simulation till last year. Depending on models, first LUmapi that represents soil occupancy of the studied area could be loaded from a file with a specific or known format e.g. ASCII, GeoTIFF, Raster or Vector format. Next LUmapi are forecasted using algorithms and learning socioeconomic and spatial data.

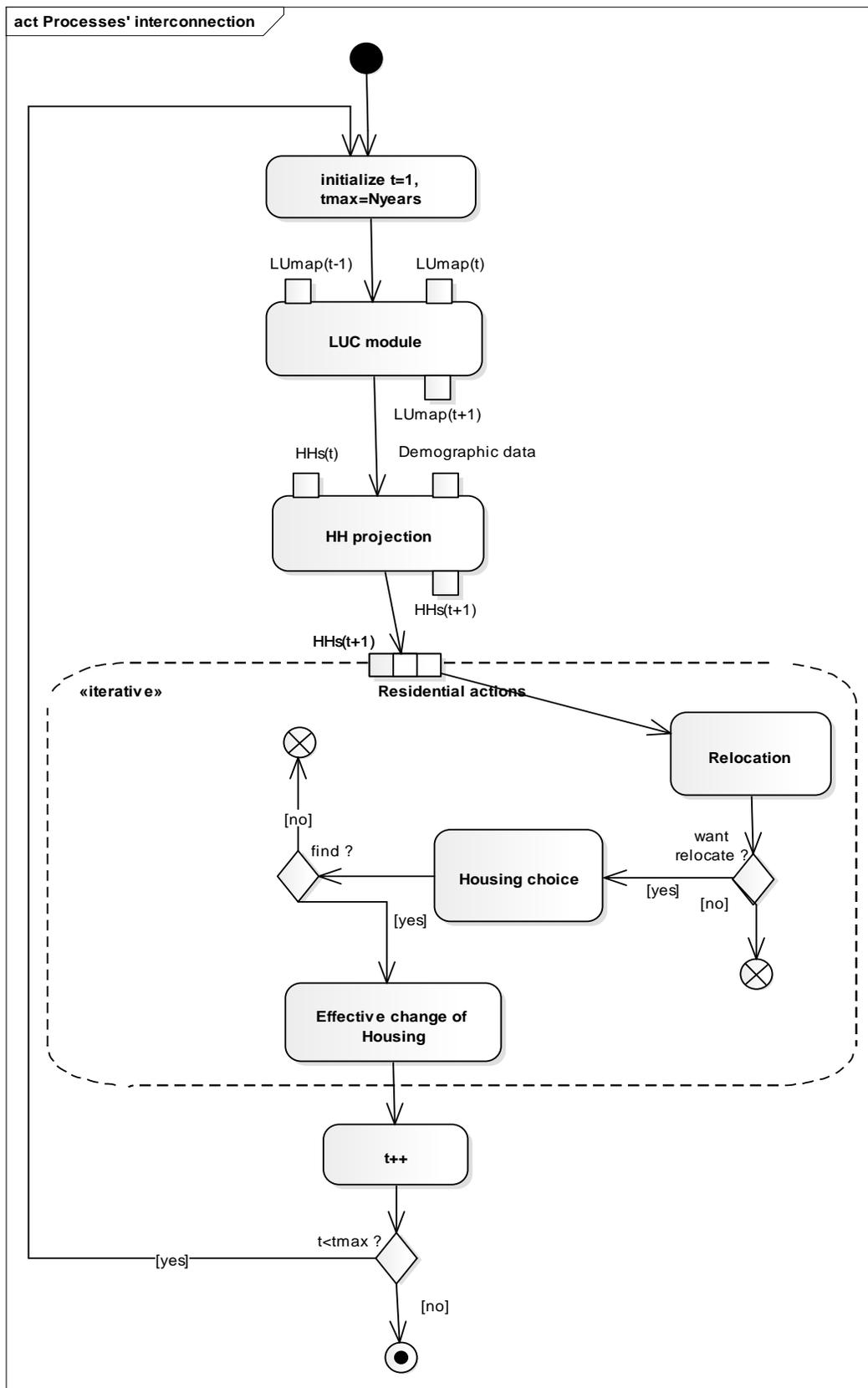


Fig. 5. Process interconnection's model.

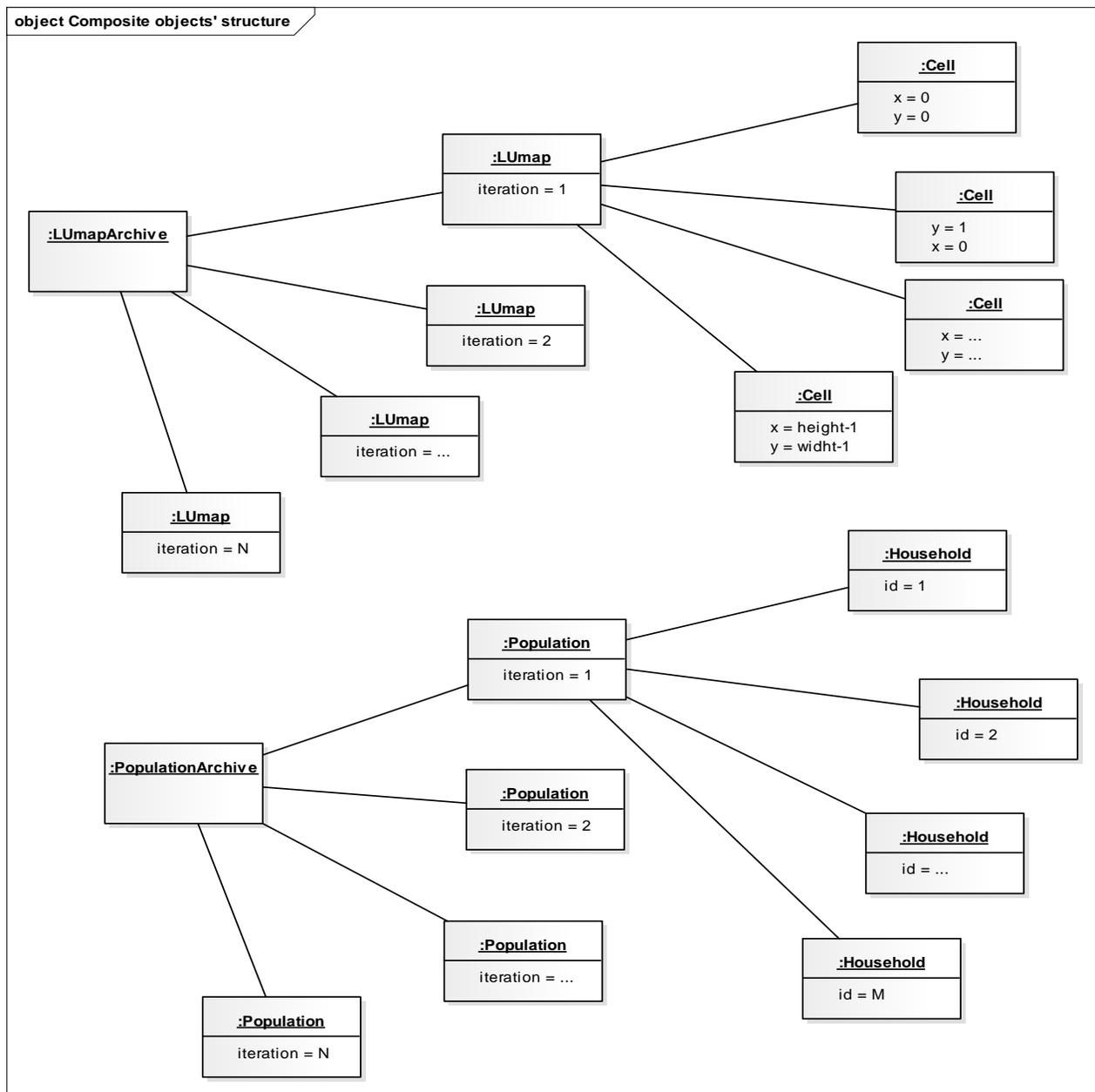


Fig. 6. Structure of composite' objects.

### E. Processes' interconnection

The model encapsulates four algorithms. To model their interconnection, the model in Fig. 5 is proposed. It describes object and control flow between algorithms which are represented by action entities. It is assumed that maps at time  $t_0$  and  $t_1$ , population at time  $t_1$ , demographic data indexes and census data are already loaded in the model. The model iterates for number of years previously parametered. In each iteration  $t_i$  the model begins by LUC module then HH projection to forecast HH at time  $t_{i+1}$ . LUC module could predict structure of next LUMaps from past LUMaps using prediction algorithms; however, in case of simulation based scenarios, next LUMaps could be parameterized manually in the model. The remaining expansion region will be executed for each HH from the list of HHs at time  $t_{i+1}$ . Relocation algorithm computes for each

household relocation propensity. Housing choice algorithm explores the neighborhood and look for the best suitable housing. All HH agents are fed by these residential actions. If a HH decide to relocate based on relocation propensity, and found a suitable housing then it will effectively change its housing.

### F. Structure of composite' objects

To represent objects' structure of the system, the model in the Fig. 6 is proposed, which depicts only composite objects. The model of the system contains one LUMapArchive which manage many LUMap objects. Each of which is a snapshot of the Land use at a given time. Each LUMap object contains a large number of Cell objects. Similarly, PopulationArchive contains many population objects. Each of which is a snapshot

of the population at a precise time. Each population object contains a large number of Household objects.

## VI. CONCLUSION

The paper presented related research works of residential mobility, which are classified on two main categories: the classical residential mobility models which are limited in goals in contrast to advanced residential mobility models that are more realistic and effective thanks to the integration of many urban processes. It outlined also specifications of advanced residential mobility applications which should be capable firstly of tracking households' mobility, households' growth, households' housing' choice and land use dynamics over a calendar of years and secondly of interconnecting urban processes on a bottom-up approach that give the model its power of micro simulation.

After that, the paper exposed the new advanced residential mobility modeling and depicted its way to handle space, population, and residential actions of individual households. What is interesting about the model is that it used a bottom-up mixed approach. The City, in which urban phenomena take place, is considered as an emerging and auto-organized system where the global system dynamics are a set of individual behavior of smallest units i.e., households and housing. The proposed modeling is of significance and it could be used to give insights on residential mobility modeling and to boost modeling, development and coding of ulterior advanced residential mobility applications. Further works are planned. The proposed modeling could be improved in a way to support more the Model View Controller (MVC) architecture. The proposed modeling is also planned to be the core of a future residential mobility simulation Framework.

## REFERENCES

- [1] El-arbi El-alaouy, Khadija Rhoulami, and Moulay Driss Rahmani, "Towards an agent based model for simulating residential mobility and urban expansion," Proceedings of the Mediterranean Conference on Information & Communication Technologies. Springer International Publishing, 2016. p. 343-351., 2015.
- [2] Dagmar Haase, Sven Lautenbach, and Ralf Seppelt, "Modeling and simulating residential mobility in a shrinking city using an agent-based approach," Environmental Modelling & Software, vol. 25, pp. 1225-1240, 2010.
- [3] Veronika Gaube and Alexander Remesch, "Impact of urban planning on household's residential decisions: An agent-based simulation model for Vienna," Environmental modelling & software, vol. 45, pp. 92-103, 2013.
- [4] Jean-Philippe Antoni and Gilles Vuidel, *MobiSim: un modèle multi-agents et multi-scalaire pour simuler les mobilités urbaines*, 2010.
- [5] Edwige Dubos-Paillard, Yves Guermond, and Patrice Langlois, "Analyse de l'évolution urbaine par automate cellulaire. Le modèle SpaCelle," L'espace g
- [6] Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, Keith Sullivan, and Gabriel Balan, "Mason: A multiagent simulation environment," Simulation, vol. 81, pp. 517-527, 2005.
- [7] Robert M. Itami, "Simulating spatial dynamics: cellular automata theory," Landscape and urban planning, vol. 30, pp. 27-47, 1994.
- [8] Frans M. Dieleman, "Modelling residential mobility: a review of recent trends in research," Journal of housing and the built environment, vol. 16, pp. 249-265, 2001.
- [9] Thierry Debrand and Claude Taffin, "Les facteurs structurels et conjoncturels de la mobilité résidentielle depuis 20 ans," Economie & Statistique, 2005.
- [10] René Jordan, Mark Birkin, and Andrew Evans, "Agent-based modelling of residential mobility, housing choice and regeneration," in Agent-based models of geographical systems.: Springer, 2012, pp. 511-524.
- [11] Arden Craig Brummell, "A theory of intraurban residential mobility behaviour," Ph.D. dissertation 1977.
- [12] Itzhak Benenson, "Multi-agent simulations of residential dynamics in the city," Computers, Environment and Urban Systems, vol. 22, pp. 25-42, 1998.
- [13] Nicolas Coulombel, "Residential choice and household behavior: State of the Art," Ecole Normale Sup
- [14] Ronghui Tan, Yaolin Liu, Kehao Zhou, Limin Jiao, and Wei Tang, "A game-theory based agent-cellular model for use in urban growth simulation: A case study of the rapidly urbanizing Wuhan area of central China," Computers, Environment and Urban Systems, vol. 49, pp. 15-29, 2015.
- [15] Helmuth Cremer, Henry-Jean Gathon, and others, "Les déterminants de la mobilité résidentielle: une analyse probit," Brussels Economic Review, vol. 115, pp. 53-75, 1987.
- [16] Sara Anderson, Tama Leventhal, and Véronique Dupéré, "Residential mobility and the family context: A developmental approach," Journal of Applied Developmental Psychology, vol. 35, pp. 70-78, 2014.
- [17] Ahmed Lahlimi Alami et al., *Les Cahiers du Plan N° 45 - Septembre / Octobre 2013 Agdal - Rabat*, 2013.
- [18] Paul Waddell and Gudmundur F. Ulfarsson, "Introduction to urban simulation: Design and development of operational models," in Handbook of transport geography and spatial systems.: Emerald Group Publishing Limited, 2004, pp. 203-236.
- [19] Itzhak Benenson, "Agent-based modeling: From individual residential choice to urban residential dynamics," Spatially integrated social science: Examples in best practice, vol. 42, pp. 67-95, 2004.
- [20] Kenneth Mubea, Roland Goetzke, and Gunter Menz, "Applying cellular automata for simulating and assessing urban growth scenario based in Nairobi, Kenya," International Journal of Advanced Computer Science and Applications, vol. 5, pp. 1-13, 2014.
- [21] Diana Mitsova, William Shuster, and Xinhao Wang, "A cellular automata model of land cover change to integrate urban growth with open space conservation," Landscape and Urban Planning, vol. 99, pp. 141-153, 2011.
- [22] Daru Kusuma Purba, "Implementation of Pedestrian Dynamic In Cellular Automata Based Pattern Generation," International Journal of Advanced Computer Science and Applications, vol. 7, pp. 65-70, 2016.
- [23] Kohei Arai and S. Sentinuwo, "Spontaneous-braking and lane-changing effect on traffic congestion using cellular automata model applied to the two-lane traffic," IJACSA) International Journal of Advanced Computer Science and Applications, vol. 3, 2012.
- [24] Hasan Omar Al-Sakran, "Intelligent traffic information system based on integration of Internet of Things and Agent technology," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 6, pp. 37-43, 2015.
- [25] Fatimazahra BARRAMOU and Malika ADDOU, "An agent based approach for simulating complex systems with spatial dynamics application in the land use planning," IJACSA) International Journal of Advanced Computer Science and Applications, vol. 3, 2012.
- [26] Harjot Kaur, Karanjeet Singh Kahlon, and Rajinder Singh Virk, "Migration dynamics in artificial agent societies," Int. J. Adv. Res. Artif. Intell, vol. 3, pp. 39-47, 2014.
- [27] Verda Kocabas and Suzana Dragicovic, "Bayesian networks and agent-based modeling approach for urban land-use and population density change: a BNAS model," Journal of geographical systems, vol. 15, pp. 403-426, 2013.
- [28] Entreprise architect Sparx, <http://www.sparxsystems.com/>

# Intelligent Diagnostic System for Nuclei Structure Classification of Thyroid Cancerous and Non-Cancerous Tissues

Jamil Ahmed Chandio, M. Abdul Rehman Soomrani  
Department of Computer Science  
Sukkur Institute of Business Administration, Sukkur IBA  
Sukkur, Pakistan

**Abstract**—Recently, image mining has opened new bottlenecks in the field of biomedical discoveries and machine learning techniques have brought significant revolution in medical diagnosis. Especially, classification problem of human cancerous tissues would assume to be one of the really challenging problems since it requires very high optimized algorithms to select the appropriate features from histopathological images of well-differentiated thyroid cancers. For instance prediction of initial changes in neoplasm such as hidden patterns of nuclei overlapping sequences, variations in nuclei structures, distortion in chromatin distributions and identification of other micro-architectural behaviors would provide more meticulous assistance to doctors in early diagnosis of cancer. In-order to mitigate all above stated problems this paper proposes a novel methodology so called “Intelligent Diagnostic System for Nuclei Structural Classification of Thyroid Cancerous and Non-Cancerous Tissues” which classifies nuclei structures and cancerous behaviors from medical images by using proposed algorithm Auto\_Tissue\_Analysis. Overall methodology of approach is comprised of four layers. In first layer noise reduction techniques are used. In second layer feature selection techniques are used. In third layer decision model is constructed by using random forest (tree based) algorithm. Finally result visualization and performance evaluation is done by using confusion matrix, precision and recall measures. The overall classification accuracy is measured about 74% with 10-k fold cross validation.

**Keywords**—Machine learning; decision support system; clustering; classification; cancer cells

## I. INTRODUCTION

Recently Image mining has become one of the well-established research area(s) of (ML) machine learning and (AI) artificial intelligence based techniques are vastly used in healthcare industry to facilitate doctors during the diagnostic and prognostic process of various kinds of malignant diseases such as breast cancer, lung cancer, thyroid cancer and so on. This paper addresses the classification problem of well-differentiated thyroid cancerous and non-cancerous cells of human tissues, which are carefully and systematically acquired from DICOM (Digital Communication in Medicine) images (Fig. 1). As per general observation, the misdiagnosis is one of the leading causes for rapid proliferation of cancer incidences among the world population and various related

research approaches [1], [2] and [4] have been seen to reduce the chances of miss-diagnosis. All above stated approaches are providing very nice services to solve the classification problem of thyroid cancer. Since the prediction of initial changes in neoplasm, such as hidden patterns of nuclei overlapping sequences, variations in nuclei structures, distortion in chromatin distribution of human cells and other micro-architectural behaviours would provide more meticulous assistance to doctors in early diagnosis cancer. By considering above stated problems, this paper proposes a data preparation algorithm so called Auto\_Tissue\_Analysis which has to perform three task (1) NSA (Nuclei Seed Approximation) because gradient base techniques generate multiple seeds due to existence of high intensities at multiple places between single object. (2) TSA (Tissue Structure Approximation) function that helps to find out the minimum mean distances between the nuclei seed points on a distance matrix where particular arrangement of nuclei specifies a particular point based sequences to declare a thyroid cancer type. For example, anaplast cancers [Fig. 1(d)] are most aggressive cancers and consists upon the defused set of nuclei having varying distances at nuclei structural level and papillary carcinoma nuclei structures are likely to be found as finger like shapes, but it become more confused [Fig. 1(b)] when these features are presented with thyroid insular carcinoma. Every cancer is treated with different therapies, since it is very essential to know the cancer type for proper prognosis of cancer and restoration of human health. (3) NFST (Nuclei Feature Selection Tray) function avoids manual object cropping and it helps to selects the appropriate behaviours from every DICOM image. In-order to mitigate all above stated issues this paper offers a CAD (Computer Added Diagnostic) system so called “Intelligent Diagnostic System for Nuclei Structural Classification of Thyroid Cancerous and Non-Cancerous Tissues” to provide assistance to doctors by recommending a second systematic opinion. The system classifies cancerous and non-cancerous nuclei by considering a system generated decision variable used as sub-class label attribute in every observation. Since the system generated recommendations provides more precise assistance to doctors to understand the hidden behaviours of cytological material which may enable the doctors to address the all types of thyroid cancer as stated above.

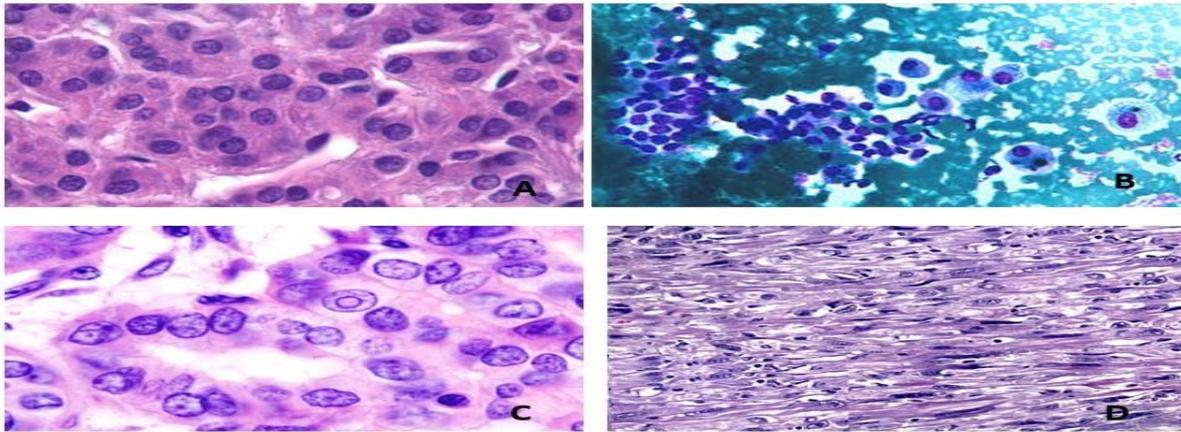


Fig. 1. Complex nuclei structures (a) Follicular (b) mixed papillary (c) papillary (d) anaplast cancers.

The methodology of proposed system comprises upon four layers and each layer is responsible to interact with each other. First two layers are responsible for data preparation. In first layer; noise reduction and classification of nuclei sequences are performed, in second layer auto-feature selection and auto-feature extraction are done, in third layer classification model have been constructed by using random forest tree based algorithm and in fourth layer the classification accuracy have been shown.

This paper is organized in several sections. Section one is used to describe introduction of this paper, related works are presented in section two, methodology is defined in section three, results are presented in section four and discussion & conclusion is discussed in section five whereas future work is presented in section six.

## II. LITERATURE REVIEW

Classification of histopathological DICOM (Digital Communication in Medicine) images is of one the active research area(s) of ML (Machine Learning) and many approaches have been proposed to solve the classification problem of malignant diseases. This paper proposes an approach which is considered as productive modelling for thyroid disease classification in the domain of machine learning. Some of the related works have been presented as bellow.

A Comparison of different three ML (Machine Learning) algorithm [1] such like artificial neural network, decision tree and logistic regression were proposed to classify follicular thyroid cancer. This paper proposes an algorithm [Auto\_Tissue\_Analysis] to decide about the sub- class label attribute and used as decision variable and it is capable to identify all the types of thyroid cancer nuclei groves / structures, since DICOM structure identification at micro-architectural level would provide more refine results. Random forest tree base machine learning technique is used to classify cancerous and non-cancerous nuclei systematically with support of fully automated phenomena. A system [2] was proposed to classify cancerous thyroid disease tissues. Convolutional neural network based machine learning technique was used. In every DICOM image nuclei is key building block component and we approximate overlapped

nuclei octo-edge distance value analysis compared with the equal to the size of neighbour set of nuclei. In most of approaches gradient base techniques have been used in segmentation stages, such kind of process would generate more than one centroids for each nuclei due to high intensity levels of pixels in nuclei regions. Proposed pre-processing algorithm selects the coordinates by considering size of each nuclei as per minimal set of edges between the set of neighbouring nuclei, if more than one edges are detected because of overlapping, regions are broken by measuring the historical data existing in the same DICOM image.

A comparison [3] of three machine learning techniques was proposed to classify the thyroid papillary cancer such as K-NN, Naive Bayesian and VPRS-CMR. In segmentation phase Otsu method is used and global level features were acquired. Micro-architectural level of nuclei detection would enhance the results and it provides more assistance to construct an unsupervised classification model for structure analysis.

A system [4] for thyroid cancer classification was presented and decision model was constructed by using support vector machine. As per their shown results malignant lesions lies between 0.97-80,  $P < .0001$ . In proposed pre-processing algorithm, the size of every nuclei is quantified with central location and compared with the nearer location by using unsupervised classification of nuclei structural states kwon as decision variable.

Using ultrasound images thyroid disease classification system [5] was proposed. Local binary patterns (LBP) base features were acquired through ROI (region of interest) feature selection was performed. There are certain limitations are associated with pixel base manual ROIs since the user may lose the important information of the content due human handling movements.

A system [6] was proposed for thyroid cancer and Otsu threshold was used to segment the nuclei from DICOM images. Using proposed method nuclei objects were detected by considering the nuclei rings and decision model is constructed for identification of set of nuclei grooves and use different levels of feature with the assistance of colour movements in DICOM image.

In an approach [7] six support vector machines were used to acquire ROI based features for thyroid disease using ultrasound images. All above stated approaches are providing very nice services but most of them consider thyroid cancer classification problem [8]-[13] at abstract levels, since thyroid disease classification may be more efficiently solved by considering the nuclei seed analysis, tissue structure analysis and nuclei feature selection tray, because structure identification needs special technique to understand nuclear grooves and arrangement of nuclei sequences, overlapping and other behaviours have significant importance in cancerous cell diagnosis. In literature thyroid follicular, medullary, papillary cancers have been classified. This paper proposes a novel technique which offers generalize algorithms for diagnosis of all thyroid cancers by proposing [Auto\_Tissue\_Analysis] algorithm which identifies all types of nuclei structures and other behaviours as stated above. Thus it is necessary to acquire the unsupervised classification by considering the nuclei regions / manual cropping with automated procedure, because manual cropping may select extra or less features from the regions DICOM [8]-[10] image. Secondly, follicular [14], [15], medullary, papillary sub cancer types have been classified in literature, whereas our proposed approach also classifies anaplastic and benign cancerous cells and tissues.

### III. METHODOLOGY

The proposed approach is basically a predictive modelling in machine learning and deals with the classification problem of the histopathological cancerous cells and tissues that are taken from the DICOM images of FNAB [17], [18], and [19]. Using proposed methodology; thyroid papillary, follicular, anaplast and benign cancerous cells have been quantified. Due to the complexities of different nuclei staining materials (such as H, E, Biomarkers and so on) that are being used by cytologists to label the nuclei and to observe the heterogenous behaviours of nuclei such as different shapes, sizes, colour schemes, structures because micro architectural structures are observed always with dissimilar patterns. The methodology of proposed system (Fig. 2) is divided into four interconnected layers where each layer is assigned several inputs to interact with each other. In first layer; edges of nuclei are detected with support of noise reduction techniques and regions are formed, in second layer unsupervised classification of nuclei structures is performed by using seed analysis and auto-feature selection is done by NFST (Nuclei Feature Selection Tray). A decision variable S is derived from the nuclei sequences by using our proposed algorithm so called Auto\_Tissue\_Analysis [Algorithm 1]. In third layer classification model is constructed by using random forest tree based algorithm and in fourth layer the classification accuracy have been measured. Proposed pre-processing algorithm [Auto\_Tissue\_Analysis] is fully automated algorithm and it classifies the cancerous and non-cancerous nuclei derived from mimic, mix and confused pattern of DICOM images of FNAB. since these micro-architectural components produces lots of confusions related to nuclei structures, because nuclei sequences have significant importance to declare specific anatomical class of thyroid cancers such as well-differentiated, un-differentiated, poorly differentiated and benign as a global ontology of cancers.

```

Algorithm 1: Auto_Tissue_Analysis
Input: DICOM image
Output: Spatial location, Size, height, width of each nuclei, Nuclei Tray, Features, Decision Variable \Nuclei Feature Selection Tray
D ← HCV DICOMnuclei components for Binarization B
Visit ← each pixel as qn(xi, yj) \Start NSA (Nuclei Seed Approximation)
for each σ2w(t) ← q1(t)σ12 + q2(t)σ22(t)
    q1(t) ← ∑i=1t P(i) ∧ q1(t) = ∑i=t+1t P(i) - μ1(t) - ∑i=1t  $\frac{IP(i)}{q_1(t)}$  ∧ μ1(t) = ∑i=t+1t  $\frac{IP(i)}{q_1(t)}$ 
    σ22(t) ← ∑i=1t [i - μ1(t)]2  $\frac{IP(i)}{q_1(t)}$  ∧ μ1(t) = ∑i=t+1t [i - μ1(t)]2  $\frac{IP(i)}{q_1(t)}$ 
    for each Pi ∈ q1(t) do
        K ← g(x, y)q1(t) ≤ 0n(k) ≤ 1
        ROI ← Radius μ(g(x, y)q1(t))
        if g(x, y)q1(t) = 1
            Count ← pi 1 + + |spatial locations (x, y)|
        if g(x, y) = 0
            Count ← pi 0 + + |measure nuclei|
        lj = ∑xi ∈ cj ||xi - μj||2 \Construct distance matrix
    end if
    for each Pi ∈ q1(t) do
        Break Nuclei ← pi 1 + +
    end if
Return μ ← Seed ← Height ← eed, Height, Width, μR, μG, μB, μRGB, size, Chro, Nucleiring, Nucleiforeground = g(x, y)q1(t)
    for each Xi = Nearer ∑xi ∈ cj ||xi - μj||2 do \Extraction of sub decision variable
        if Near = papillary < 0.30 max numerical distance
        if Near = Follicular < 0.40 max numerical distance
        if Near = Anaplast < 0.50 max numerical distance
        if Near = Benign < 0.60 > .70 max numerical distance
    end if
Return μ ← Class (papillary, Follicular, Anaplastic, Benign) \ decision variable
    
```

#### Layer 1: Noise Reduction & formation of regions

For example well differentiated cancer class consists upon the follicular, papillary and other types of cancers. Papillary class can be further subdivided into papillary tall cell carcinoma, glass ground appearance of cells [Fig. 1(c)] and so on.

#### A. Noise Reduction

Set of Nuclei groves contain sufficient information to classify despite of having heterogenous behaviours and shapes. In literature lots of noise reduction techniques have been proposed. Some of them use image segmentation techniques (such as Watershed, Graph-cut, Super pixels and so on) meanwhile our technique is motivated from the graph-cut segmentation because colour information found in DICOM images is really important for doctors during the diagnostic process traditionally. For example in [Fig. 1(a-d)] lots of noise is persisting and nuclei sequences are heterogenous, especially; [Fig. 1(b)] is very difficult to interpret traditionally either it is belonging to papillary cancer class or other classes. Evidently; on applying of graph-cut segmentation algorithm, we observed that most significant object were detected as fore-ground and the remaining objects were removed as back-ground (Fig. 3). This phenomena is not only supporting to reduce the noise but also efficient to saved time and space complexities.

#### B. Grey Scale Conversion

For further image analysis [20], [21], DICOM images were converted into binary form, we use adaptive threshold algorithm as shown in [Fig. 4(a)] to covert the image into binary format. Since the Otsu method has several advantages in binary representation of medical image [22] because clustering base unsupervised classification of each pixel p = {x, y<sub>1</sub> ... .. x, y<sub>n</sub>} have to be represented as an object known The method follows the weighted pixels corresponding to either one or zero class as presented in (1) and (2).

$$\omega_0(t) = \sum_{i=0}^{t-1} p(i) \tag{1}$$

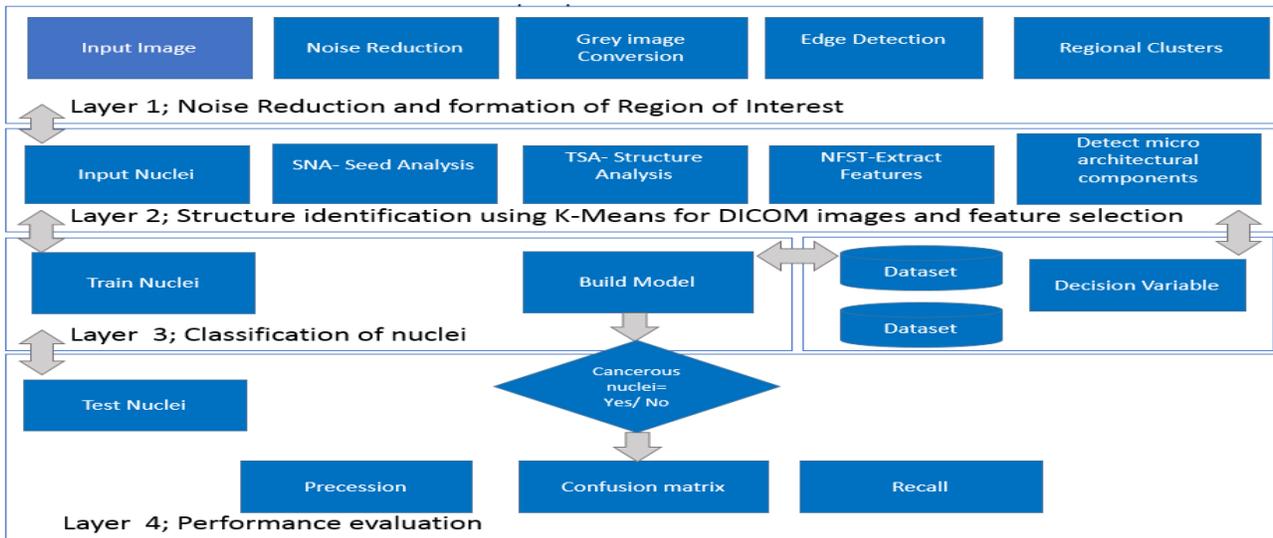


Fig. 2. Intelligent diagnostic system for nuclei structural classification of thyroid cancerous and non-cancerous tissues workflow.

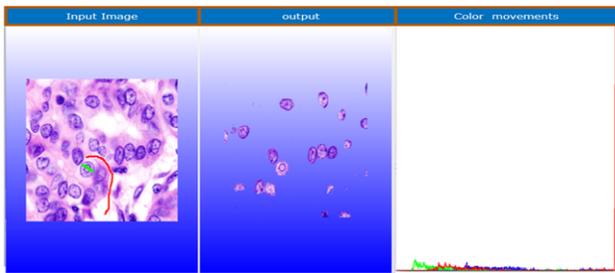


Fig. 3. Noise Reduction using graph-cut segmentation

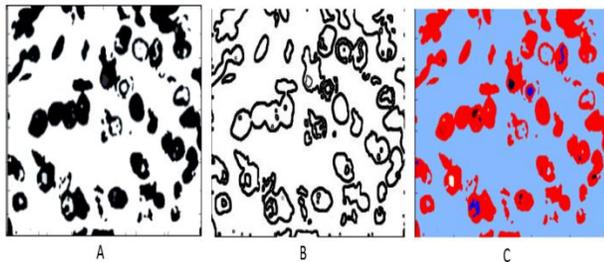


Fig. 4. (a) image banalization using adaptive threshold (b) Canny edge detection (c) Nuclei clusters

$$\omega_1(t) = \sum_{i=t}^{L-1} p(i) \quad (2)$$

The covariance of these corresponding pixel classes would be represented as per (3), where each label has significant impact upon the generated pixels.

$$(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t) \quad (3)$$

So two regions may be constructed for classification of pixels and some regions would be appeared with the pixels having the value 1 whereas others may be represented with 0 quantity as shown in (4), (5) and (6)] where mean values are potted to visualize the  $w\mu_0 + w\mu_1 = \mu T$  and  $w_0 + w_1 = 1$

$$\mu_0(t) = \sum_{i=0}^{t-1} ip(i)/\omega_0 \quad (4)$$

$$\mu_1(t) = \sum_{i=0}^{L-1} ip(i)/\omega_1 \quad (5)$$

$$\mu_T = \sum_{i=0}^{L-1} ip(i) \quad (6)$$

Canny edge detection algorithm was used to find out the edges of all nuclei, since geometrical and morphological features have great importance in machine learning to let the classifier learn about the possible shapes of different objects as shown in [Fig. 4(b)] where each object is considered as region. Since the pixel base orientation is not enough to represent a medical component, therefore regions [Fig. 4(b)] are useful tool to apply further operations such as extracted centroid may be used for further operations as per Fig. 6.

### Layer 2: Structure identification and feature selection

Appropriate Nuclei seed approximation is required to find out the spatial coordinates such as x and y but at pixel level the energy of the pixel may be found from 0 to 255. If we fix some of the ranges of DICOM images intensities we may find more the one point for each class, since our algorithm considers every nuclei as a region, detected by using canny edge detection algorithm in following way.

#### 1) Nuclei Seed Approximation (NSA)

Proposed nuclei separation technique is inspired from the LBP (Linear Binary Patterns) at regional levels because there are several limitations associated with gradient base techniques because due to same intensity threshold may detect multiple cancroids for same nuclei such as DNA feature measuring algorithm. Let's consider a ring is created around the nuclei using canny edge detection so called region of interest and we have to acquire the radius of all regions [Fig. 5].

Above expression is plotted with eight points on the detected edges [23] region ROI= {  $a_1, a_2, \dots, \dots, a_7, a_8$  } where radius of each nuclei is the result of cross correlation and division of opposite points as per following equations [(7), (8) and (9)].

$$\sum_{a=8}^{a=0} g(x, y) \quad (7)$$

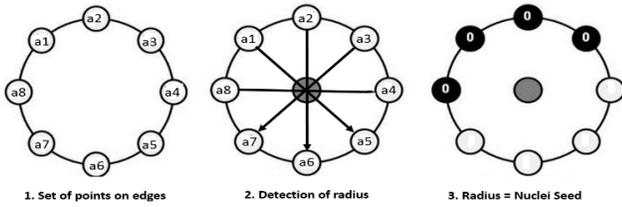


Fig. 5. Seeds extraction process.

$$ROI = R_{adius} \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\} \quad (8)$$

$$\text{Where Radius} = \frac{\mu\{a_1 + a_5\}}{\mu\{a_2 + a_6\}} \left| \frac{\mu\{a_3 + a_7\}}{\mu\{a_4 + a_8\}} \right| 8 \quad (9)$$

The father analysis on these seeds is performed in (Section 3.4) to visualize the probabilities of different cancer sequences.

### 2) Tissue Structure Approximation (TSA)

DICOM distance matrix (Fig. 6) is very useful tool to measure the pixel coordinates where individual distances of each nuclei are considered by using dot product of spatial location x and y. By summarizing all the distances of nuclei and constructing a decision model to find out the nuclei sequences designated as S would assist to find the optimal sequences of nuclei. Since the optimal values of f(k) may be quantified which lies between  $M = \{m_1 \dots \dots m_n\}$  distances of pixels of each object by considering edges of each nuclei. The ideal central location of each nuclei in DICOM dataset D is lying between the  $x^1, y^1$  locations of  $g(x)$ , which may be detected by considering the mean pixel values on gradient of every geometrical shape of nuclei. After acquiring the K values of each location the same may be separated by considering every nuclei as a partition. For example DICOM datasets with  $m^n$  number of nuclei could be considered in same group lies between locations 1...n but the central point is more significant to separate piece of medical image in an automated process. We may analyse k values of n objects designated as spatial location as decision variable and that may consider the clustering points by the adjustment of k clusters. Since the DICOM nuclei structures is an study of nuclei connected components could be used as decision variable S and central points would be assumed as following notation eq(12), where sk deviation of distances could achieved as global impact of clusters.

$$I_j = \sum_{x_i \in C_j} \|x_i - \mu_j\|^2 \quad (10)$$

The global impact of all clusters' distortions is given by the quantity

$$S_k = \sum_{j=1}^K I_j \quad (11)$$

The function of f (k) becomes optimal for DICOM structure as per following notation eq. (12).

$$f(k) = \begin{cases} 1 & \text{if } K < .30 \\ \frac{S_k}{a_k S_{k-1}} \text{if } S_{k-1} \neq 0, \forall K < .40 \\ 1 & \text{if } S_{k-1} = 0, \forall K < .50 \\ & \text{if } S_{k-1} = 0, \forall K < .60 \\ & \text{if } S_{k-1} = 0, \forall K < .50 > .70 \end{cases} \quad (12)$$

Display Data	1	2	3	4	5	6	7	8	9	10	11
1	0	709.0	759.0	128.15615...	342.81481...	341.97660...	399.29735...	183.86951...	169.83815...	170.32028...	170.1
2	705.0	0	54.0	589.12562...	369.53619...	370.66561...	1086.5656...	846.73077...	621.51427...	622.71181...	623.91024...
3	759.0	54.0	0	642.94712...	423.08509...	424.19924...	1140.2074...	906.14721...	674.27294...	675.45688...	676.64770...
4	128.15615...	589.12562...	642.94712...	0	221.00276...	220.00909...	498.14455...	261.88547...	90.027773...	91.219515...	92.417530...
5	342.81481...	369.53619...	423.08509...	221.00276...	0	1.4142135...	717.14573...	477.25255...	255.28219...	256.56383...	257.84685...
6	341.97660...	370.66561...	424.19924...	220.00909...	1.4142135...	0	716.04469...	476.10503...	254.00196...	255.28219...	256.56383...
7	399.29735...	1086.5656...	1140.2074...	498.14455...	717.14573...	716.04469...	0	242.0	471.10169...	472.12816...	471.15284...
8	183.86951...	846.73077...	900.14721...	261.88547...	477.25255...	476.10503...	242.0	0	231.21634...	230.52689...	229.31919...
9	169.83815...	621.51427...	674.27294...	90.027773...	255.28219...	254.00196...	471.10169...	231.21634...	0	1.4142135...	2.8084771...
10	170.32028...	622.71181...	675.45688...	91.219515...	256.56383...	255.28219...	472.12816...	230.52689...	1.4142135...	0	1.4142135...
11	170.32028...	623.91024...	676.64770...	92.417530...	257.84685...	256.56383...	471.15284...	229.31919...	2.8084771...	1.4142135...	0
12	170.37842...	625.10959...	677.82741...	93.621978...	259.13124...	257.84685...	470.17975...	228.37031...	4.2426006...	2.8284271...	1.4142135...
13	669.96029...	162.86497...	187.52333...	544.03308...	329.51176...	330.15148...	1026.3293...	784.43100...	553.23141...	554.20303...	555.17654...
14	690.74163...	171.60710...	190.11838...	564.47143...	351.06979...	351.08167...	1044.7660...	802.99688...	571.78754...	572.73466...	573.68371...
15	529.96698...	1216.0010...	1269.4061...	629.42910...	846.51343...	845.36915...	135.14806...	369.27090...	597.77753...	596.69925...	595.62740...
16	636.63254...	234.34163...	262.80030...	508.81135...	306.47185...	306.81101...	978.44212...	737.57236...	507.13706...	507.88622...	508.83789...
17	700.48911...	217.34764...	232.93775...	572.97207...	366.26902...	366.72196...	1044.54633...	803.61495...	573.06195...	573.92159...	574.78343...
18	701.74710...	218.16049...	233.46620...	574.21685...	367.66984...	368.06113...	1045.6256...	804.21677...	574.18938...	575.04434...	575.90400...
19	842.39658...	219.14054...	231.15795...	712.72410...	503.98908...	504.33398...	1189.8600...	948.97079...	718.28824...	718.16086...	720.03772...
20	366.23680...	478.85279...	526.85881...	245.53918...	186.96791...	185.72291...	666.84705...	429.29360...	207.50603...	207.56634...	208.43224...
21	524.13738...	334.46275...	374.66918...	396.55337...	275.53981...	275.70902...	848.89810...	609.84251...	381.79182...	382.30339...	383.21925...
22	525.48073...	334.46225...	374.52503...	397.70592...	226.92730...	226.64840...	850.01764...	610.80438...	383.01566...	383.72516...	384.43855...

Fig. 6. Distance matrix by using x, y location of nuclei.

In the above equation  $N_d$  are the number of attributes in decision matrix where  $a_k$  is the weighted distance points such as papillary structures with mean distances of  $<0.30$ , follicular with  $<0.40$ , anaplastic cancers  $<0.40$  and benign with  $<0.50 >0.70$ . Above 0.70 values were eliminated due to no significant set of object could be considered for partition as a DICOM structure class. In results section [Fig. 8] is shown to represent the distance matrix and [Fig. 9] is visualized to show the optimum percentage of split for each DICOM decision variable.

### 3) Nuclei Feature Selection Tray (NFST) algorithm

Manual selection of ROI (region of interest) and cropping image into sub-images as feature selection is labourous work and there are number of limitations are associated with manual cropping / ROIs because user may select reduce set of features or human handling may not crop properly. Thus there are fare chances for the loss of useful information. NFST (Nuclei Feature Selection Tray) algorithm avoids manual cropping and ROI (region of interest) of nuclei selection, thus lots of laborious work of feature selection is saved in terms of time and complexity. Besides we have presented results of our approach from the perspective of feature selection in Fig. 6 so called Nuclei Feature Selection Tray (NFST) algorithm. Let's consider every DICOM image is consisting upon the several number of nuclei where  $DICOM = \{D^1, D^2, \dots \dots \dots D^n\}$  on feature vector space. The spatial coordinates at this stage are really challenging to separate because overlapping of nuclei does not allow finding individual nuclei from a set of nuclei. At this stage we recorded sizes of each object and approximated the overlapped nuclei size by constructing spatial boundaries between the nuclei to count them as closed individual object. After acquiring the individual set of nuclei as objects we recorded colour movements with RGB mean values. Due to in-availability of such datasets for DICOM images of FNAB, we prepared our own datasets for training and testing purposes. The datasets comprise upon RGB movements [24]-[26] of nuclei, sizes, shapes and sub-class (SC) decision derived from (Section 3.3), which provides second opinion to doctors that what is the structure of particular DICOM image.

Therefore central point p of nuclei in a DICOM image is unique location containing x, y may be represented by using distance matrix (Fig. 6) where distance of each nuclei may be measured by using the [(13) and (14)] where height (h) and width (w) may be cropped with by considering the central

point known as seed.

$$h, w = \sum_{a(y)=n}^{a(x)=1} + \mu h/w \quad (13)$$

$$seed(\mu x, y) = \{Sum(a(x, y) \frac{1}{5}, \dots, \frac{4}{8})\} \quad (14)$$

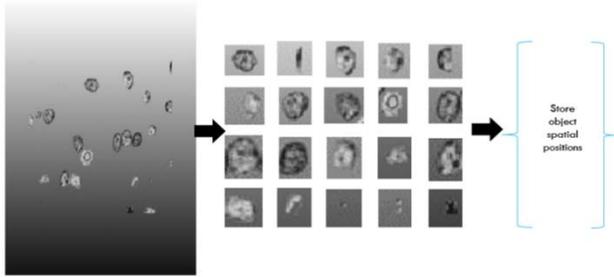


Fig. 7. Auto Nuclei Feature Selection Tray (NFST) for diagnosis.

### Layer 3: Classification of nuclei

A wide variety of machine learning algorithms is available since the feature engineering needs special contributions to select the related observations from the medical images. This research uses random forest machine learning algorithm for classification of cancerous disease because tree base algorithms have been found with higher rank relationships and deep decision advantages.

Let's suppose medical image dataset D contains a number of significant class instances designated as  $D = \{(X_i, Y_i)\}_{i=1}^n$  to construct a decision model by using the aggregated measured variables fitted values during the training phase by considering the  $J_{th}$  feature in terms of deep decision trees based upon the aggregated permutation functions since the out of bag errors have been used to represent overall trees.

$$D = \{(X_i, Y_i)\}_{i=1}^n \quad (15)$$

$$\hat{y} = \sum_{i=1}^n W(x_i, \hat{x}) y_i \quad (16)$$

In (16) and (17)  $\hat{y}$  variable is used to represent the cancerous and non-cancerous disease where x is the collection of formulized weighted and associated variables designated as W. the rest of function have assigned several number of class label attributes to classify by considering the highest level of tree base ranks.

$$\hat{y} = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n W_j(x_i, \hat{x}) y_i = \sum_{i=1}^n \left( \frac{1}{m} \sum_{j=1}^m W_j(x_i, \hat{x}) \right) y_i \quad (17)$$

### Layer 4: Performance Evaluation

In performance evaluation we use confusion matrix (Table 1). The classifier was trained by parsing total number of 602 observations, we selected 60% for training and 40% for testing purposes, where (Table 2) 218 instances were classified as true positive out of 301 and 813 instances were classified as true positive out of 857 instances. The precision (18) and recall (19) was recorded for cancerous classes 83.2%, 73.2% and 91.00%, 94.90% was measured for non-cancerous classes.

$$Precision = \frac{NumberofTrePositives}{NumberofTruePositives+FalsePositives} \quad (18)$$

$$Recall = Sensitivity = \frac{TruePositive}{TruePositive+FalseNegative} \quad (19)$$

$$Sepecify = \frac{TrueNegative}{TruePositive+FalsePositive} \quad (20)$$

## IV. RESULTS

Fig. 7 is plotted to present the results of fully automated object separation process where each nuclei is fully cropped without human interaction. In Fig. 8, Column 4 is describing the presentation of distance matrix quantities in a leaner set of nuclei sequences, where each nearer colour have been plotted in blue spectrum to form the identity matrix whereas deviated values on an identity matrix are presented with different colour scales ranging from 120-540 colour schemes where most blue colour is object value. Pre-process method have been illustrated in Fig. 8 where column 1 is showing input image and column 2 demonstrates TSA estimation where each most probable centre of normal and abnormal nuclei is detected. Further pre-processing of initial guess about the spatial locations have been depicted in column 3 where group of nuclei / nuclei grooves have been captured with the assistance of NSA algorithm. Column 4 presents the results of extracted nuclei for measuring the nuclei grooves and sequences to determine the class label of particular image. Four type of object threshold have been observed deeply by the proposed algorithm to determine the structure of nuclei sequences such as Papillary class <0.30 with k-distances and finger like structure, Follicular class <0.40 with k-distances having random distances, Anaplast class <0.50 with k-distances with random vector spaces and values of <0.60 having k-distances shows un-known class in Dataset D.

Image input	TSA	NSA	NFST	Output of Preprocess
				No of Nuclei: 69 No of nucleolus: 40 No of overlaps: 08 Nuclei Grooves: 05 Structure type: A Class Label: Anaplast
				No of Nuclei: 125 No of nucleolus: 99 No of overlaps: 23 Nuclei Grooves: 07 Structure type: A Class Label: Anaplast
				No of Nuclei: 48 No of nucleolus: 35 No of overlaps: 35 Nuclei Grooves: 08 Structure type: P Class Label: Papillary
				No of Nuclei: 30 No of nucleolus: 17 No of overlaps: 07 Nuclei Grooves: 10 Structure type: P Class Label: Papillary
				No of Nuclei: 200 No of nucleolus: 190 No of overlaps: 30 Nuclei Grooves: 15 Structure type: A Class Label: Anaplast
				No of Nuclei: 70 No of nucleolus: 70 No of overlaps: 70 Nuclei Grooves: 08 Structure type: F Class Label: Follicular

Fig. 8. Pre-process results, Column 1 shows input image, column 2 represents TSA estimation, Column 3 demonstrates group of nuclei / nuclei grooves, column 4 shows distance matrix estimation developed to determine the structure of nuclei and column 5 is represents the summary of finding with class label attribute.

TABLE. I. CONFUSION MATRIX

	Papillary	Follicular	Anaplast	Benign
Papillary	158	20	25	16
Follicular	27	145	30	17
Anaplast	25	15	258	26
Benign	24	23	24	215

TABLE. II. CONFUSION MATRIX

	CANCEROUS NUCLEI = YES	CANCEROUS NUCLEI = NO
CANCEROUS NUCLEI = YES	217	70
CANCEROUS NUCLEI = NO	202	559

OVERALL ESTIMATED ACCURACY = 74%

TABLE. III. OVERALL PERFORMANCE OF PROPOSED METHODOLOGY

	Raw DICOM images	No of Extracted nuclei	No of classified Nuclei	No of miss-classified Nuclei	Precession	Recall
Papillary	20	219	158	61	67.52%	72.14%
Follicular	20	219	145	74	71.42%	66.21%
Anaplast	20	324	258	66	76.55%	79.63%
Benign	20	286	215	71	87.46%	75.15%

Overall estimated accuracy = 74%

TABLE. IV. COMPARISON OF OUR SYSTEM WITH LITERATURE

Approaches	Image Type	Cancer Type	Technique	Accuracy
1	Ultrasound Image	Follicular	SVM	97.50 %
2	FNAC Images	Follicular	SVM	91.00 %
3	FNAC Images	Medullary	Decision Tree	98.00%
4	FNAB Images	Papillary	SVM	93.30%
5	FNAC Images	Papillary	SVM	95.00%
Our Proposed approach	FNAB Images	All Thyroid Cancers	Random forest	74.00%

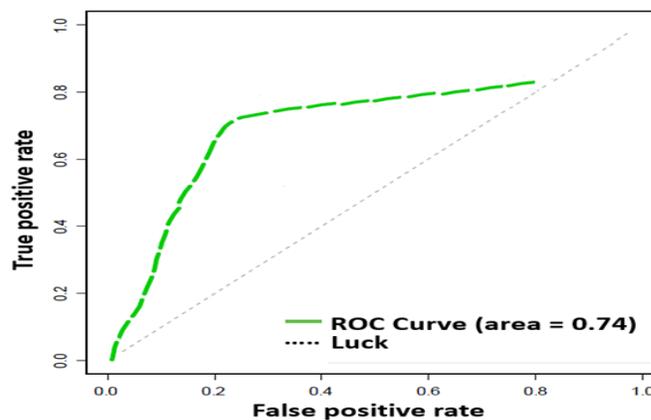


Fig. 9. Estimated ROC Curve for proposed system.

Confusion matrix (Table 1) show that 158 observations have been classified for papillary class label attribute and 145 instance have been classified for follicular class whereas 258 instance have been classified for anaplast class label attribute and 215 observation have been classified for benign class label attribute. The malignant and non-malignant observations have been presented in (Table 2) where 217 number of nuclei observation have been classified as cancerous cells yes and 559 instances of cancerous cells No have been classified by the classifier. The precision measure for each class such as papillary, follicular, anaplast and benign were recorded

respectively 67.52%, 71.42, 76.55% and 87.46% whereas recall measure was approximated, respectively 72.14%, 66.21%, 89.63% and 75.15%. ROC Curve have been shown in Fig. 9 for overall system performance and measured classification accuracy (Table 3) of our system is about 74% with 10-k fold cross validation. Table 4 is presented for comparison with literature which shows that our proposed approach classifies all the classes of well-differentiated thyroid cancers. Overall performance of the pre-process method have been described in Fig. 8 which show decision variable after examining the number of detected nuclei,

number of nucleolus, number of overlaps, number of nuclei grooves / sequences determined by the algorithm and type of sequence as P for papillary, F for follicular, A for anaplast and B for benign class label attribute.

## V. DISCUSSION AND CONCLUSION

Due to the heterogeneous and complex nature of micro-architectural components of histopathological DICOM (Digital Communication in Medicine) images, automated nuclei structure identification is one of the significant problems. Since follicular, medullary, papillary classification approaches are reported in literature and automated segmentation with classification of thyroid disease structure is yet not reported, thus it is direly needed to propose a system to detect the nuclei grooves by considering micro-architectural component analysis as a decision output or sub-class label attribute (such as well differentiated, poorly differentiated and benign cancers) followed by constructing a classification model for thyroid cancer variants. This paper proposes an automated computer based decision support system as second opinion to doctors which may enrich the assistance during the diagnostic process of cancer. Finally reproducibility of results would assume to be one of the convinced advantages to save the precious time and finance of the patients. This paper proposes a novel methodology for nuclei structure identification which selects the most significant DICOM (Digital Communication in Medicine) image behaviours from thyroid papillary, follicular carcinoma and anaplast cancers by using our algorithm *Auto\_Tissue\_Analysis* which is combination of our three proposed jobs (1) NSA (Nuclei Seed Approximation), (2) TSA (Tissue Structure Approximation) and NFST (Nuclei Feature Selection Tray). Over all methodology of our approach is comprising upon four layers, In first layer; noise reduction is by grey scale segmentation, edges of nuclei are detected, regions are transformed and nuclei seeds are extracted with respect to edges of every nuclei, in second layer unsupervised classification of tissue structures are performed by using the seed analysis and auto-feature selection is done. A decision variable is derived from the nuclei sequences which generates sub-class label attribute for every cancerous and non-cancerous class. In third layer we construct the decision model by using random forest (tree based) algorithm to extract the decision variable dependencies. Finally result visualization and performance evaluation is conducted by using confusion matrix, precision and recall measures respectively. The overall classification accuracy is measured about 74% with 10-k fold cross validation.

## VI. FUTURE WORK

There are several types of histopathological DICOM (Digital Communication in Medicine) images each have its own properties and behaviours. This research work addresses the classification problem of thyroid cancer. As a future work; the different real-world datasets will be trained and tested for various types of histopathological images of cancers such as classification of lung cancer, breast cancer, brain cancer, blood cancer, skin cancers and others. In skin cancer especially zero derma pigmentosa will be one of the important dimension because of heterogynous images may contain the same intensity of nuclei representation.

## REFERENCES

- [1] Pourahmad, S., Azad, M., Paydar, S., & Reza, H. Prediction of malignancy in suspected thyroid tumour patients by three different methods of classification in data mining. In First International Conference on advanced information technologies and applications: pp.1-8, 2012.
- [2] Xu, Y., Mo, T., Feng, Q., Zhong, P., Lai, M. and Chang, May. Deep learning of feature representation with multiple instance learning for medical image analysis. In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE, pp. 1626-1630, 2014.
- [3] Jothi, J.A.A. and Rajam, V.M.A. Effective segmentation and classification of thyroid histopathology images. Applied Soft Computing, 2016.
- [4] Iakovidis, D. K., Keramidas, E. G., & Maroulis, D. Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns. Journal of Artificial Intelligence in Medicine 50 (2010): pp.33-41, 2010.
- [5] Ding, J., Cheng, H., Ning, C., Huang, J., & Zhang, Y. (2011). Quantitative measurement for thyroid cancer characterization based on elastography. Journal of Ultrasound in Medicine, 30(9): pp.1259-1266, 2011.
- [6] Bell A.A, Kaftan, J.N, Schneider, T. E. Imaging and Image Processing for Early Cancer Diagnosis on Cytopathological Microscopy Images towards Fully Automatic AgNOR Detection, WrithArchin science journal, 2006.
- [7] Chuan, Y. Application of support-vector-machine-based method for feature selection and classification of thyroid nodules in ultrasound image. Journal of Pattern Recognition 43 (2010): pp.3494-3506, 2010.
- [8] Dimitris G., Panagiota S., Stavros T., Giannis K., Nikos D., George N., Dionisis C. Unsupervised segmentation of fine needle aspiration nuclei images of thyroid cancer using a support vector machine clustering methodology, 1<sup>st</sup> IC-SCCE Athens, pp. 8-10 September, 2004.
- [9] Suzuki, H., Saita, S., Kubo, M., Kawata, Y., Niki, N., Nishitani, H., & Moriyama, N. An automated distinction of DICOM images for lung cancer CAD system. In SPIE Medical Imaging (). International Society for Optics and Photonics, pp. 72640Z-72640Z, 2009.
- [10] Glotsos, D, Tsantis, S. Kybic J, Daskalakis, A. Pattern recognition based segmentation versus wavelet maxima chain edge representation for nuclei detectionin microscopy images of thyroid nodules. Euromedica medical center, Department of Medical Imaging, Athens, Greece 2013.
- [11] Gopinath, B, Shanthi, N. Support Vector Machine Based Diagnostic System for Thyroid Cancer using Statistical Texture Features, Asian Pacific J Cancer Prev, 14 (1), pp. 97-102, 2013.
- [12] Gopinath, B, Shanthi, N. Computer-aided diagnosis system for classifying benign and malignant thyroid nodules in multi-stained FNAB cytological images, Australas Phys EngSci Med pp.36:219-230, 2013.
- [13] Stavros T. Improving diagnostic accuracy in the classification of thyroid cancer by combining quantitative information extracted from both ultrasound and cytological images. 1<sup>st</sup> IC-SCCE-Athens, 8-10 September, 2004.
- [14] Smith, Russell B., et al. "Preoperative FDG-PET imaging to assess the malignant potential of follicular neoplasms of the thyroid." Otolaryngology-Head and Neck Surgery Vol 138(1), pp. 101-106, 2008
- [15] Gupta, Nidhi, et al. "Development and Validation of a Pediatric Endocrine Knowledge Assessment Questionnaire: Impact of ac Pediatric Endocrine Knowledge Assessment Questionnaire Intervention Study." Journal of clinical research in pediatric endocrinology Vol-8 (4), pp. 411-416, 2016.
- [16] Han J., Kamber M., and Pei P. "Data Mining: Concepts and Techniques". 3<sup>rd</sup> (ed.) the Morgan Kaufmann Series in Data Management Systems, 2011.
- [17] Rafeal, C. G., Richard, E. W. "Digital Image Processing", 3<sup>rd</sup> (ed). Prentice Hall. 2007.
- [18] Hussein, Mohamed, Amitabh Varshney, and Larry Davis. "On implementing graph cuts on cuda." First Workshop on General Purpose Processing on Graphics Processing Units. Vol. 2007. 2007.

- [19] Boykov, Y., & Funka-Lea, G. Graph cuts and efficient ND image segmentation. *International journal of computer vision*, 70(2), pp.109-131, 2006.
- [20] Malik, J., Belongie, S., Leung, T., & Shi, J. Contour and texture analysis for image segmentation. In *Perceptual Organization for artificial vision systems*. Springer US, pp. 139-172, 2000.
- [21] Al-Kofahi, Y., Lassoued, W., Lee, W., & Roysam, B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *Biomedical Engineering, IEEE Transactions on*, 57(4), pp. 841-852, 2010.
- [22] Liu, X., Huo, Z., & Zhang, J. Automated segmentation of breast lesions in ultrasound images. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the IEEE* pp. 7433-7435, 2006.
- [23] Al-Kofahi, Y., Lassoued, W., Lee, W., & Roysam, B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *Biomedical Engineering, IEEE Transactions on*, 57(4), pp.841-852, 2010.
- [24] Xu, N., Ahuja, N., & Bansal, R. Object segmentation using graph cuts based active contours. *Computer Vision and Image Understanding*, 107(3), pp.210-224, 2007.
- [25] Freedman, D., & Zhang, T. Interactive graph cut based segmentation with shape priors. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 755-762)*. IEEE, 2005.
- [26] Moukaddam, H., Pollak, J. and Haims, A., 2009. MRI characteristics and classification of peripheral vascular malformations and tumors. *Skeletal Radiology*, 6(38), pp..535-547, 2009.

# Mobility for an Optimal Data Collection in Wireless Sensor Networks

EZ-ZAIDI Asmaa  
Laboratory of Applied Mathematics and  
Computer Science (LAMAI)  
Cadi Ayyad University  
Marrakesh, Morocco

RAKRAK Said  
Laboratory of Applied Mathematics and  
Computer Science (LAMAI)  
Cadi Ayyad University  
Marrakesh, Morocco

**Abstract**—Sensor nodes located in the vicinity of a static sink drain rapidly their batteries since they have to carry more traffic burden. This situation results in network partition, holes as well as data losses. To mitigate this issue, many research proposed the use of mobile sink in data collection as a potential solution. However, due to its speed, the mobile sink has very short communication time to pick up all data from the sensor nodes within the network, therefore the sink is forced to return back to gather the remaining data. In this paper, we propose a new data collection scheme that aims to decrease the latency and enlarge the staying time between the mobile sink and the meeting points that buffer data originated from the other sensor nodes. We have also handled the case of urgent data so that they can be delivered without any delay. Our proposed scheme is validated via extensive simulations using NS2 simulator. Our approach significantly decreases the latency and prolongs the contact time between the mobile sink and sensor nodes.

**Keywords**—Contact time; mobile sink; wireless sensor networks; meeting point; data gathering

## I. INTRODUCTION

The growth of the micro-electro-mechanical systems (MEMS) technology as well as wireless communications have resulted to the development of low-cost, low-power, multifunctional sensor nodes which are characterized by their small size and that communicate untethered over short distances. A wireless sensor network (WSN) is composed of hundreds of sensor nodes that are distributed over a large area to monitor and track physical phenomenon like humidity, temperature, sound and so forth. WSNs have a wide range of applications such as traffic monitoring [1], patient healthcare monitoring [2], target tracking [3], indoor living monitoring [4], localization [5], and many other interesting applications.

In traditional wireless sensor networks, collected data is forwarded by sensor nodes to a static base station via multi-hop routing. Nodes located near the static sink bear more traffic burden, consequently they drain their energy faster than the other nodes causing the hot spot problem [6], [7] (Fig. 1). The death of these nodes causes network disconnections, holes and data losses since the sink will be no more in connection with the rest of the network.

Sink mobility has been introduced in several works as a potential solution to overcome the issues cited above. The sink motion improves the network's performances such as energy,

connectivity, reliability, security in data collection and many others benefits.

The remainder of this paper will be organized as follows: Section 2 presents some advantages and challenges of using sink mobility, Section 3 presents some related work, Section 4 describes the proposed scheme, Section 5 depicts the results of our scheme, conclusion and future works are presented in Section 6.

## II. ADVANTAGES AND CHALLENGES OF SINK MOBILITY IN DATA GATHERING

### A. Enhancement of Network Lifetime

Typical wireless sensor networks use multi-hop routing to transfer data toward the base station. Nonetheless, sensor nodes located in the sink's vicinity deplete their energy rapidly than the other nodes, leading to network degradation, disconnection as well as holes. Mobility was introduced to balance the energy consumption and minimize failures. The mobile sink moves in the network and pulls data buffered in sensor nodes that are within its communication range. By doing so, the multi-hop communication is reduced, and the traffic forwarding load is spread in the whole network. Several works have been conducted to decrease the network energy consumption by using sink mobility. For example, Shrivastava et al. [8] proposed a technique that consists on repositioning multiple mobile sinks towards regions with heavy traffic. Another work was introduced in [9], in which the authors examined how the mobility of sensor nodes operates in the wireless sensor network. They concluded that using mobile sensors results in better energy saving.

### B. Reliability

As the number of hops increases, the probability of transmission errors as well as data losses increases too. A reliable data transmission is ensured by the use of mobile sink, this is because sensor nodes located near the sink's trajectory upload their data directly through fewer hops.

### C. Security

Since the mobile sink keeps on changing its location around the sensor field, the chances of being attacked by external adversaries, and overhearing the collected information is decreased. In an attempt to avoid the injection of fake data, authors [10] proposed a random data collection scheme to protect the mobile sink from being tracked and

becoming the target of attacks. The random motion of the sink keeps its location private and hard to track or predict. The mobile sink moves randomly around the the network, and collects the sensed data stored in sensor nodes.

#### D. Coverage and Connectivity

Coverage has a great impact on WSN performances. It is one of the most important measurements in quality of service (Qos). Random deployment of sensor nodes, environmental disasters, presence of obstacles and power depletion of nodes all result in coverage holes. These holes may cause network dysfunction and interruption, which disturb the data collection process. Mobile sinks are used to overcome this problem; they visit disconnected regions where sensor nodes cannot operate to gather data. A tracking mechanism and a repairing robot algorithm were introduced in [11] to solve the coverage issue in the network. The moving robot carries a set of sensor nodes and performs hole-healing and repairs failure regions whenever it receives a request. In the same time, the mobile robot performs patrolling tasks to collect data.

Despite the various advantages that a mobile sink has nevertheless, many challenges may arise.

#### E. Data Dissemination

The communication overhead increases since the mobile sink has to broadcast each time its location within the network. Overhead causes energy wastage, that is why, it is important while designing data collection schemes to deal with this issue since sensor nodes have constrained resources.

#### F. Obstacles

The presence of obstacles inhibits the sink's movement in the network and degrades the communication quality between the sink and the sensor nodes.

#### G. Data Delivery Latency and Packet Losses

The data collection process depends on the contact time between the mobile sink and sensor nodes. When the sink moves at a high speed, it has very short time to pull data. Thus, sensor nodes have to wait for the sink to return, which prolong the delay of delivery. Moreover, an important packet losses and communication errors may occur due to the instability of the signal strength. Furthermore, since sensor nodes have limited buffer space, they may experience data losses, because they have to wait long time before the mobile sink comes in their vicinity again.

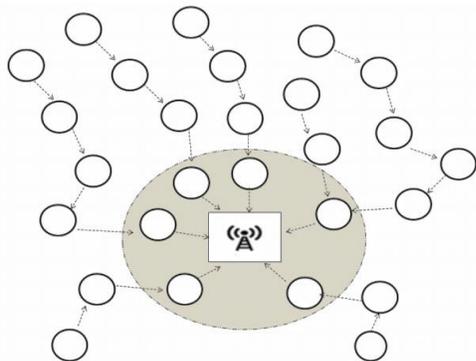


Fig. 1. The hotspot problem.

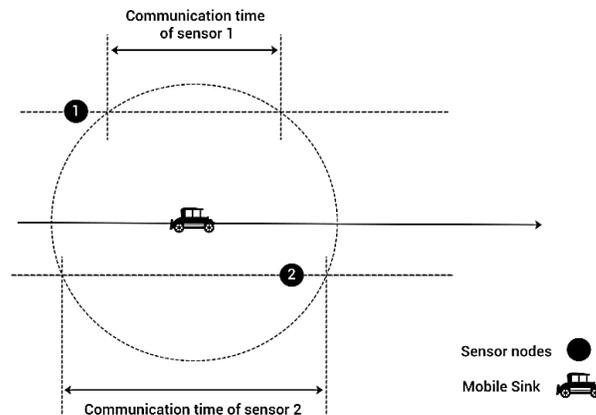


Fig. 2. Communication time between nodes and the sink.

#### H. Communication Time

Due to the high speed of the sink, the data collection process may be affected because of the short communication time between sensor nodes and the sink. As a consequence, only a small amount of data packets will be delivered to the mobile sink. So, nodes have to wait the return of the sink to complete the delivery of the rest of data. Fig. 2 shows the contact time between the mobile sink and the sensor nodes. Node 1 and 2 as instance can communicate with the sink from instant T1 to T2. However, node 2, which is located near the trajectory of the mobile sink, has relatively long contact time in comparison with node 1.

Aiming to decrease the latency while collecting a large amount of data, a data collection scheme is proposed in which a mobile sink travels the sensor network to pull data from Meeting points when it gets closer to them. These meeting points are special nodes that are elected according to some criteria, and have the ability to move when the mobile sink is within their range. Since we try to improve latency, we also took into consideration the case of urgent data that should be delivered in a fast manner.

Through this article, several questions will be answered:

- How to deliver maximum amount of data knowing beforehand that the communication time between the sensor nodes and the mobile sink is very limited? And in the same time avoid delays while collecting data.
- How to deliver urgent data effectively and without latency?

### III. RELATED WORK

This section exposes some related works having used the mobility of the sink for an efficient data collection.

The use of mobile sinks (MSs) has been investigated in much research. Authors in [12], [13] have classified the MS motion according to three different types of mobility patterns, which are: Predictable, Random and Controlled.

In random mobility pattern, the base station is attached to entities that move in a random manner such as animals or people moving in a certain area. In this type of pattern, the

mobile sink makes random movements in terms of direction and velocity. The controlled mobility refers to mobile devices that are guided in the network to accomplish specific tasks such as healing coverage, decreasing the energy consumption or repairing network failures. In the last type of mobility i.e. predictable that is used in our approach, the sink's movement can be predicted. In general, in this mobility pattern, the sink can be mounted on a public transportation such as busses, trains or vehicles, and follow a predetermined fixed path. By using this type of mobility, sensor nodes can anticipate the time of the sink's visit, and thus they can switch to sleep mode to save energy until the sink comes in their range.

In [14] authors studied the tradeoff between the energy conservation and data collection latency in wireless sensor networks. The approach's main concept consists in finding a set of special sensor nodes called polling points (PP) in such a way that all sensor nodes can send their data through a certain number of relay hops. The approach has demonstrated its effectiveness in shortening the tour length of the mobile collector.

Xu et al. [15] used a mobile sink travelling along a fixed trajectory to gather data without stopping. The sink, moving at a constant speed, collects data from gateway nodes that are located nearby its trajectory. These gateway nodes are relay nodes for the other sensors. Since the communication time with the gateway nodes is limited, the mobile sink cannot upload all the data sensed by nodes. To overcome this issue, data originated from only a subset of sensor nodes referred as packet nodes are collected by the sink and used to estimate those of the others. The maximum packet nodes are allocated to gateways having a large intersection time with the mobile sink.

A proactive data reporting protocol called SinkTrail was proposed by Liu et al. [16] in which the mobile sink moves in the field with a low speed, sojourns at some positions for a short time to collect data, and then moves to another location. Each position visited by the sink is viewed as a footprint. These footprints represent the logical coordinate of the sink in the network; they are used to guide the sensor nodes to report their data without knowing the physical locations as well as speed of the mobile sink. However, in a large scale network, huge delay may be induced especially for sensor located far away from the mobile sink.

A mobile sink based routing protocol "MSRP" was introduced in [17] in which the mobile sink visits locations of ClusterHeads having high residual energy to gather data. However, the approach may experience significant delays since the mobile sink favors zones of network that are rich in terms of resources, and overlook the other regions.

Authors in [18] proposed a Detour-Aware Mobile Sink Tracking (DAMST) for collecting data in a low overhead and in an energy efficient way. The mobile sink, while crossing a region, nominates a specific sensor node as a region agent called RA, which is in charge of gathering data around. In the same time, the sink records the location of nodes near its trajectory as footprints. When the sink needs information from

the region agent, it constructs an energy efficient path from itself to that region; the path is established by analyzing the sink's movement angles and eliminating the footprints on detours by comparing adjacent movement angles.

Different from the approaches cited above, authors in [19] used a flying mobile sink that hover above a set of terrestrial sensors to collect data. The mobile sink is carried by a quadcopter because it is more flexible to move, descent and lift. Authors studied the speed, time of sojourning, flying trajectory, height of the sink as well as the amount of data to be transmitted to it.

Wang et al. [20] used several mobile sinks that move along the periphery of the network with a constant speed and sojourn at some particular parking positions for a certain time to gather data sensed from the sensor nodes. The next sojourn locations  $P_1 \dots P_n$  are picked from a PP\_Table that stores all the parking positions. When the mobile sink is on its way to a parking position, it broadcasts a notification message to inform about its coming. Upon the movement of the mobile sink from a parking position  $P_i$  to  $P_j$  it will not collect data.

Authors in [21] used three mobile sinks to mitigate the problem of hot spots during data collection. The network is divided into two parts, an inner concentric circular region of radius  $r$  called area A and an outer circular region called region B that is divided into 8 portions  $B_1, B_2 \dots B_8$ . One of the mobile sinks will span along the diameter of the circle, while the two others will travel along the arc line. These mobile sinks will move back and forth along their determined trajectories and sojourn at some fixed points to pick data.

In an attempt to collect the maximum amount of data, authors in [22] proposed a biased sink mobility scheme with adaptive pauses time for efficient data collection. When moving, the sink visits each region to gather data and adaptively stops for a time interval, which is proportional to the local data traffic. The introduction of pauses time increases the delivery success rate; however in large scale networks this scheme will incur important latencies while delivering data because of the large pauses made by the mobile sink. Furthermore, sensor nodes located in far regions will have to wait for a long time before the mobile sink comes again into their vicinity. This situation may cause buffer overflows since these nodes may not be able to hold all the generated data during this lapse of time.

Clearly by decreasing the mobile sink's velocity or by introducing the concept of stop times or favoring regions with high traffic density, the total amount of data collected will be improved, however the problem of latency will occur especially in large scale networks. In our approach, we do not need to lower the speed of the mobile sink or introduce pauses to collect large amount of data, or even visit regions having high density. We exploit the mobility of the sink as well as the meeting points so as to provide an efficient data collection without latency. We also take into account the priority of data when a set of nodes detects some alarming events. Our algorithm aims at optimizing the total amount of data collected and in the same time lowering the latency.

#### IV. PROPOSED SOLUTION

##### 1. System model

Our sensor network as shown in Fig. 3 is composed of a set of sensor nodes deployed in a grid and a mobile sink that moves along a straight line.

- A set of sensor nodes denoted as SN deployed within the region of interest. These nodes generate data and relay it towards special nodes called meeting points according to multi-hop communication.
- Nodes located near the trajectory of the mobile sink (black dots) are called meeting points (MPs). In addition to sensing tasks, they are also responsible of delivering data at the mobile sink when this latter comes within range. MP nodes are used to enable the sink to collect a large amount of data without having to visit the whole network; this will minimize the delay caused by the long-distance that the sink travels.
- A mobile sink referred as MS, moves back and forth along a straight line at a constant speed. Its motion follows a predictable mobility model [23], which is used in various applications such as public transportation: buses and trains. We also assume that there are no obstacles that hinder the sink's movement.

The mobile sink is aware of its speed and trajectory, and all the sensor nodes within the network have the same communication range, and aware of their positions. Each sensor node generates the same amount of data.

The mobile sink has unlimited energy, memory and strong computational capacity. So, all the computations will be performed by the mobile sink.

- Proposed approach

The proposed data collection process is divided into several phases described below:

- First trip of the mobile sink

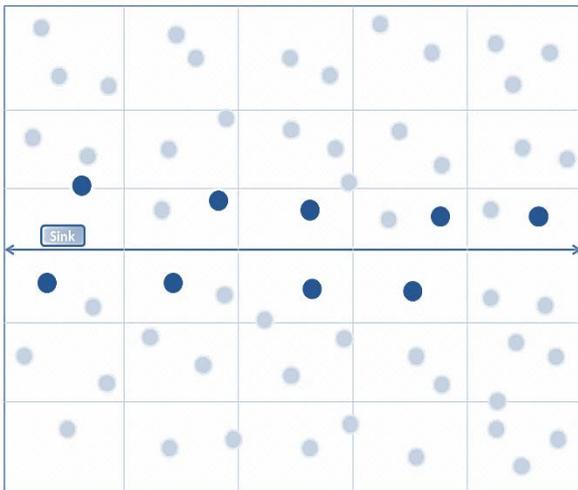


Fig. 3. Diagram of the wireless sensor network.

In its first trip, the mobile sink moves along the network and broadcasts periodically beacon messages to inform about its presence. Sensor nodes having received the broadcast send a message to the sink to express their candidacy to be meeting points. During this round, the mobile sink also records its communication duration with each sensor node.

- Second trip (Election of meeting points)

In this trip, the mobile sink travels the trajectory for a second time to broadcast results about the elected meeting points. The computation of the most suitable nodes to be nominated as MPs is done by the mobile sink since it has very strong power computations. In fact, the meeting points guarantee the non-disconnection of the network since they act like a bridge between the mobile sink and the other sensors within the network. They buffer incoming data until the sink passes by to collect it.

The mobile sink elects the most convenient nodes to be meeting points; they should have a high energy level, and located near the sink's trajectory to deliver data reliably and within one hop, and finally they must have a large intersection time with the mobile sink so as to transfer important amount of data as shown in Fig. 4.

##### // executed code in sensor nodes

```
Initialize TfirstBeacon = TlastBeacon = 0, Eresid=0, Distances [],  
shortestDistance = 0;  
// Waiting until a beacon message is received  
If (beacon is received) {  
    d = estimated distance between MS and CurrentNode;  
    Distances[i] = d;  
    Eresid = CurrentNode.getResidualEnergy();  
    If ( is first beacon ) {  
        // record the time when the first beacon has been received  
        TlastBeacon = TfirstBeacon = beacon.timestamp ;  
        shortestDistance = d;  
    }  
    else {  
        TlastBeacon = beacon.Timestamp;  
        If (shortestDistance > d) {  
            shortestDistance = d;  
        }  
    }  
    SendCandidacy (TfirstBeacon, TlastBeacon, ΣDistances [i], Eresid);  
}
```

##### // executed code in Mobile sink

```
While (MS is still in the first trip) {  
    If ( a candidacy message is received ) {  
        nodeId = extractNodeIdFromCandidacyMessage();  
        If (MS has already received a candidacy msg from nodeId )  
        {  
            updateCandidacyInformationOfNode (nodeID);  
        }  
        else {  
            storeCandidacyInformationOfNode (nodeID);  
        }  
    }  
    For all candidate nodes  
    ΔT = TlastBeacon - TfirstBeacon ;  
    Score = x * (1/(ΣDistances [i] / Distances.Length)) + y * Eresid + z *  
    ΔT;
```

Fig. 4. Depicts the algorithm of meeting points' election.

The parameters of choosing the meeting points are based on [24]. After the reception of candidacies, the mobile sink calculates the cost value of each candidate node according to the algorithm below and sorts it in an increasing order. Elected nodes are those who have the highest cost value. Finally, the sink informs each chosen node through a message about its election. Once elected, each meeting point builds the shortest path rooted from them. That way, each sensor node can relay its collected data towards the meeting point to which it is attached.

X, Y and Z are coefficients of the distance, energy level and time of communication, respectively.

- Data collection

After the meeting points' election, the data collection process begins. The meeting points are used because they enable the mobile sink to pull data without having to visit the entire network. They serve as data collection points for many sensor nodes. Furthermore they play a major role in decreasing the data collection delay.

Each sensor node forwards its sensed data, using multi-hop communication to its attached meeting point, this latter stores data until the arrival of the mobile sink to fetch it. Since the mobile sink has a very short time to communicate with the meeting points due to its speed, so only few amounts of data will be collected. Besides, when the size of data is important that it cannot be entirely delivered during the contact time, the MPs have to wait for the sink to return again which induces large latencies. In an attempt to find a compromise between collecting a large volume of data, and in the same time avoiding long delays in data gathering, our proposal consists on moving the meeting points for a definite distance with the mobile sink to remain the longest possible in contact so as to deliver maximum quantity of data.

Upon its movement back and forth along the path, the mobile sink sends periodically a Notification\_Message that contains information about its speed, position as well as its direction (forth or back) to announce its presence. As the mobile sink enters the communication range of the meeting point, the data collection process begins (Fig. 5(a)).

When the signal strength (shortest Distance) between the mobile sink and the meeting point reaches its peak, the meeting point starts moving for a certain distance parallel to the sink (Fig. 5(b)) with the aim to prolong the staying time between each other.

The distance travelled by the meeting point is equal to  $2/3 * d$  where d is the distance between two consecutive sensor nodes within the grid. In fact, the choice of such distance is to avoid collisions and the overlap that may occur between the meeting points that try to deliver data simultaneously to the mobile sink.

After crossing the specified distance, the meeting point stops and wait for the sink to come back. When the latter returns, the MP will move again another  $2/3 * d$  to return back to its initial position (start position). By doing so, we ensure that more packets are delivered to the mobile sink without having to make stops.

- Priority of data

We considered two types of data, the normal ones (temperature) and high-priority ones (detection of a critical event) that must be sent within a shorter delay. All along the data collection, each meeting point filters the received data packets according to a flag to distinguish between sensitive and normal ones.

When the mobile sink enters within range, the urgent data are first delivered then the normal ones. In case, the size of the priority data is big enough and cannot be gathered completely by the sink within the contact time, the meeting point forwards the rest of it toward the future location of the mobile sink to avoid any delays.

To do so, the current MP (the meeting point that possesses the priority data) looks for a near meeting point according to the sink's direction movement (forward or backward). If the meeting point is located within its communication range, the rest of priority data is sent directly to it, otherwise multi-hop communication via relay nodes is used to relay the remaining data as depicted in Fig. 6. By doing so, it is guaranteed that data that has priority reaches the sink in a timely manner. Fig. 7 depicts the algorithm of the data collection process.

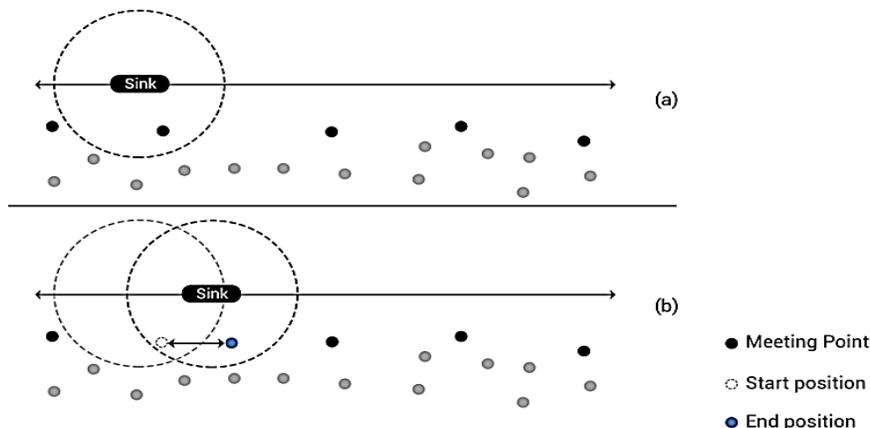


Fig. 5. Meeting point movement.

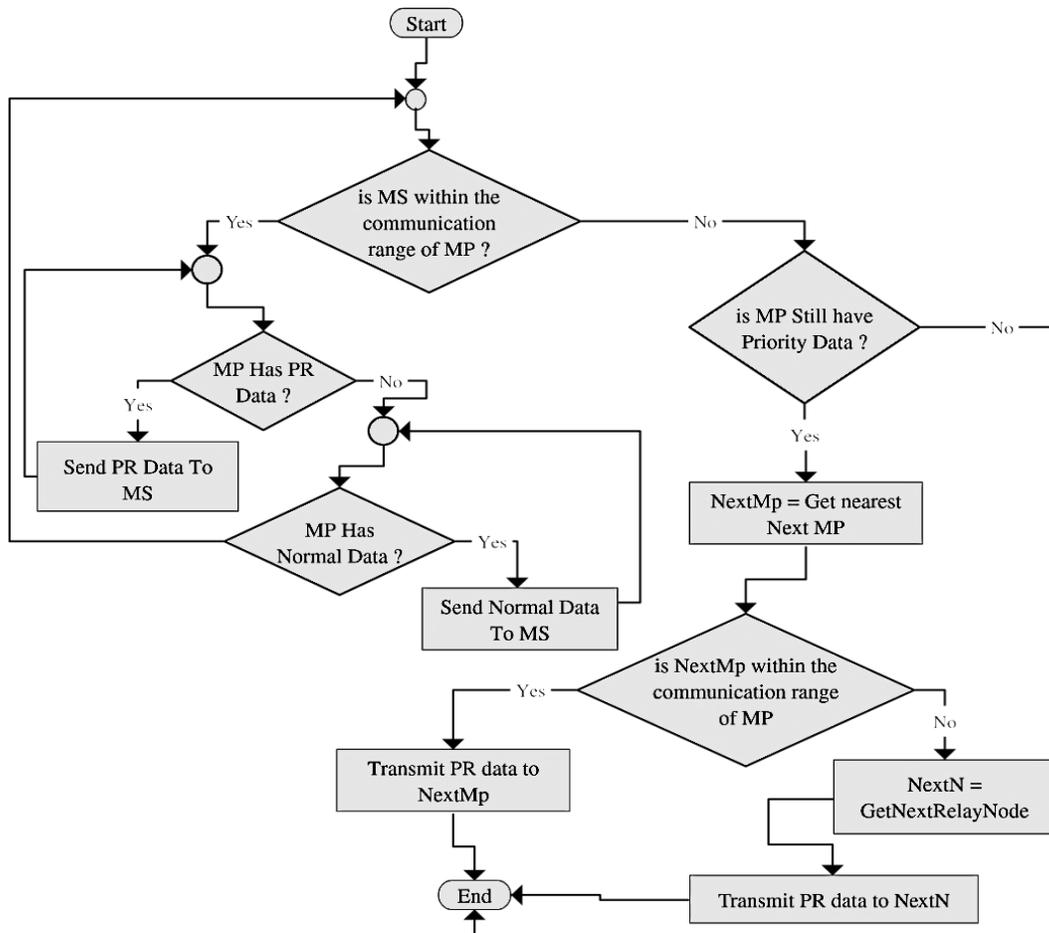


Fig. 6. Priority of data diagram.

```

Initialize V_MP, V_MS, Pr, isMoving = false;
// the variables are respectively: velocity of MP, MS,
priority and a Boolean to check whether the MP is moving
or not

While (First NotificationMsg has not been received yet) {
    MP waits in its actual position;
    MP collects data from sensor nodes;
    MP performs filter based on data priority flag;
}
// First Notification message is received
V_MP = V_MS // set the velocity of the MPs
While (MS is still within range of MP) {
    D = estimatedDistanceBetweenMP_And_MS ();
    If (D <= shortestDistance - μ) {
        If (isMoving) {
            Start moving 2/3*d;
            isMoving = true;
        }
    }
    If (Pr) // priority data exists.
        Transfer Priority data ();
    Else
        Transfer non-priority data ();
}
If (Still Pr data) { // MS is out of range (timeout expired)
    Send remain priority data to the Next MP (cf. figure 6)
}
isMoving = false;
    
```

Fig. 7. Data collection algorithm.

## V. PERFORMANCE EVALUATION

To further evaluate the efficiency of the proposed approach, we have used NS-2.35 which is a discrete event simulator.

### A. Simulation setup

We have considered a 200 \* 200 network area; the number of sensor nodes is 300. All the sensor nodes have the same communication range  $R=15$  m. Each sensor node generates at random times about 1 message per 10 seconds and sends it to its parent. The simulation time is 2500 seconds. The approach is evaluated by considering different speeds of the mobile sink which are 4, 10 and 20 m/s.

*Number of mobile sinks: 1 mobile sink.*

*Number of nodes: 300.*

*Simulation area: 200\*200 m*

*Speed of the mobile sink: 4, 10 and 20 m/s.*

*Communication range: 15m*

*Simulation runtime: 2500 s*

*Mac protocol 802.15.4*

We have considered three parameters for evaluation which are data latency, packet delivery ratio and contact time duration.

- Packet delivery ratio represents the ratio of packets which are successfully delivered to the sink compared to the total packets that have been generated.
- Data latency: the time elapsed between the creation of the message and its delivery to the sink.
- Contact time duration: The communication time between the mobile sink and the meeting points.

The simulation was run several times; we took the average of the runs to show the results.

### B. Findings

1) *Duration of the communication:* Fig. 8 shows how the mobility of the meeting points increases the contact time duration while gathering data.

As discussed above, the limited communication time between sensor nodes and the mobile sink impacts the data collection process since only small amount of data will be gathered.

We notice from the graph that when the meeting points are static the contact duration with the sink is relatively very short, therefore the MPs upload only few amount of data packets and have to wait until the sink returns back to upload the rest of data. However, when the meeting points are mobile the communication time is enlarged, so in one hand more data is collected and in the other hand the latency will be minimized as shown in the next graph (Fig. 9).

2) *Latency:* Fig. 9 shows the latency of the network for different speeds of the mobile sink. As expected, when moving the meeting points, our scheme achieves a much better performance in terms of latency in all cases of sink's velocity {4, 10 and 20m/s}.

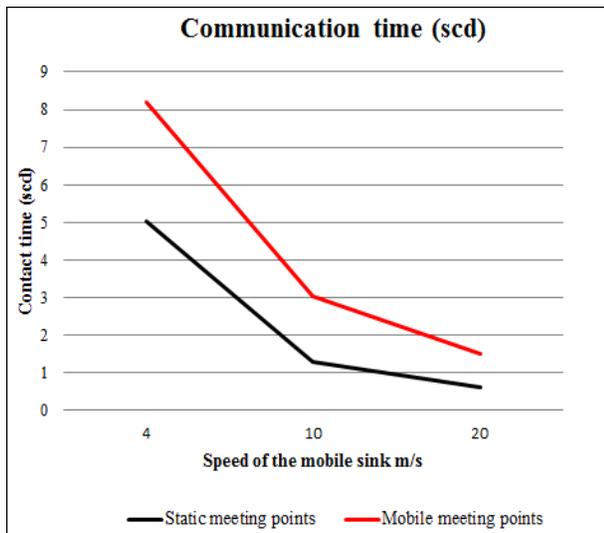


Fig. 8. Communication time.

Conversely, the latency is very high in the case of static meeting points. This goes back to the fact that the communication time between the mobile sink and the meeting points is very short, thereby data is not fully collected. So, the sink is forced to make several trips back and forth to gather the rest of packets. This situation obviously increases the latency.

It is also noticed that latency is slightly high when the sink's velocity is 4m/s or 20m/s. With a speed of 4m/s, the MS moves slowly and takes a long time to collect data within the network. With a speed of 20m/s, the sink moves at a high speed which reduces the contact time. The sink is forced to return repeatedly to collect data buffered in the meeting points. Conversely, the mobile sink takes less time to visit all the meeting points when it moves with 10m/s. The speed of 10m/s balances between latency and quantity of collected data.

3) *Packet delivery ratio (%):* In Fig. 10 the number of collected data using static meeting points is almost equal to the case of mobile meeting points.

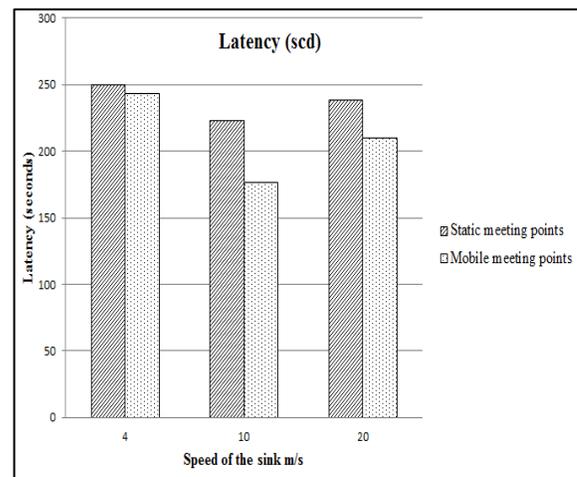


Fig. 9. Latency.

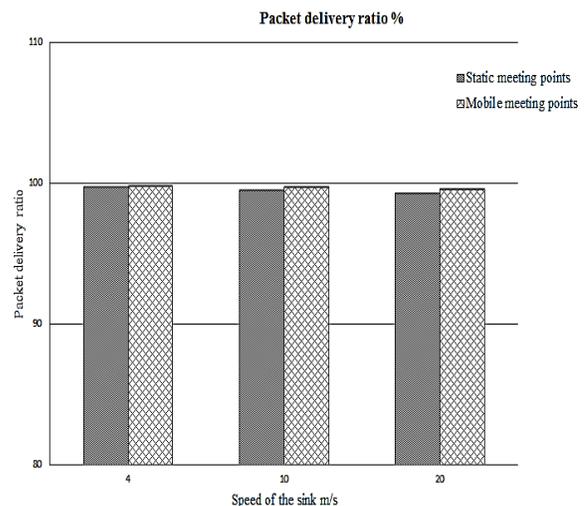


Fig. 10. Packet delivery ratio.

We used the mobility of meeting points to ensure a maximum contact time with the sink which increase the quantity of collected data. The plus in our approach is that we have achieved a balance between how to have a high packet delivery ratio while reducing the latency. From the figure below, the number of collected packets is high regardless of the sink's speed. When the mobile sink spans the network at a speed of 20m/s, the success rate is slightly low in comparison with the speed of 4m/s and 10m/s. The high speed of the sink causes packet drops which subsequently impact the ratio of delivered packets.

## VI. CONCLUSION AND PERSPECTIVES

In this paper, an efficient data collection using the mobility of both the mobile sink and the meeting points was proposed. Our approach aims to collect large amount of data in a very short time without having to travel the whole network or make pauses time. Urgent data are also handled while collecting data, they are delivered without any delay.

Simulation results have shown that we achieve a high success rate without delays. Our future work will focus on the energy aspect; the frequent mobility of meeting points consumes much energy, as well as the retransmission of packets.

### REFERENCES

- [1] Bottero, M., Dalla Chiara, B., Deflorio, F. P. (2013). Wireless sensor networks for traffic monitoring in a logistic centre. *Transportation Research Part C: Emerging Technologies*, 26, 99-124.
- [2] Manvi, S. S., Manjula, R. B. (2014). A Survey on Health Care Services Using Wireless Sensor Networks. *Healthcare Administration: Concepts, Methodologies*, 132.
- [3] Ez-Zaidi, A., & Rakrak, S. (2015). A Comparative Study of Target Tracking Approaches in Wireless Sensor Networks. *Journal of Sensors*, 2016.
- [4] M. Bhuiyan and G. Wang, "Deploying wireless sensor networks with fault-tolerance for structural health monitoring," *Computers, IEEE Trans.*, 2015.
- [5] Halder, S., & Ghosal, A. (2015). A survey on mobile anchor assisted localization techniques in wireless sensor networks. *Wireless Networks*, 1-20.
- [6] Kuo, C. H., Chen, T. S., & Lo, Y. H. (2015). Efficient traffic load reduction algorithms for mitigating query hotspots for wireless sensor networks. *International Journal of Ad Hoc and Ubiquitous Computing*, 18(3), 153-163.
- [7] Jannu, S., & Jana, P. K. (2016). Energy Efficient Algorithms for Hot Spot Problem in Wireless Sensor Networks. In *Proceedings of the Second International Conference on Computer and Communication Technologies* (pp. 509-517). Springer India.
- [8] Shrivastava, P., Pokle, S.B.: A multiple sink repositioning technique to improve the energy efficiency of wireless sensor networks. 2013 15th International Conference on Advanced Computing Technologies (ICACT). pp. 1-4. IEEE (2013).
- [9] Yang, T., Mino, G., Barolli, L., Durresi, A., & Xhafa, F. (2011, June). Energy-saving in wireless sensor networks considering mobile sensor nodes. In *Complex, Intelligent and Software Intensive Systems (CISIS)*, 2011 International Conference on (pp. 249-256). IEEE.
- [10] Ngai, E., Rodhe, I.: On providing location privacy for mobile sinks in wireless sensor networks. *Wirel. networks*. (2013).
- [11] Chang, C.-Y., Lin, C.-Y., Yu, G.-J., Kuo, C.-H.: An energy-efficient hole-healing mechanism for wireless sensor networks with obstacles. *Wirel. Commun. Mob. Comput.* 13, 377-392 (2013).
- [12] Tunca, C., Isik, S.: Distributed mobile sink routing for wireless sensor networks: a survey. *Communications Surveys & Tutorials*, IEEE. (2014).
- [13] Khan, M., Gansterer, W., Haring, G.: Static vs. mobile sink: The influence of basic parameters on energy efficiency in wireless sensor networks. *Comput. Commun.* (2013).
- [14] Zhao, M., and Yang, Y. (2012). Bounded relay hop mobile data gathering in wireless sensor networks. *Computers, IEEE Transactions on*, 61(2), 265-277.
- [15] Xu, X., Liang, W., and Wark, T. (2011, June). Data quality maximization in sensor networks with a mobile sink. In *Distributed Computing in Sensor Systems and Workshops (DCOSS)*, 2011 International Conference on (pp. 1-8). IEEE.
- [16] Liu, X., Zhao, H., Yang, X., and Li, X. (2013). SinkTrail: a proactive data reporting protocol for wireless sensor networks. *Computers, IEEE Transactions on*, 62(1), 151-162.
- [17] Nazir, B., and Hasbullah, H. (2010, December). Mobile sink based routing protocol (MSRP) for prolonging network lifetime in clustered wireless sensor network. In *Computer Applications and Industrial Electronics (ICCAIE)*, 2010 International Conference on (pp. 624-629). IEEE.
- [18] Yang, G., Xu, H., He, X., Wang, G., Xiong, N., & Wu, C. (2016). Tracking Mobile Sinks via Analysis of Movement Angle Changes in WSNs. *Sensors*, 16(4), 449.
- [19] Chen, Y., Chen, J., Zhou, L., Du, Y.: A data gathering approach for wireless sensor network with quadrotor-based mobile sink node. *Adv. Wirel. Sens.Networks*. (2012).
- [20] Wang, J., Li, B., Xia, F., Kim, C. S., & Kim, J. U. (2014). An energy efficient distance-aware routing algorithm with multiple mobile sinks for wireless sensor networks. *Sensors*, 14(8), 15163-15181.
- [21] Wang, J., Zuo, L., Shen, J., Li, B., and Lee, S. (2015). Multiple mobile sink based routing algorithm for data dissemination in wireless sensor networks. *Concurrency and Computation: Practice and Experience*, 27(10), 2656-2667.
- [22] Kinalis, A., Nikolettseas, S., Patroumpa, D., and Rolim, J. (2014). Biased sink mobility with adaptive stop times for low latency data collection in sensor networks. *Information fusion*, 15, 56-63.
- [23] Dong, Q., Dargie, W.: A Survey on Mobility and Mobility-Aware MAC Protocols in Wireless Sensor Networks. *IEEE Commun. Surv. Tutorials*. 15, 88-100 (2013).
- [24] Konstantopoulos, C., Pantziou, G., Gavalas, D., Mpitziopoulos, A., Mamalis, B.: A Rendezvous-Based Approach Enabling Energy-Efficient Sensory Data Collection with Mobile Sinks. *IEEE Trans. Parallel Distrib. Syst.* 23, 809-817 (2012).

# Estimating True Demand in Airline's Revenue Management Systems using Observed Sales

Alireza Nikseresht  
Computer Science and Engineering  
Shiraz University  
Shiraz, Iran

Koorush Ziarati  
Computer Science and Engineering  
Shiraz University  
Shiraz, Iran

**Abstract**—Forecasting accuracy is very important in revenue management. Improved forecast accuracy, improves the decision made about inventory and this lead to a greater revenue. In the airline's revenue management systems, the inventory is controlled by changing the product availability. As a consequence of changing availability, the recorded sales become a censored observation of underlying demand, so could not depict the true demand, and the accuracy of forecasting is affected by this censored data. This paper proposed a method to estimate true demand from censored data. In the literature, this process is referred to as unconstraining or uncensoring. Multinomial Logit model is used to model the customer choice behaviour. A simple algorithm is proposed to estimate the parameters (customers' preference) of the model by using historical sales data, product availability info and the market share. The proposed method is evaluated using different simulated datasets and the results are compared with three benchmark models that are used commonly in airline revenue management practice. The experiments show that proposed method outperforms the others in terms of execution time and accuracy. A 47.64% improvement is reported in root mean square error between simulated and estimated demand in contrast to the benchmark models.

**Keywords**—Demand estimation; demand modelling; forecasting; revenue management; inventory control; unconstraining; uncensoring

## I. INTRODUCTION

After airline deregulation in 1978, many low cost airlines were born. They started to offer much cheaper tickets than the major airlines, nevertheless they had benefit, because they had lower operation cost [1]. The major airlines were not able to price their seats below or at least near to the new born airlines. Although they did not lose all of their customers, because some customers are price sensitive while the others are sensitive to quality of service and brands, they loosed a significant amount of revenue. In result, they offered their seats with different prices and features to capture all types of customers, and the problem changed to how to optimally price the seats and controlling seat inventory? To find the optimal solution for this problem, airlines should know the exact amount of demands for each offered seat. Knowing the exact amount of demand is impossible because it is something that will happen in future, so by looking at historical sales data, they try to estimate demand.

In the airline's revenue management systems, the inventory is controlled by means of setting booking limits for each

product. The goal of setting these booking limits is to protect a specific amount of seat for higher profit customers. Although the booking limits cause more revenue, it censors the true demand, which is needed for accurate demand forecasting. In most cases the observed sales data do not reflect the true demand because no sales are recorded after a product has been sold out, or in other word if a product's booking limit has been reached. Therefore, transactional sales data are usually denoted as censored demand [2]. Using the censored data, as true demand has two consequences: 1) under estimation of true demand, which causes a spiral down effect over time; 2) over estimation of true demands [3]–[5]. In both situations the companies lose revenue. Although ignoring the censorship results in significant reductions in revenue, observing demand after its booking limit reached is impractical. So, unconstraining methods are used to estimate the true demand [6]. In addition to censorship, ignoring the correlation of demand between related products leads to inefficient estimates of the true demand [7], [8]. Empirical studies show that between 45% and 84% of demand can be substituted [9]–[12]. Nevertheless, measuring true demands using the available sales data is not an easy task [13]. If the accuracy of forecasting grows about 20%, the resulting revenue will improve about 1% [14]. In a research conducted by Weatherford, it was found that if there is a negative bias in forecast, up to 3% of the potential profit may be lost [15]. So it is rational to look forward to methods measuring the true demands from the available data and use these data for forecasting and managing seats [16].

This paper proposed an unconstraining algorithm to estimate true demand using censored sales transaction data in a fast, accurate, and very simple manner. The proposed algorithm is an extension of MSEG algorithm which is an iterative demand estimation algorithm for airline sales transactions data [17]. MSEG is actually a minimum square error algorithm which runs iteratively and in each iteration tries to reduce the amount of error between estimated and observed sales data, until a minimum error is reached or a certain iteration is elapsed. Gradient descent is used to minimize the error function, and in each iteration, to approach the minimum, it moves toward the opposite of the gradient vector with a small step size  $L$ .

Unlike MSEG, in the proposed method the objective function is changed to convex form and the optimal solution is founded easily by a simple derivative. As MSEG, our method needs transactional sales data and product availability, plus

market share which were optional in MSEG. In the proposed method, customer arrival is estimated using a simple heuristic instead of assuming a priori distribution, so the risk of misspecification of distribution is resolved and it benefit the advantages of not requiring any assumption on the form of the distribution function. To model the demand, the multinomial logit choice model is employed and the proposed algorithm estimates the choice model parameters. The proposed method is evaluated using two types of datasets: 1) dataset with full observation of demand occurrences; and 2) dataset with censored observation of demands. The dataset 1 which contains unconstrained observation is served as a benchmark to evaluate our work. The results were compared with the other unconstraining methods which are applied commonly in revenue management practice, such as projection de truncation (PD), expectation maximization (EM) and another version of EM proposed by Vulcano et al. in 2012 [18]. The rest of the paper is structured as follows: In Section 2, similar works have been reviewed and Section 3 discusses the problem description. Section 4 describes how the demand is modelled. Section 5 is dedicated to the solution and algorithm. In Section 6, the simulation process and the datasets are described. The results are demonstrated in Section 7, and finally, Section 8 concludes the paper.

## II. LITERATURE REVIEW

Since 1990 in the context of demand unconstraining, there is a vast literature that address the issue of estimating true demand using historical sales data [19]. Choice modelling attempts to model the decision process of customers. The theory of choice modelling assume customers are rational agents who intelligently make decisions when, what, and how much to purchase to achieve the maximum benefit. An important aspect of this rationality assumption is that customer behaviour can be predicted [1]. Multinomial logit model (MNL) is the most popular approaches to choice modelling. The MNL is a discrete customer choice model which presumes customers are rational utility maximizers, and predicts the customer's behaviour [20].

Andersson, Algiers, and Beser, presented a customer choice based method for optimal seat allocation and the parameters of choice model are captured from interview with experts and historical sales data. In the proposed model they considered recapture and buy-up [21], [22]. Ratliff devised a customer choice model-based heuristic to estimate demand, spill and recapture. His heuristic needs sales data and market share to do unconstraining. It is capable of computing recapture alongside different flights and products, and its main advantage is that by taking into account the portion of substitution between products or flights, due to the fact that some products are not available, it prevents double counting of demands [23].

Many researchers in their works, employed expectation maximization (EM) to find the customer choice model parameters. Talluri and Van Ryzin applied a discrete choice model to model customer behaviour and found the related parameters using EM method. In their proposed method, they assumed a distribution function for demand and then tried to find its parameters from the observed data, indicating that their methods perform well [1]. In other work, they developed an

estimation method based on the expectation maximization which is able to jointly estimate arrival rates and choice model parameters when no-purchase outcomes are unobservable [24]. Haensel and Koole utilized the idea of customer choice set to model the buying behaviour of customers. They applied EM algorithm to unconstrain the censored data, not assuming a fixed arrival rate for customers as regularly done. Instead, they estimated a demand function for each group of customers by analysing observed data [2], [25]. Newman et al. presented a parameter estimation method for multinomial logit model in which one alternative is never observed. Their method is based on decomposing the log-likelihood function into marginal and conditional components. They showed that their proposed method is computationally efficient and provides consistent parameter estimates. Simulations based on industry data set demonstrate their method computationally outperform the other alternative estimation methods [26]. Vulcano et al. developed a maximum likelihood estimation algorithm that uses a variation of the EM method to account for unobservable data. With simulation study, they showed that revenue improves about 1%–5% using choice based revenue management [27]. Vulcano et al. proposed an EM-based method estimating spilled and substitute demands. Their method only needs the observed sales data, product availability, and company market share. Their main idea is to consider the problem as primary demand or customer first choice. They supposed that each customer has a set of choices with a primary or first choice. Each customer could buy his/her first choice, and if the first choice is not available, the next choice is substituted or leaves the system without purchasing anything. Then they tried to estimate primary demands, substitute demands, and no purchase count, using EM [18]. Agrawal et al. modeled the consumer choice behavior using the multinomial logit model for assortment selection problem and dynamically estimated the model parameters [28].

Modelling customer choice is a high dimensional problem and it is difficult to dealing with it, so to deal with its difficulty, many researchers prefer to assume a priori distribution and parametric model which they think is able to adequately capture choice model behaviour [29]. The side effects of all parametric approaches are misspecification of the model and overestimation or underestimation true demand. Van Ryzin and Vulcano proposed a nonparametric choice model approach to estimate, customer preferences for a set of substitutable products. With the numerical experiments on a real dataset, they showed that their method perform well and improves root mean square error between predicted and observed sales about 67% [19]. Farias et al. presented an approach to predict the expected sales from historical data. They used a nonparametric approach to model the customer choice. With empirical study using simulated and real dataset, they showed that their method is able to produce accurate revenue forecast without over/under fitting [29]. For a detail review on choice modelling you can refer to: [1], [30], [31]. A good categorization and review of the majority of demand unconstraining methods is presented in [32].

In the context of demand estimation and forecasting in airline revenue management systems, our paper brings three main contributions to the literature. First, this work contributes

to the literature by introducing an easy-to-implement non-parametric estimation algorithm for estimating multinomial logit choice model parameters in a reasonable computation time and accuracy. Second, the objective function is perturbed to a convex form by devising a simple heuristic which, shortens the execution time and guarantees the convergence. Third, a simple method proposed to estimate the customer arrival and no purchase in each period, without considering any a priori distribution, and just by the means of looking at the aggregated sale and market share in each period.

### III. PROBLEM DESCRIPTION

Airlines start to sell their flights about a year before departure. To analyse the customer behaviour this long selling time horizon is divided to some periods and discretized. These periods may differ in length, for example a period may be a week while the other may be a month. Then the aggregated sales of each offered product (seats with different conditions) are observed and recorded in each period. So the available dataset, consisting of aggregated sales for each product in each period. Table 1, shows a sample of such a dataset which is from an airline that offers 4 classes (C1 to C4) in 10 periods before flights. As you can see in this table, the aggregated sales of each product are recorded in each period.

TABLE I. A SAMPLE OF AGGREGATED RECORDED SALES IN DIFFERENT PERIODS BEFORE FLIGHT. N MEANS THE PRODUCT IS NOT AVAILABLE IN THAT PERIOD. PERIOD 0 IS THE DEPARTURE DAY

Products	9	8	7	6	5	4	3	2	1	0
C1	13	15	13	N	N	N	N	N	N	N
C2	11	6	7	9	10	N	N	N	N	N
C3	1	8	3	5	8	4	9	N	N	N
C4	4	9	4	7	7	3	9	6	7	5

TABLE II. CLASSES AND THEIR FEATURES OFFERED BY AN AIRLINE

Classes	Miles Earned	Change	Cancellation	Price \$
C1	100%	Charge 0%	Charge 0%	600
C2	100%	Charge 10%	Charge 20%	400
C3	50%	Charge 40%	Charge 50%	300
C4	30%	NO	NO	240

As it alluded to above, based on the theory of choice model the customers are rational utility maximizer and when they face with a choice list, they try to choose the choice which has the most benefit for them. If their first choice was not available, they may choose the second choice to buy in lieu to their first choice (this action is called substitution) or leave the system without purchase. Table 2 presents a sample of offered classes by an airline, along with their features.

To illustrate the problem, suppose that an airline offers 4 classes, C1 to C4 (see Table 2) in an aircraft with the capacity of 100, and the nested booking limits are set to 100, 80, 45 and 15. This could be translated to protecting 20 seats for class C1, 55 seats for classes C1 and C2 and 85 seats for classes C1, C2 and C3. There is no protection for class C4 as like as there is no limitation to sale class C1 except maximum aircraft

capacity. Suppose that 35 customers arrive. The observed sales for product C1 to C4 is (0, 0, 10, 15). The current information systems just register the successful sales and do not record any information about unsuccessful requests. Therefore, in such cases, the sales data could not show the true demands; for example, the sales data here shows 15 requests for class C4 (the observed demand), but due to C4 booking limit of 15, if the requests for this class were more than 15, system just able to record 15 successful sales and the other requests is not recorded. So this number of sales does not show the true amount of customer's request for C4. The observed sales for product C3 are 10. Although it is lower than its booking limit, this number does not show the true demand, because some of customers, whose first choice was C4, prefer to buy C3 instead of leaving the system without purchase. So this number of sales for class C3 is the mixture of customers whose their first choice are C3 and those whom their first choice are C4 but substitute to C3.

Here if these recorded sale is used as true demand, it is obvious that the true demand of C4 is under estimated while the true demand of C3 is over estimated. If these data are used as true demands and are fed into the forecasting module, a spiral down effect will occur in estimation, and the total revenue begins to decrease [3]. In this paper our goal is to find that how the customers prioritize the offered products and choose among them. Hence it is important to model the customer's buying behaviour. In this paper, it is supposed that the customers behave like in the customer choice model. The next section describes the approach to model the demand based on a discrete customer choice model.

### IV. DEMAND MODEL AND PROBLEM FORMULATION

Discrete choice model is the theoretical basis of customer behaviour. In 2000, the Nobel prize in economics has been awarded to American economists James Heckman and Daniel McFadden. Daniel McFadden proposed the discrete choice model, which is the basis for the development of the customer choice behaviour [33], [34]. The MNL model of McFadden was the primary basis for analysis of multinomial choice for many years. MNL assumes consumers have homogeneous tastes for observed product attributes, and that the random (unobserved) part of utility is iid. In other word all customers use the following equation to calculate the product j's utility with the same parameters.

$$i. u_j = \beta^T x_j + \epsilon_j \tag{1}$$

Here in (1),  $\beta$  is a vector of parameters and  $x_j$  is a vector of attribute values for product j which could include factors such as prices, rewards, time of departure, length of flight, cancellation and change policy and etc.  $\epsilon_j$  is the random component. In the MNL model, the probability that an alternative j is chosen from a set of offered products S, which contains product j is given by:

$$P_j(S) = \frac{e^{u_j}}{\sum_{i \in S} e^{u_i}} \tag{2}$$

The MNL possesses a restrictive property known as the independence from irrelevant alternatives (IIA) property [1]. For a famous example which describes IIA (red bus/blue bus paradox) please refer to [35]. As a result of IIA, the MNL model must be used with caution. It should be restricted to choice sets that contain alternatives that are dissimilar. Despite this deficiency, the MNL model is widely used in estimating travel demand. The popularity of MNL goes back to being analytically tractable, relatively accurate, and can be estimated easily using standard statistical techniques. The product utility is needed to calculate the probability of selecting an item from choice set. The product utility could be estimated from the sales data. For simplicity, the product preference weight is defined same as in [18]:

$$v_j = e^{u_j} \quad (3)$$

And the probability of choosing product j from choice set S is become:

$$P_j(S) = \frac{v_j}{\sum_{i \in S} v_i} \quad (4)$$

Now, assume that an airline offers n classes to its customers. The selling time horizon before departure, is divided into T periods:  $t=1 \dots T$ . In each period t, a set of classes ( $C_t$ ) is offered. The number of observed sales of class j in the period t is  $o_{jt}$ , and  $O_t$  is the observed sale vector in interval t:  $O_t = (o_{1t}, o_{2t}, \dots, o_{nt})$ . The customers may decide to not purchase anything when facing with the offered choices, so  $c_0$  is added to the choice set as no purchase alternative. The utility of no purchase is zero, so, the preference weight of no purchase is  $v_0 = 1$ . Using these preference weights the MNL choice probability is written as in [36], the probability of choosing class j in period t:

$$P(j, t) = \frac{v_j}{\sum_{i \in C_t} v_i + v_0} \quad (5)$$

The probability of choosing class j which is not available in period t is:  $P(j, t) = 0$ .

Based on no purchase preference weight or  $v_0 = 1$  the no purchase probability is defined as:

$$P(0, t) = \frac{v_0}{\sum_{i \in C_t} v_i + v_0} = \frac{1}{\sum_{i \in C_t} v_i + 1} \quad (6)$$

Suppose that the number of arrivals in each period t is known to be  $A_t$ , and if all classes were available in period t then, the probable number of sales of class j in the period t or true demand, is:

$$d_{jt} = P(j, t) \cdot A_t = \frac{v_j}{\sum_{i=1}^n v_i + v_0} A_t \quad (7)$$

The total sale in period t is:

$$E_t = \sum_{j \in C_t} d_{jt} = \frac{\sum v_j}{\sum v_j + v_0} A_t \quad (8)$$

But in the real world application the number of arrival is not visible, so it should be estimated. Suppose the preference weights are known. As an estimate, consider that the aggregated number of sales in period t ( $q_t$ ) is equal to the aggregated number of estimated sales in period t ( $E_t$ ), so the number of arrival could be found using:

$$q_t \approx E_t \Rightarrow q_t \approx \sum_{j=0}^n d_{jt} = \frac{\sum v_j}{\sum v_j + v_0} A_t \Rightarrow A_t \approx q_t \cdot \frac{\sum v_j + v_0}{\sum v_j} \quad (9)$$

The number of customers who could not buy their first choice in period t is estimated by using:

$$s_t = \sum_{j \notin C_t} d_{jt} \quad (10)$$

The number of requests which choose class k as a substitute in period t, could be estimated using:

$$r_{kt} = s_t \cdot P(k, t) = s_t \cdot \frac{v_k}{\sum_{j \in C_t} v_j + v_0} \quad k \in C_t \quad (11)$$

The number of observed sales of class j in period t or  $o_{jt}$  consists of two components: 1) the number of customers interested in class j as their first choice (true demand or  $d_{jt}$ ); 2) the number of customers buying j because their first choice was not available (recapture or  $r_{jt}$ ) [21], [37]. Thus, the demand mass balance equation is as follows:

$$o_{jt} = d_{jt} + r_{jt} \quad (12)$$

It supposed that the preference weights is known, but the fact is that, the values of preference vector is not known, so at first the preference vector V is needed to compute arrival in each period, spill and recapture. Therefore, first the preference vector is estimated, and then by using it and applying foregoing equations the number of true demands, recaptures, and spills are estimated.

## V. ALGORITHM

The Airline's true demand estimation algorithm is an iterative algorithm which starts with an initial value for preference vector and tries to make a better estimate for preference vector in each iteration by minimizing an error function which is the gap between estimated and observed sales in each period:

$$e_{jt} = o_{jt} - (d_{jt} + r_{jt}) \quad (13)$$

The objective function which should be minimized is a quadratic error function (Least Square Error):

$$\min_{v_j} (E = \sum_{t=1}^T \sum_{j=1}^N \frac{1}{2} \cdot e_{jt}^2) \quad (14)$$

### A. Analysis of Objective Function

To analysis the objective function, the selling time horizon is divided into two parts: 1) periods which all classes are available; and 2) periods which some classes are not available. In the first part, since all classes are available the spill and recapture is zero, so the observed sales are equal to true demand. Using (7), (9) and (14) the objective function is rewritten as:

$$\min_{v_j} (E = \sum_{t=1}^T \sum_{j=1}^N \frac{1}{2} \left( o_{jt} - q_t \cdot \frac{v_j}{\sum_{i=1}^n v_i} \right)) \quad (15)$$

Because of the fractional term, the objective function is not convex, also the Hessian matrix is not positive definite. To illustrate its non-convexity, assume a company offers two classes A and B in 10 periods. At the first 6 periods both of the classes were available, but in the rest of periods, just class B was available. Based on (15) the error function in the first 6 periods which all classes were available (so no spill and recapture exist) becomes:

$$e = \left( x[t, 1] - v1 * \frac{\sum_{i=1}^2 x[t,i]}{v1+v2} \right)^2 + \left( x[t, 2] - v2 * \frac{\sum_{i=1}^2 x[t,i]}{v1+v2} \right)^2 \quad (16)$$

The 3D plot of (16) is shown in Fig. 1. In the second part, which some classes are not available, some of the arrived customers prefer to buy substitute classes (recapture) or leave the system without purchase (spill). Using (7), (9), (13) and (14) the objective function is rewritten as:

$$\min_{v_j} (E = \sum_{t=1}^T \sum_{j=1}^N \frac{1}{2} \left( o_{jt} - q_t * \frac{\sum_{i \in C_t} v_i + v_0}{\sum_{i \in C_t} v_i} * \frac{v_j}{\sum_{i=1}^n v_i + 1} - s_t * \frac{v_j}{\sum_{i \in C_t} v_i + v_0} \right)) \quad (17)$$

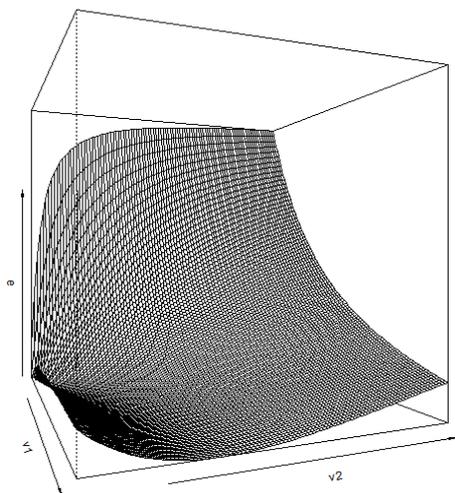


Fig. 1. The 3D plot of error function in the periods that all products are available, for a company with two offered classes A and B in 10 periods.

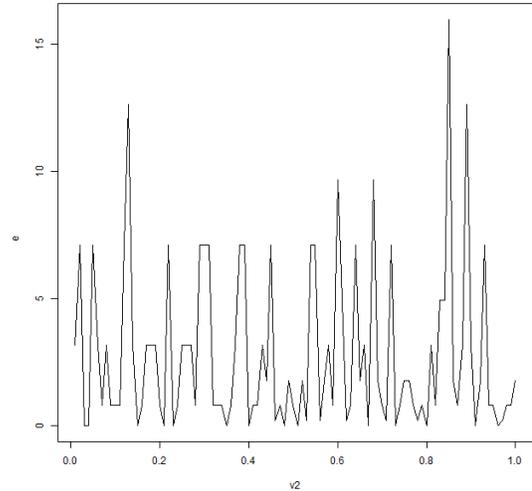


Fig. 2. Plot of error function in the second part of periods that some products are not available, for a company with two offered products A and B in 10 periods.

In above foregoing airline company example, class A is not available in the last 4 periods. So, because of spilled and recaptured demands, the error function is like:

$$e = \left( x[t, 2] - (v2 + 1) * \frac{x[t, 2]}{v1 + v2 + 1} - v1 * \frac{x[t, 2]}{v1 + v2 + 1} \right)^2 \quad (18)$$

The value of v1 is not updated as class A is not available, so it assumed that, the value of v1 is set to its optimal value, then the error function with fixing v1 at its optimal value is plotted. The plot is shown in Fig. 2 and obviously is ill structured and not convex.

### B. Convexification Process

Because of the fractional term in the demand (7), arrival (9) and the recapture (11), the objective function is not convex. A simple heuristic is devised to make that function convex. As it said before, the nonconvexity is because of the existence of the fractional term in the demand function, so if the denominator ( $\sum v_j + v_0$ ) is replaced with a fixed value, the objective function becomes convex. This value could be extracted from market share. The market share (M) is considered same as that in [38]:

$$M = \frac{\sum v_j}{\sum v_j + v_0} \xrightarrow{v_0=1} \sum v_j = M / (1 - M) \quad (19)$$

Fig. 3 shows the error function after convexification.

The proposed method initiated the values of preference weights equally using (19). For example, the product j's initial preference weight is:

$$v_j = \left( \frac{M}{1 - M} \right) / (\text{number of products}) \quad (20)$$

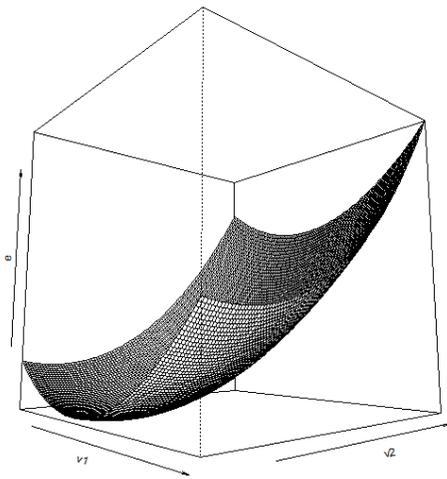


Fig. 3. 3D plot of error function after convexification.

---

**Pseudo Code of the algorithm:**

---

1. Ct: Set of available classes in period t.
2. V: Preference weight vector.
3. At: The number of arrival in period t.
4. Ot: The vector of purchases in period t.
5. M: The market Share
6. n: the number of offered classes
7. Initialize :
  - Estimating  $v_j$
  - For each class j in [1..n]
 
$$v_j = \left( \frac{M}{1-M} \right) / n$$
8. For each period t in [1..T]
  - Compute At based on  $V_j$ s:  $A_t = q_t \cdot \frac{\sum v_j + v_0}{\sum v_j}$
  - Compute no purchase:  $d_{0t} = A_t - q_t$
  - For each class j in [1..n]
    - $d_{jt} = \frac{v_j}{\sum v_j + v_0} A_t$
    - if (cjt not available)
      - $V_{j,new} = V_{j,old}$
      - $S = S + d_j$
    - else
      - $r_j = s \cdot \frac{v_j}{\sum_{j \in Ct} v_j + v_0}$
  - $V_{j,new} = \text{mean} \left( \frac{O_{jt}}{\left( \frac{A_t}{\sum v_j} \right) + \left( \frac{S}{\sum_{j \in Ct} v_j + v_0} \right)} \right)$

---

**VI. DATA AND SIMULATION PROCESS**

This section describes the simulation process and the data which is used for evaluating proposed algorithm. To have a dataset of customer sales, 6 types of information are needed: 1) offered product set; 2) product availability information during the selling time horizon; 3) the number of periods before flight; 4) the customer arrival process; 5) the customer buying behaviour; 6) the market share of the company. So the simulation process is as follows:

1) It is considered that an airline company offers a set of products in each flight. So for example, it is assumed that the company offers 4 classes in 10 periods. Then 100 flight data are generated.

2) In the simulation process, the product availability info is set exogenously. The product availability info is in matrix format, with the number of columns equal to the number of periods and the number of rows is equal to the number of products. For example  $AVL_{i,j} = 1$  means that the product i is available in period j and vice versa. This information also could be generated randomly.

3) The number of periods which the aggregated sales of each class are observed is considered different for each dataset. For example, assume 10 periods for selling the seats before the flight.

4) In the simulation process, it assumed that, customers arrive based on a poison process. The mean of poison distribution is different for each dataset. Here for 10 periods, 4 classes example, a poison arrival process with the mean of 60 is considered.

5) To simulate the customer's buying behaviour, a customer choice model is used. This is a predefined model with predefined products and their preference weights. This model may pick from real-world data or synthetic datasets. For example, consider a model with offered classes, c1 to c4 and the preference vector of  $V = (0.85, 0.68, 0.33, 0.14)$ .

6) The market share of the company is considered different for each dataset. In our example, the market share is considered 67%.

7) To simulate the demands, the selected model is used to simulate the customer's request to buy the products in each period. By having the simulated number of arrivals in each period, the number of sales for each product and the number of no purchase could be simulated using foregoing equations. This generated dataset, which contains the unconstrained demands is served as a benchmark to evaluate the accuracy and performance of the proposed method.

8) By applying the product availability info on simulated sales from previous step a censored observation of the dataset is achieved. For example, if the product i was not available on period j (i.e.  $AVL_{i,j}=0$ ) the simulated sales of that product on foregoing period reset to zero (ie.  $d_{ij} = 0$ ). This dataset is used as input to our unconstraining algorithm.

9) By executing the proposed algorithm on the generated dataset in step viii, the preference vector is estimated.

10) Using estimated preference vector, the true demand, spill and recapture could be calculated.

A sample of simulated dataset is available in Table 3. The number of arrivals and no purchase is not visible in real-world application; thus, these values are separated in two rows entitled Hidden Data.

Each dataset contains 100 instances of flights, which each instance contains a set of classes and the number of sales of each class on every period. For example, in the 4X10 dataset which contains a set of 4 offered classes and 10 intervals, 00 instances of flight exists. This means totally  $100 \cdot 10$  intervals that 4 products is offered so 4000 sales data records is generated in this dataset.

TABLE III. A SAMPLE OF SIMULATED SALES DURING PERIODS BEFORE DEPARTURE. N MEANS THE CLASS IS NOT AVAILABLE IN THAT PERIOD

Classes	9	8	7	6	5	4	3	2	1	0
C1	13	15	13	N	N	N	N	N	N	N
C2	11	6	7	9	10	N	N	N	N	N
C3	1	8	3	5	8	4	9	N	N	N
C4	4	9	4	7	7	3	9	6	7	5
↓ Hidden Data ↓										
No Purchases	47	42	46	72	58	76	75	68	83	9
#Arrivals	76	80	73	93	83	83	93	74	90	4

VII. RESULTS

To evaluate the performance and the accuracy of the proposed algorithm, the algorithm is executed on the observed sales of all simulated datasets. As a result of applying the proposed algorithm on simulated dataset, preference values are estimated for each dataset. Table 4 shows a sample of estimated preference vector in contrast to the true value of preference vector which is used to generate the simulated dataset. There are two favourable attributes for an estimator: accuracy and precision. Accuracy is lack of bias and precision is small variance. If an estimator is unbiased, then its variance is investigated. If it is biased, it is good to look at the mean squared error. As the Table 4 shows, the proposed estimator has a negative bias on estimated values and this is because of perturbing the objective function to a convex form and the new convex function is always below the true values.

Table 5 shows the mean, variance, mean square error, accuracy and precision of the proposed algorithm for estimated values of preference vector. As Table 5 shows the values of MSE and Variance are almost equal and this is because of the mean of the estimated preference vector is almost equal to the true value of preference vector. Here in Fig. 4, the box plot of estimated preference values for 100 executions of the algorithm over simulated datasets is shown. The line in the middle of the box is the median. The box itself represents the middle 50% of the data. The box edges are the 25th and 75th percentiles. There is some disk visible below or above of some boxes which is shown in some cases the simulated sales were far from the regular values. So, the estimated preference of that simulated dataset is far from the median.

TABLE IV. ESTIMATED PREFERENCE VECTOR

Classes	Preference Vector	True Value	Estimated Value	Estimator Bias
C1	V1	0.85	0.822	-0.028
C2	V2	0.68	0.659	-0.020
C3	V3	0.33	0.327	-0.002
C4	V4	0.14	0.139	-0.0009

TABLE V. MEAN, VARIANCE AND MEAN SQUARE ERROR OF DIFFERENCE BETWEEN ESTIMATED AND TRUE VALUES OF PREFERENCES.

Classes	Mean	Variance	MSE	Accuracy%	Precision%
C1	-0.028	0.0023	0.0031	97.16	99.7
C2	-0.020	0.0020	0.0024	97.91	99.8
C3	-0.002	0.0014	0.0014	99.73	99.8
C4	-0.0009	0.0004	0.0004	99.90	99.9

TABLE VI. ROOT MEAN SQUARE ERROR BETWEEN ESTIMATED TRUE DEMAND AND SIMULATED TRUE DEMAND IN CONTRAST TO EM AND PD.

Methods	C1	C2	C3	C4
Our Method				
RMSE	3.39	2.49	1.29	0.59
EM				
RMSE	2.34	3.68	3.9	2.34
PD				
RMSE	2.20	3.59	3.78	2.34
Vulcano EM				
RMSE	3.57	2.61	1.39	0.66

Figure 5 shows the cumulative sum of simulated demand, estimated demand, observed sales and estimated sales over 30 periods for products c1 to c4. It is obvious that always estimated sales is close, but a bit upper than observed sales. One of the important aspect of the simulation is that even the hidden data which are not observable in the real world are generated, hence it is easy to measure the accuracy of the estimation process. Here, the estimated true demand is compared with simulated true demand data. In this way, we have the ability to show how much accurate the forecasting is, and this will indicate the accuracy of estimated preference vector. To have a better insight about the results and the accuracy of proposed method, the results are compared to three other methods which are usually applied in practice: 1-Expectation Maximization or EM, 2-Projection Detruncation or PD methods and 3-EM method that proposed by Vulcano et al in 2012 [18],[39]–[41]. To compare the results, three above methods are applied on a dataset for 1000 times, then the mean of root mean square error for each method is calculated. The result of the comparison is depicted in table VI. The product C4 which is the most expensive product, is always available in all periods. The EM and PD methods, assume that the observed sales of product 4 is equal to true demand because these methods do not consider the substitution. The value of 2.34 which is seen in the table VI for the root mean square error of product c4 in EM and PD methods shows that the true demand is differ from observed sales. Here the observed sales are a mixture of true demand and recaptured demand, hence, the observed sales are greater than the true demand. Our proposed method and the EM method which is presented by Vulcano et al. are able to compute the spill and recapture from observed sales data and so, they are more accurate than EM and PD. If there is no substitution, the EM and PD methods are also accurate, even more accurate than our method or the Vulcano’s EM method.

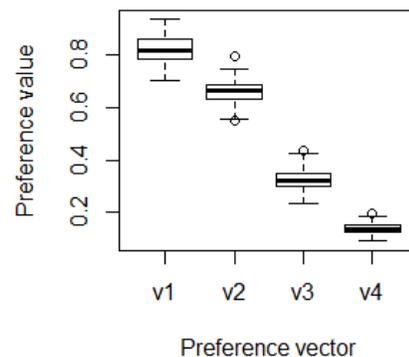


Fig. 4. The Boxplot of preference values.

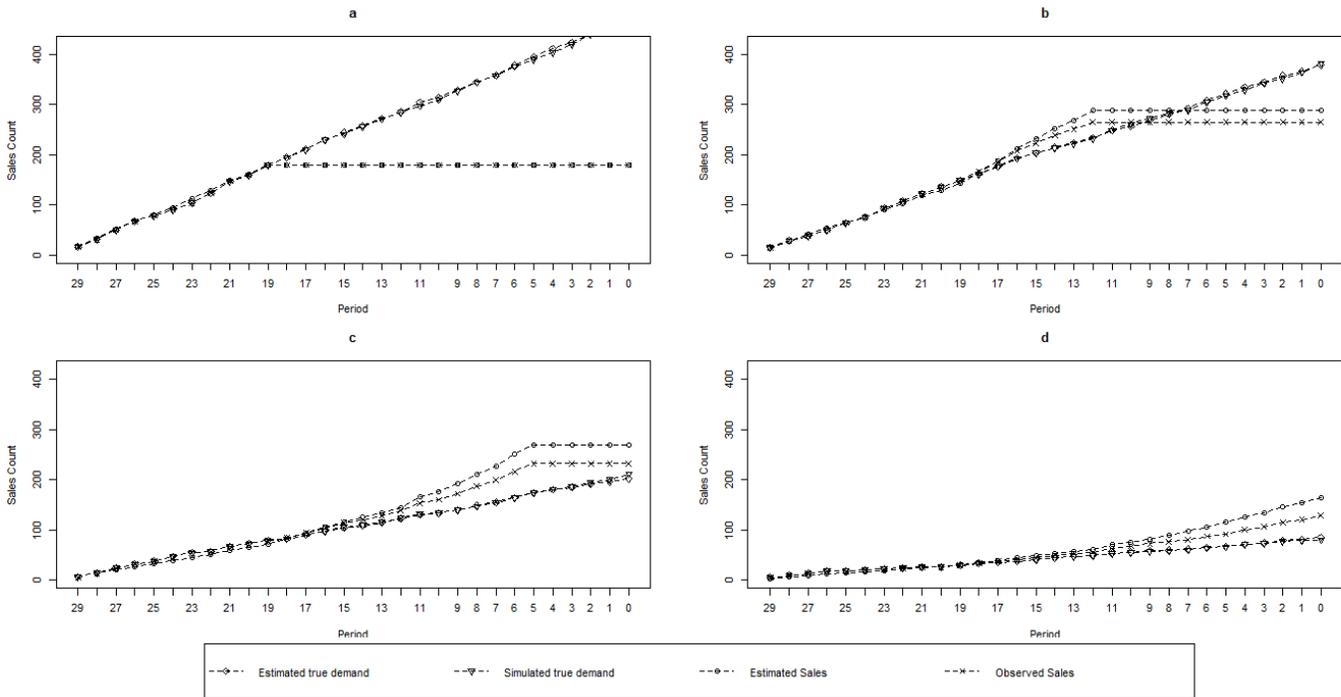


Fig. 5. The cumulative sum of simulated and estimated demand, observed sales and estimated sales over 30 periods for classes c1(a), c2(b), c3(c) and c4(d)

TABLE. VII. EXECUTION TIME FOR DIFFERENT DATASETS

Dataset	Execution Time for 100 flights (sec)
5X15	2.01
4X30	3.61
5X30	3.72
10X30	4.70
10X60	9.12
20X60	12.77

This could be deduced by looking at column C1 of Table 6. As you can see in this column the root mean square error of EM and PD are smaller than the others, because product C1 is the least expensive and more favourable product in the choice set and in the periods that this product is available, all of the other classes are available too. So the observed sales of product C1 are always equal to its true demand. The results show that our proposed method out performs the other three, and also has up to 47.64% improvement in root mean square error in contrast to the other mentioned methods.

Different size datasets used to show the efficiency and speed of the proposed algorithm. At first it is evaluated with a rather small dataset with 5 products and 15 periods, in this case the algorithm lasts 2.01 seconds to estimate 100 instances of flights. The second dataset is composed of observed sales from a flight with 4 classes and 30 periods. The third dataset is for a flight having 5 classes sold in 30 intervals. Finally, in the last dataset which is our biggest dataset, there exist 20 classes in a flight with 60 time intervals. For the last datasets algorithm takes about 12.77 seconds to solve 100 instances, which is very good for this size of dataset in contrast to current demand

unconstraining methods. Table 7 shows the execution time for all six datasets. R revision 3.2.2 used to implement the algorithm on a computer with an Intel core2 Quad Q8300 and 4 GB of internal memory. Code execution time is measured by `proc.time()` in R[42]. Column 1 of Table 7 shows the datasets and column 2 is the mean convergence time.

## VIII. CONCLUSION

In this paper, a new approach for demand unconstraining is proposed. Multinomial logit model is used to model the customer choice behaviour. Transactional sales data, product availability and market share are the only data used to estimate the choice model parameters. Spilled and recaptured demands are estimated alongside with true demand. The problem of double counting demands in available classes is resolved by considering spill and recapture. The customer arrival is estimated using a simple heuristic. In this heuristic method, the customer arrival count is assumed to be a fraction of aggregated sales in each period (see (9)).

Unlike the proposed method, most of current in practice choice based methods, assume a priori distribution to estimate, customer arrival rate in each period, so suffer from misspecification of the distribution function and its parameters. Numerical experiments prove that proposed method performs well in terms of the speed and the accuracy of estimation. Based on the simulation results, our method improved root mean square error between simulated and estimated demands by 47.64%, in contrast to other methods such as PD, EM and another version of EM proposed by Vulcano et al. It is observed that if there is no substitution in the datasets, the EM and PD methods perform well and even better than the other methods. But in most cases the substitution exists between classes and this affects the true demand. So the EM and PD

methods are not able to perform well in these kind of datasets. For a rather large dataset of size 20 products and 60 periods, it takes 0.127 seconds to estimate the true demand, spill and recapture, which is fast in this context.

Based on the experiment's result, we believe that the proposed method would be a good replacement for current demand unconstraining methods in airline revenue management systems.

#### REFERENCES

- [1] K. T. Talluri and G. J. Van Ryzin, *The Theory and Practice of Revenue Management*, vol. 1. 2005.
- [2] A. Haensel and G. Koole, "Estimating unconstrained demand rate functions using customer choice sets," *J. Revenue Pricing Manag.*, vol. 10, no. February, pp. 438–454, 2011.
- [3] W. L. Cooper, T. Homem-de-Mello, and A. J. Kleywegt, "Models of the Spiral-Down Effect in Revenue Management," *Oper. Res.*, vol. 54, no. December 2014, pp. 968–987, 2006.
- [4] R. Saleh, "Estimating lost demand with imperfect availability indicators," in *AGIFORS Reservations and Yield Management Study Group*, 1997.
- [5] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*. 2002.
- [6] A. M. Fouad and A. F. Atiya, "A New Unconstraining Method for Demand Forecasting," pp. 16–19, 2012.
- [7] C. Stefanescu, "Multivariate Demand: Modeling and Estimation from Censored Sales," pp. 1–33, 2009.
- [8] B. Tan and S. Karabati, "Retail inventory management with stock-out based dynamic demand substitution," *Int. J. Prod. Econ.*, vol. 145, no. 1, pp. 78–87, 2013.
- [9] T. Gruen, D. Corsten, and S. Bharadwaj, "Retail Out-of-Stocks: A Worldwide Examination of Extent, Causes and Consumer Responses," 2002.
- [10] K. Campo, E. Gijsbrechts, and P. Nisol, "The impact of retailer stockouts on whether, how much, and what to buy," *Int. J. Res. Mark.*, vol. 20, no. 3, pp. 273–286, 2003.
- [11] J. Aastrup and H. Kotzab, "Analyzing out-of-stock in independent grocery stores: an empirical study," *Int. J. Retail Distrib. Manag.*, vol. 37, no. 9, pp. 765–789, 2009.
- [12] G. Xin, P. R. Messinger, and J. Li, "Influence of soldout products on consumer choice," *J. Retail.*, vol. 85, no. 3, pp. 274–287, 2009.
- [13] R. R. Wickham, "Evaluation of forecasting techniques for short-term demand of air transportation," MIT, 1995.
- [14] K. T. Talluri, G. J. van Ryzin, I. Z. Karaesmen, and G. J. Vulcano, "Revenue management: Models and methods," in *2008 Winter Simulation Conference*, 2008, pp. 145–156.
- [15] L. R. Weatherford, "A review of optimization modeling assumptions in revenue management situations," in *AGIFORS Reservations and Yield Management Study Group*, 1997.
- [16] A. Nikseresht and K. Ziarati, "Review on the Newest Revenue Management Demand Forecasting Methods," in *International Conference on Management, Economics and Industrial Engineering*, 2015, vol. 1, no. 1.
- [17] A. Nikseresht and K. Ziarati, "A Non-Parametric Customer Choice-Based Method to Unconstrain Airline Sales Transaction Data," *J. Air Transp. Manag.*, 2016.
- [18] G. Vulcano, G. van Ryzin, and R. Ratliff, "Estimating Primary Demand for Substitutable Products from Sales Transaction Data," *Oper. Res.*, vol. 60, no. 2, pp. 313–334, 2012.
- [19] G. van Ryzin and G. Vulcano, "A Market Discovery Algorithm to Estimate a General Class of Nonparametric Choice Models," *Manage. Sci.*, vol. 61, no. 2, pp. 281–300, 2015.
- [20] A. Hübner, H. Kuhn, and S. Kühn, "An efficient algorithm for capacitated assortment planning with stochastic demand and substitution," *Eur. J. Oper. Res.*, vol. 250, no. 2, pp. 505–520, Apr. 2016.
- [21] S.-E. Andersson, "Passenger choice analysis for seat capacity control: A pilot project in Scandinavian Airlines," *Int. Trans. Oper. Res.*, vol. 5, no. 6, pp. 471–486, 1998.
- [22] M. B. Staffan Algers, "Modelling choice of flight and booking class - a study using Stated Preference and Revealed Preference data," *Int. J. Serv. Technol. Manag.*, vol. 2, no. 1/2, 2001.
- [23] R. M. Ratliff, B. Venkateshwara Rao, C. P. Narayan, and K. Yellepeddi, "A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings," *J. Revenue Pricing Manag.*, vol. 7, no. 2, pp. 153–171, 2008.
- [24] K. Talluri and G. van Ryzin, "Revenue Management Under a General Discrete Choice Model of Consumer Behavior," *Manage. Sci.*, vol. 50, no. 1, pp. 15–33, Jan. 2004.
- [25] A. Haensel, G. Koole, and J. Erdman, "Estimating unconstrained customer choice set demand: A case study on airline reservation data," *J. Choice Model.*, vol. 4, no. 3, pp. 75–87, 2011.
- [26] J. P. Newman, M. E. Ferguson, L. A. Garrow, and T. L. Jacobs, "Estimation of choice-based models using sales data from a single firm," *Manuf. Serv. Oper. Manag.*, vol. 16, no. 2, pp. 184–197, 2014.
- [27] G. Vulcano, G. van Ryzin, and W. Chaar, "OM Practice--Choice-Based Revenue Management: An Empirical Study of Estimation and Optimization," *Manuf. Serv. Oper. Manag.*, vol. 12, no. 3, pp. 371–392, 2010.
- [28] S. Agrawal, V. Avadhanula, V. Goyal, and A. Zeevi, "A near-optimal exploration exploitation approach for assortment selection," in *EC 2016 - Proceedings of the 2016 ACM Conference on Economics and Computation*, 2016, pp. 599–600.
- [29] V. F. Farias, S. Jagabathula, and D. Shah, "A Nonparametric Approach to Modeling Choice with Limited Data," *Manage. Sci.*, vol. 59, no. March 2015, p. 44, 2013.
- [30] B. Wierenga, *Handbook of marketing decision models*. Springer, 2008.
- [31] S. R. Chandukala, J. Kim, T. Otter, P. E. Rossi, and G. M. Allenby, "Choice Models in Marketing: Economic Assumptions, Challenges and Trends," *Found. Trends Mark.*, vol. 2, no. 2, pp. 97–184, 2007.
- [32] S. Sharif Azadeh, P. Marcotte, and G. Savard, "A taxonomy of demand uncensoring methods in revenue management," *J Revenue Pricing Manag.*, vol. 13, no. 6, pp. 440–456, 2014.
- [33] W. Qi, X. Luo, and X. Liu, "Discrete choice model of customer behavior and empirical study," in *2016 Chinese Control and Decision Conference (CCDC)*, 2016, pp. 5677–5682.
- [34] D. L. McFadden and P. Zarembka, "Frontiers in econometrics," *Cond. logit Anal. Qual. choice Behav.*, p. 105, 1974.
- [35] G. Debreu, "Review of R. D. Luce, individual choice behavior: A theoretical analysis," *Am. Econ. Rev.*, vol. 50, pp. 186–188, 1960.
- [36] M. N. Lee G. Cooper, *Market-Share Analysis*. Boston Dordrecht London: Kluwer Academic Publishers, 2010.
- [37] S. Ja, S., Rao, B. V. Chandler, "Passenger recapture estimation in airline RM," in *AGIFORS 41st Annual Symposium*, 2001.
- [38] T. S. G. and D. Sudharshan, "Equilibrium Characteristics of Multinomial Logit Market Share Models," *Mark. Res.*, vol. 28, no. 4, pp. 480–482, 1991.
- [39] R. H. Zeni, "Improved Forecast Accuracy In Revenue Management By Unconstraining Demand Estimates From Censored Data," *The State University of New Jersey*, 2001.
- [40] P. Guo, B. Xiao, and J. Li, "Unconstraining methods in revenue management systems: Research overview and prospects," *Adv. Oper. Res.*, vol. 2012, 2012.
- [41] T. Abdallah and G. Vulcano, "Demand Estimation under the Multinomial Logit Model from Sales Transaction Data," no. May, 2016.
- [42] S. L. Niël J le Roux, *A Step-by-Step R Tutorial: An introduction into R applications and programming*, 1st ed. 2015.

# 2.5 D Facial Analysis via Bio-Inspired Active Appearance Model and Support Vector Machine for Forensic Application

Siti Norul Huda Sheikh Abdullah  
Faculty of Information Science and  
Technology,  
Universiti Kebangsaan Malaysia,  
Bandar Baru Bangi, Malaysia

Nazri Ahmad Zamani  
Digital Forensics lab,  
CyberSecurity Malaysia,  
Seri Kembangan,  
Malaysia

Khairul Akram Zainol Ariffin  
Faculty of Information Science and  
Technology,  
Universiti Kebangsaan Malaysia,  
Bandar Baru Bangi, Malaysia

Mohammed Hasan Abdulameer  
Department of Computer Science,  
Faculty of Education for Women,  
University of Kufa,  
Iraq

Fasly Rahim  
Faculty of Science and Technology  
Universiti Kebangsaan Malaysia,  
Bandar Baru Bangi,  
Malaysia

Zulaiha Othman  
Faculty of Information Science and  
Technology,  
Universiti Kebangsaan Malaysia,  
Bandar Baru Bangi, Malaysia

Mohd Zakree Ahmad Nazri  
Faculty of Information Science and Technology,  
Universiti Kebangsaan Malaysia,  
Bandar Baru Bangi, Malaysia

**Abstract**—In this paper, a fully automatic 2.5D facial technique for forensic applications is presented. Feature extraction and classification are fundamental processes in any face identification technique. Two methods for feature extraction and classification are proposed in this paper subsequently. Active Appearance Model (AAM) is one of the familiar feature extraction methods but it has weaknesses in its fitting process. Artificial bee colony (ABC) is a fitting solution due to its fast search ability. However, it has drawback in its neighborhood search. On the other hand, PSO-SVM is one of the most recent classification approaches. However, its performance is weakened by the usage of random values for calculating velocity. To solve the problems, this research is conducted in three phases as follows: the first phase is to propose Maximum Resource Neighborhood Search (MRNS) which is an enhanced ABC algorithm to improve the fitting process in current AAM. Then, Adaptively Accelerated PSO-SVM (AAPSO-SVM) classification technique is proposed, by which the selection of the acceleration coefficient values is done using particle fitness values in finding the optimal parameters of SVM. The proposed methods AAM-MRNS, AAPSO-SVM and the whole 2.5D facial technique are evaluated by comparing them with the other methods using new 2.5D face image data set. Further, a sample of Malaysian criminal real case of CCTV facial investigation suspect has been tested in the proposed technique. Results from the experiment shows that the proposed techniques outperformed the conventional techniques. Furthermore, the 2.5D facial technique is able to recognize a sample of Malaysian criminal case called “Tepuk Bahu” using CCTV facial investigation.

**Keywords**—Face recognition; active appearance model; ant bee colony; particle swarm optimization; support vector machine

## I. INTRODUCTION

The reliability of evidence in a court of law is dependent upon how the evidence is handled, how it is interpreted, and how it is presented. The fundamental principle behind digital forensics is maintaining integrity and provenance of media upon seizure, and throughout the analysis and handling process [1]. The evidence has to undergo a number of processes to ensure its integrity and admissibility into the court of law. In the past, a number of models have been considered as the best practices for conducting digital forensics and the common processes in those models are evidence identification, analysis and interpretation of the result and finally presentation of the findings [2]. In forensics, evidence can be explained with Locard’s Exchange Principle, which holds that the perpetrator of a crime will bring something into the crime scene and leave something from it [3]. This principle is the foundation of forensic sciences and this extends to the digital forensics, and video forensic is no exception. The role of forensic science is to uncover the traces left behind from a crime scene, whether it is physically, chemically or digitally.

Facial recognition is the application of biometrics [4], and in video exhibit, it is an essential forensic analysis in order to improve the clarity of the face for a human observer. The facial clarified is said to be important, as it will be used to assist law enforcement agency in their investigation, litigation or legal process. The challenges faced by this particular discipline of forensics contributed mostly by the quality of the exhibit itself. For example, surveillance video from CCTV

system is always backed up with down-sampled resolution. The surveillance video stored inside the CCTV's Digital Video Recorder (DVR) is normally kept at resolution 720 pixels wide with H.264 compression. After a certain period of time, the video resolution is down-sampled at lower size to accommodate more space for the recording. The qualities of the recording were found to be low with signal noise, color noise, compression artefacts, illumination problems and many kinds of blurring. Further, the people are usually walking free, and it is impossible to always keep faces frontal to the cameras. Hence, face images captured by the CCTV are non-ideal due to many factors: pose, illumination, expression, distance and weather [5]. These degradations have caused many dead ends to video forensic analysis. Consequently, the forensic investigation reports were challenged or rejected by court due to inefficient evidence presented by the reports.

The paper highlights on the methodology of 2.5 D Facial Analysis via Bio-Inspired Active Appearance Model and Support Vector Machine. It is organized as follows: In Section 2, the discussion on the application of face recognition in digital forensics will be outlined. In Section 3, the state of the art for face recognition is highlighted. Then it continues with the proposed model in Section 4 and the result is discussed in Section 5. Finally, the overall work of this paper is summarized in the last section.

## II. APPLICATION OF FACE RECOGNITION IN DIGITAL FORENSICS

Forensics facial identification system is a biometric system that analyses face-matching result for investigation or for litigation process. In other variation, this can also be defined as a face recognition system for the forensics purposes. As the term *identification* is being used, therefore it emphasizes the 1: N matching.

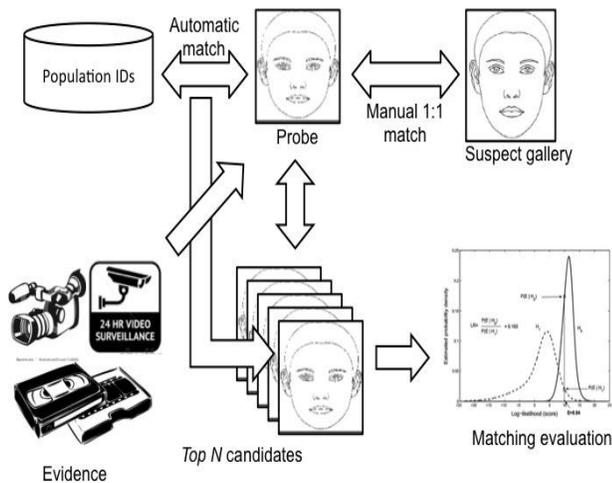


Fig. 1. The paradigm of forensics face recognition system

The aim of the analysis is to find the top  $N$  matches for a probe, in which is extracted from a photo or video exhibit. For the database of IDs, two types of IDs are prepared- 1) the population face enrolment IDs and 2) a gallery of suspect IDs face photos. The matches will look into the top  $N$  matches result and see if the suspect IDs comes up at the top or not.

The analysis would also examine on the matching accuracy and the numbers of false matches. The overall work flow for face recognition is shown in Fig. 1.

There are four applications of facial identification in digital forensics such as:

- Criminal justice: The facial identification system is used to ensure that the database does not contain multiple records for a single individual.
- Identity checks in the field: To get automated decisions without requiring expert analysis.
- Criminal investigation and information: Analysing surveillance videos, witness' camera, Internet sites or smartphones.
- Prevention: Background check to prevent people with criminal history. The technology can also be used for passport control's watch list.

## III. STATE OF THE ART

**From the point of forensic analysis view:** The first attempt of face recognition by humans was reported in a British court in 1871 where the suspect was identified by comparing a pair of facial photographs [6]. Then, the forensic tool for face recognition is introduced by Bertillon's work where in the practice, two photographs of a subject are captured, one from the front and one in profile. The same approach is still being used in law enforcement's mug shots [7]. The application of face recognition for forensic is at an advantage if the evidence is presented – whether it is a still image or a video. The image and video can be from both analogue and digital camera, provided the original storage or the negative are also seized. Furthermore, face recognition is a good modality for forensic as the biometric is not intrusive and non-contact [8]. Another advantage of face modality in forensic is that the evidence and the reference that contain the modality is at abundance. Face image is already in existence in any country's civil and criminal files. In addition to mug shots, face images are recorded and digitized for citizens' identity, licenses, passport and visas [9]-[11]. There is also a potential wealth of facial images recorded by alert observers using smartphones, and ever growing number of websites, blogs and social medias. Just like the old Bertillon's system, face recognition is used to establish the identity of the person inside, as evidence to identity which is a reference, for example, inside a surveillance video footage of a robbery. The video can be tended to court as primary evidence against apprehended suspects. To verify the men inside the video footage is the men apprehended face recognition is used.

**From the point of image processing view:** In verifying the evidence in the forensic work related to image and CCTV, the digital forensic analysts perform manual operations to match facial images with huge databases of mug-shots, which are most challenging and consumes huge time [12]. Nevertheless, a number of studies have focused on 2D and 3D face recognition. However, using 2D images for face recognition is sensitive to various factors, such as, illumination, pose variations, and facial expressions [5] [13]. On the other hand, using 3D image has been adversely

impacted by changes in facial expressions [14]. Furthermore, 3D techniques also encounter few other restrictions, including, costly gadgets, and computational complication [15]. In contrast, earlier studies have highlighted that 2.5D face image is capable to address the limitations of using 2D and 3D face image recognition [16]-[18]. One of the essential processes in face recognition is the feature extraction. Previously, the Active Appearance Model (AAM) is frequently used for the facial extraction. Although it is a popular method for the facial identification, this model has weaknesses in terms of fitting process [19].

**From the perspective of Bio-inspired machine learning algorithms:** Artificial Bee Colony (ABC) is an effective fitting solution and has been extensively used in many fields, owing to its rapid search potential [20]. The ABC has also been effectively applied to address challenging optimization problems [21]. However, this approach has a setback in terms of their neighborhood search, where it randomly generates new food sources. Classification is another crucial part in any face recognition technique. Previously, most of classification techniques use Particle Swarm Optimization with Support Vector Machine (PSO-SVM) [22] which employs PSO to find the optimal parameters for SVM, even though its performance has deteriorated by the usage of random values for calculating velocity.

#### IV. PROPOSED MODEL

The proposed face recognition technique is performed in two main phases, feature extraction and classification. The feature extraction stage uses the proposed active appearance model approach based on Maximum Resource Neighborhood Search (MRNS) which is an enhanced ABC algorithm (AAM-MRNS), and the classification stage is based on the proposed AAPSO-SVM [23]. These two phases are performed repeatedly on the input database face images, and thus the face images are recognized more effectively. The basic structure of the proposed 2.5D facial technique is shown in Fig. 2.

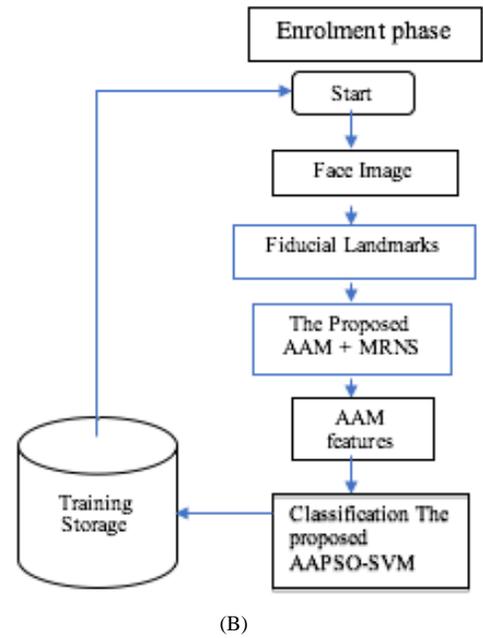
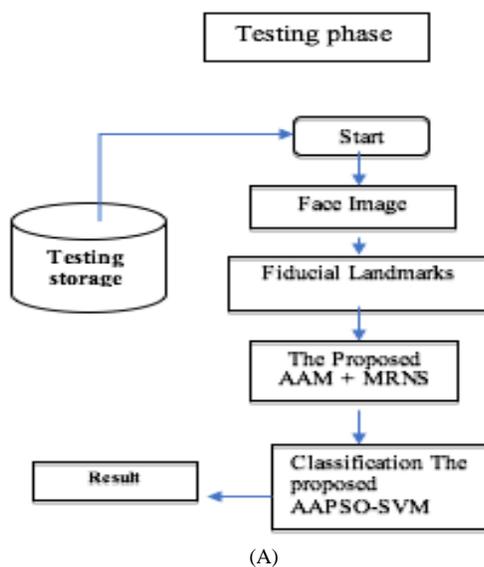


Fig. 2. The main components of the proposed 2.5D facial forensic technique; (A) for training, (B) for enrolment.

The proposed feature extraction technique utilizes AAM for extracting the shape and appearance features from the database images. In addition, it is crucial to have appropriate fitting for extracting AAM based features. Consequently, the fitting performance has been enhanced by presenting a new maximum resource neighborhood search (MRNS) algorithm, where the searching performance has been accelerated by considering the quality of the best food source. Depending on the quality of best food source, the new food sources are generated by the neighborhood search of the algorithm. The proposed feature extraction method is mainly comprised of 1) Feature Extraction using AAM Modeling; 2) the proposed MRNS Algorithm; 3) fitting using MRNS algorithm.

The standard PSO method has been utilized in many research works to obtain optimal problem solutions. The major drawback of the PSO is the random value selection during new particle generation; that is, in the velocity computation, the acceleration coefficients are generated randomly. The random value selection in the velocity process means that the generated particles will also be random. Random populations do not produce more accurate results. Hence, to acquire a more accurate result and to reduce this PSO drawback, adaptive acceleration particle swarm optimization (AAPSO) is proposed. In order to obtain more accurate classification results, the SVM parameters will be optimized by our AAPSO. The utilization of AAPSO in the SVM parameter optimization will reduce the PSO drawback and improve the classification result accuracy.

#### V. EXPERIMENTAL RESULT

This section presented the experimental results and performance of the proposed 2.5D facial forensic technique

which depends on the proposed AAM-MRNS and AAPSO-SVM methods. The performances have been evaluated based on the fitting error in feature extraction and accuracy rate of the whole classification process in this technique. The experiment has been conducted using the property 2.5D face dataset and a sample of Malaysian real case of CCTV facial investigation. Two evaluation steps have been used to analyze the final 2.5D facial forensic technique. The first step has evaluated the proposed 2.5D technique using fitting errors based on the performance of the proposed AAM-MRNS method; while the second step has evaluated the 2.5D technique using recognition accuracy rate based on the performance of proposed AAPSO-SVM method.



Fig. 3. Ten sample images from 2.5D face dataset prepared by UKM-Cybersecurity Digital Forensic Lab.

The performance of the proposed 2.5D technique has been evaluated by UKM property dataset namely 2.5D face images dataset. This dataset has been divided to training and testing images. 18 images have been utilized for training and 6 images have been used for testing. The performance of the proposed technique has been analyzed and compared to the conventional methods like, conventional AAM and Adaptive AAM [24]. The samples of 2.5D images utilized in the recognition process are shown in Fig. 3.

During the analysis of the performance, 10 rounds of experiments have been conducted with the 2.5 D face and iris images. Table 1 and Fig. 4 illustrate the results of fitting error values for the proposed and conventional techniques, acquired in 10 rounds of experiments. Based on the result in Table 1 and Fig. 4 the proposed 2.5D technique has achieved minimum fitting error as average as compared to the conventional AAM and adaptive AAM. The results reflect the stability in different datasets not only 2D datasets.

Fitting Error versus Methods

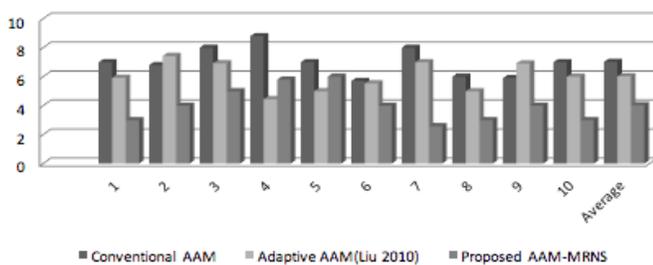


Fig. 4. Fitting error values of the proposed AAM-MRNS and conventional AAM techniques.

TABLE. I. FITTING ERROR FOR PROPOSED AND CONVENTIONAL RECOGNITION TECHNIQUES ON FACIAL DATASET

Exp	Fitting Error		
	Conventional AAM	Adaptive AAM [24]	Proposed AAM-MRNS
1	7	5.93	3
2	6.8	7.43	4
3	8	6.93	5
4	8.8	4.45	5.8
5	7	5	6
6	5.7	5.54	4
7	8	7	2.58
8	6	5	3
9	5.9	6.92	4
10	7	6	3
Average	7.02	6.02	4.038

In the accurate rate evaluation, the performance of proposed 2.5D technique has been evaluated by the accuracy rate based on AAM.MRNS+AAPSO.SVM. The recognition performance has been analyzed and compared with AAM ABC + OPSO.SVM, AAM.MRNS + OPSO-SVM, AAM.ABC + AAPSO-SVM methods, and with PSO-SVM which is one of the most recently developed face recognition approaches. Table 2 and Fig. 5 illustrate the accuracy of face images recognition of the proposed technique and others.

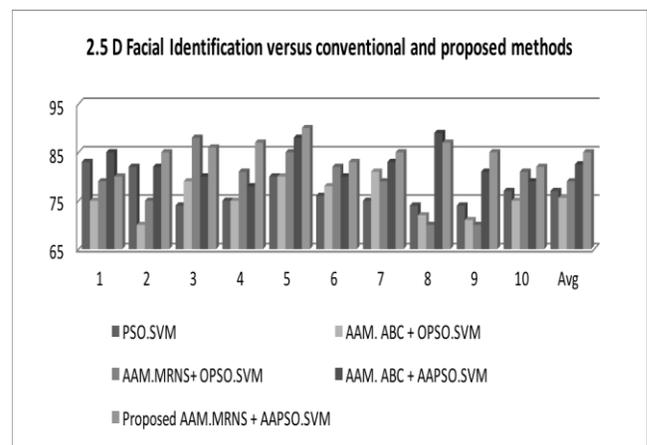


Fig. 5. Accuracy performance of the proposed 2.5 technique against the other techniques.

Based on the Table 2 and Fig. 4, it is evident that the proposed 2.5D technique that used AAM-MRNS for feature extraction and AAPSO-SVM for recognition, has given higher face image recognition accuracy, than the other techniques.

The high-performance results show that, the proposed 2.5D face identification technique is capable of more accurately recognizing the face images.

TABLE. II. PERFORMANCE OF PROPOSED 2.5D TECHNIQUES AND OTHER RECOGNITION TECHNIQUES FOR FACE DATASET.

Exp	Accuracy (%)				
	PSO.SVM	AAM.ABC + OPSO.SVM	AAM.MRNS + OPSO.SVM	AAM.ABC + AAPSO.SVM	Proposed AAM.MRNS + AAPSO.SVM
1	83	75	75	85	80
2	82	70	75	82	85
3	74	79	88	80	86
4	75	75	81	78	87
5	80	80	85	88	90
6	76	78	82	80	83
7	75	81	79	83	85
8	74	72	70	89	87
9	74	71	70	81	85
10	77	75	81	79	82
Ave	77%	75.6%	79%	82.5%	85%

In addition, a sample of Malaysian real case of CCTV facial investigation has been tested on the proposed 2.5D technique. Several images taken from camera at different crime scene but related to single stealing case namely “Tepuk bahu” case had been observed to be matched. The suspects (Fig. 6) conducted a serial of stealing crimes at various ATM banks in different states in Malaysia. Three Bolivian suspects conducted the crime by setting a plot with one person observing the pin number of a victim, then another person would steal the victim’s ATM card by bowing after clapping the victim’s shoulder and informing her of her fallen money on the floor. Another suspect would then withdraw the money using stolen ATM card after receiving all details. The source video image has been extracted to tiff image with 480 × 320 image size. We cropped all suspects’ face, resizing them and evaluated by using cross validation approach. We conducted two evaluations based on fitting error and precision recall. The result showed that the proposed 2.5D technique was able to recognize the suspect from CCTV sample.

### VI. CONCLUSION

In this paper, a fully automatic 2.5D facial technique for forensic applications had demonstrated its performance on the new property 2.5D data set as well as Malaysian real case of CCTV facial investigation sample.

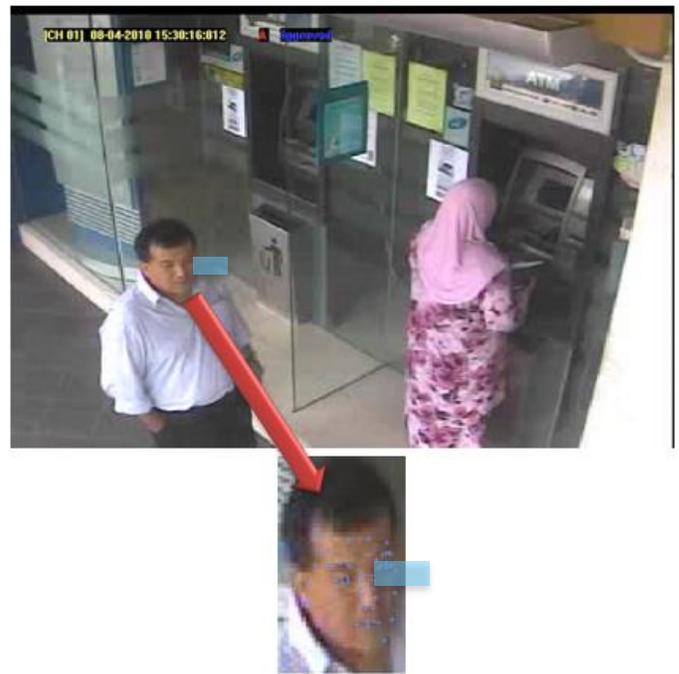


Fig. 6. Examples of (above) life incident dated 8 April 2010 timed 15:30 and (below) extracted face images with annotated landmarks Malaysian real case from CCTV.

A number of contributions were presented while addressing major problems in the area of face recognition. Our contributions include: a modified feature extraction method using AAM based on new algorithm called MRNS (AAM-MRNS), classification method named as Adaptively Accelerated Particle Swarm Optimization based on SVM (AAPSO-SVM) and fully automatic 2.5D facial technique based on AAM-MRNS and AAPSO-SVM methods.

In the experimental results, the comparisons were done with the conventional AAM and adaptive AAM in feature extraction part; and with the PSO-SVM as well as the other techniques for the recognition part. The average error rate for the proposed 2.5D technique based on (AAM-MRNS) on the 2.5D face dataset was 4.038% and 7.02%, 6.02% for conventional AAM and adaptive AAM respectively. On the other hand, the accuracy rate was used to evaluate the recognition performance for the final proposed 2.5D technique. In the recognition stage, the proposed 2.5D technique (AAM.MRNS+AAPSO.SVM) technique was compared with (AAM.ABC+ OPSO.SVM), (AAM.MRNS + OPSO.SVM), (AAM.ABC + AAPSO.SVM), and with PSO-SVM methods. The average accuracy rate for the proposed 2.5D technique on 2.5D face images were 85% and 77% for PSO-SVM; while 75.6, 79, 82.5 for (AAM.ABC+ OPSO.SVM), (AAM.MRNS + OPSO.SVM), (AAM.ABC + AAPSO.SVM) correspondingly. Furthermore, the 2.5D technique was able to recognize the suspect of Malaysian real case of CCTV investigation.

For the future work, the research will look into several fields to ensure the methodology can assist in helping the

current video forensics problem. Further, the research is currently exploring in the Graphics Processing Unit (GPU) to enhance and optimize the algorithm performance.

#### ACKNOWLEDGMENT

This research was initially funded by the Ministry of Education through ERGS/1/20 IISTG/UKM/2/48 (TK) under the title of 2D-3DHybrid Face Matching via Fuzzy Bees Algorithm for Forensic Identification, Universiti Kebangsaan Grant DIP 2015-023 and FRGS/2/2014/ICT02/UKM/02/1 entitled Crime Prediction Using Spatio-Temporal Pattern with Rough Fuzzy Sets Based Ensemble Learning”.

#### REFERENCES

- [1] Daniel B. Garrie, Digital Forensic Evidence in the Courtroom: Understanding Content and Quality, Northwestern Journal of Technology and Intellectual Property, Vol 12, Issue 2, 2014.
- [2] M. D. Kohn, M.M Eloff, J.H.P Eloff, Integrated Digital Forensic Process Model, Computer & Security, Vol 38, pp. 103-115, 2013.
- [3] Anil K. Jain, Arun Ross, “Bridging the Gap: from Biometrics to Forensics”, Philosophical Transactions of The Royal Society, 2015.
- [4] Dong Yi, Zhen Lei, Stan Z. Li, Towards Pose Robust Face Recognition, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3539-3545, 2013.
- [5] Rahul D. Dhotkar, Prakash R. Chandore, Prashant N. Chatur, “Face Recognition techniques and its application”, International Journal of Application in Engineering & Management (IAIEM), Vol 3(2), 2014.
- [6] G. Porter and G. Doran, “An Anatomical and Photographic Technique for Forensic Facial Identification,” Forensic Science International, Vol. 114, pp. 97–105, 2000.
- [7] Kevin J. Chan, Stephen J. Elliott, “Mugshot Compliance for Face Image Quality, International Conference on Computational Science and Computational Intelligence, IEEE, pp. 428-432, 2015.
- [8] Certified Biometric Professional-Module 1, 2012
- [9] Thirimachos Bourlai, Arun Ross, Anil Jain, On Matching Digital Face Images Against Scanned Passport Photos, IEEE International Conference on Biometrics, Identity and Security, 2009.
- [10] Anil K. Jain, Brendan Klare, Usang Park, Face Recognition: Some Challenges in Forensics, 9<sup>th</sup> IEEE International Conference on Automatic Face and Gesture Recognition, 2011.
- [11] Rabia Jafri, Hamid R. Arabnia, A Survey of Face Recognition Techniques, Journal of Information Processing Systems, Vol. 5(2), pp. 41-68, 2009.
- [12] Anil K. Jain, Brendan Klare, Usang Park, Face Matching and Retrieval in Forensics Applications, Multimedia in Forensics, Security and Intelligence, pp. 20-28, 2012.
- [13] Mian, A.S, Bennamoun, M. & Owens, R., An Efficient Multimodal 2D-3D Hybrid Approach to Automatic Face Recognition, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 29(11), pp.1927-1943, 2007.
- [14] Kusuma, G. P. & Chua, C.-S., PCA-based image recombination for multimodal 2D+3D face recognition. *Image and Vision Computing* 29 (5), pp. 306–316, 2011.
- [15] Chang, K., Bowyer, K., & Flynn, P. 2003. Multi-modal 2d and 3d Biometrics for Face Recognition. *Proc. of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures*. France, October, pp. 187-195, 2003.
- [16] Conde, C., A. Serrano, & E. Cabello, MULTIMODAL 2D, 2.5D & 3D FACE VERIFICATION. International Conference on Image Processing, Oct, Atlanta, GA , pp. 2061 – 206, 2006.
- [17] Charoenpong, T., Tokai S. & Hase, H., Facial Expression Recognition from 2.5D Partial Face Data by Using Face Plane. *ECTI Transactions on Electrical Eng., Electronics , and Communications* 8(1), pp. 32-42, 2010.
- [18] Liu, P., Woo, W.L. & Dlay, S.S., One Colored Image Based 2.5d Human Face Reconstruction, 17th European Signal Processing Conference (EUSIPCO), Glasgow, Scotland, August .pp. 2584- 2588, 2009.
- [19] Lee, H.-S. & Kim, D., Tensor-Based AAM with Continuous Variation Estimation: Application to Variation-Robust Face Recognition, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 31(6) pp. 1102 – 1116, 2009.
- [20] Baykasoulu, A., Ozbakir, L. & Tapkan, P. 2007. Artificial Bee Colony Algorithm and Its Application to Generalized Assignment Problem. in *Swarm Intelligence: Focus on Ant and Particle Swarm Optimization*, F. T. Chan and M. K. Tiwari, Eds. Vienna, Austria: Itech Education and Pub., pp. 113–144
- [21] Ma, M., Liang, J., Guo, M., Fan, Y. & Yin, Y. 2011. SAR Image Segmentation Based on Artificial Bee Colony Algorithm. *Applied Soft Computing* 11 (8): 5205–5214
- [22] Jin, W., Zhang, J. & Zhang, X. 2011. Face Recognition Method Based on Support Vector Machine and Particle Swarm Optimization. *Expert Systems with Applications* 38(4): 4390–4393
- [23] Nazri A. Zamani, A. D Zaharudin, Siti Norul Huda Sheikh Abdullah, Md Jan Nordin, Sparse representation super-resolution method for enhancement analysis in video forensics, proceeding of 12th International Conference on Intelligent Systems Design and Applications, pp. 921-926, 2012.
- [24] Liu, X. 2010. Video-based face model fitting using Adaptive Active Appearance Model. *Image and Vision Computing* 28: 1162–1172.

# Efficient Feature Selection for Product Labeling over Unstructured Data

Zeki YETGİN

Computer Engineering Department,  
Mersin University  
Mersin, Turkey

Abdullah ELEWI

Computer Engineering Department,  
Mersin University  
Mersin, Turkey

Furkan GÖZÜKARA

Computer Engineering Department,  
Çukurova University  
Mersin, Turkey

**Abstract**—The paper introduces a novel feature selection algorithm for labeling identical products collected from online web resources. Product labeling is important for clustering similar or same products. Products blindly crawled over the web sources, such as online sellers, have unstructured data due to having features expressed in different representations and formats. Such data result in feature vectors whose representation is unknown and non-uniform in length. Thus, product labeling, as a challenging problem, needs efficient selection of features that best describe the products. In this paper, an efficient feature selection algorithm is proposed for product labeling problem. Hierarchical clustering is used with the state of the art similarity metrics to assess the performance of the proposed algorithm. The results show that the proposed algorithm increases the performance of product labeling significantly. Furthermore, the method can be applied to any clustering algorithm that works on unstructured data.

**Keywords**—Product labeling; product clustering; feature selection; similarity metrics; hierarchical clustering

## I. INTRODUCTION

With recent developments in web technologies, online shopping sites are changing to powerful product search engines to integrate various attractive services for sellers and customers such as product recommendation and comparison systems [1]-[3]. These web platforms can get data from online marketplaces, unify them, and provide e-commerce services using this data for their customers as in online comparison shopping engines [2]-[3].

One general concern in these search platforms is the product labeling problem for clustering identical products [4], generally referred to as record linkage [2]. Usually, customers want to compare the same products from different sellers, e.g. to see their prices. The product labeling requires assigning labels to those products that have identical features. However, one product is commonly described in different ways by different online web sources. Moreover, in order to describe the structure of the product information, each web source should have its own schema. Currently, ontology mapping approaches [5]-[6] are used to unify the product information from various resources. Ontology mapping is the schema matching approach in order for the web sources to learn their structure description of product information. However, a separate ontology should be developed manually for each web source, and importantly the ontology mapping may not provide perfect data collection. That is, the collected product data may

still include unstructured or incomplete features. To address the problem of unstructured or imperfect data, new decision methods, apart from ontology matching approaches, are required in product clustering and labeling.

Formally, product labeling can be considered as a clustering problem to group the identical products into the same category using some similarity metrics. Each product is described by a feature vector and the similarity metrics define the degree of similarity between any pair of feature vectors. If these vectors contain only descriptive and relevant features that contribute much to its identification, the performance of clustering identical products will be improved significantly. However, with unstructured vectors where no vector metadata is available, selecting most descriptive and important features becomes crucial and challenging for product labeling. Thus, product labeling requires efficient feature selection methods to cope with unstructured nature of product data. In this paper, a web crawler is implemented to blindly collect products' features in many categories. Then product labeling is demonstrated using hierarchical clustering algorithms of various types with the proposed feature selection method applied.

The paper is organized as follows: Related works are investigated in the second section. The third section provides the system model and the proposed algorithm. The fourth section demonstrates the experimental results of the proposed methods. Finally, conclusion and future directions are given.

## II. RELATED WORK

Most works related to product clustering usually focus on analysis of customer behaviors [7] to cluster recommended products of interest [1] or analysis of product reviews [8]-[9] to study human opinions about the products features. These works usually use sentiment analyses or opinion mining [9]-[16] where human subjects are involved for assessment of the product features. For example, the authors in [15] cluster the similar features and try to find the correlation between the human opinions and set of features of the products. Commonly, feature selections of the products are also studied with respect to opinion mining or human behaviors. In [9], [12]-[14], information extraction systems are introduced, which extracts fine features with respect to associated opinions. However, none of the works in the literature studied feature selection for product labeling for unstructured web crawled over product dataset.

In literature, the problem of categorizing identical products, referred to here as product labeling, is expressed using different terminologies such as record linkage, entity resolution, duplicate detection [17], clustering of identical products [4], and product normalization [2], [18]. To the best of our knowledge, only few works [2], [4], [19] addressed the record linkage problem for ecommerce products. In [2] the record linkage problem is addressed by using supervised-learning of a similarity function, which is costly and not practical due to continuous need of training. Also, in [4] clustering algorithm is used to label identical ecommerce products where new similarity and performance metrics for clustering of identical products are proposed. Moreover in [19], an incremental hierarchical clustering system for record linkage in ecommerce domain is proposed. Although there are many works related to record linkage, almost none of them consider product labeling or product identification for web crawled products taking feature selection into account.

### III. SYSTEM MODEL

In this section, the proposed feature selection algorithm and its application to hierarchical clustering is described for product labeling problem. Hierarchical clustering is used, as described in [4], to solve the product labeling problem where the feature vectors are formed using the proposed feature selection algorithm. The dataset containing the product features are obtained from [4] where each line represents feature vectors, some samples are shown in Table 1. The proposed method selects important product features and removes the others that don't contribute to identification of the product, which results in final feature vectors.

TABLE I. SAMPLES OF INITIAL FEATURE VECTORS

'intel' 'core' 'i5' '2300' '80ghz' '6mb' 'vga' '1155p'
'bx80623i52300' 'intel' 'lga1155' 'core' 'i5' '2300' '80ghz' '6mb' 'cache'
'intel' 'core' 'i5' '2300' '80' 'ghz' 'lga1155' 'i?lemci'
'intel' 'ci5' '2300' '80ghz' 'mb' 'vga' '1155p' 'core' 'i5' 'i?lemci'
'intel' 'core' 'i5' '2300'

#### A. Proposed Feature Selection Algorithm

The proposed feature selection algorithm has 3 phases. In the first phase, it divides the feature space into overlapping clusters where same vectors might be referred to in different clusters. In the second phase, most informative features of the vectors in each cluster are selected and ordered using a weight function which adopts two criteria: 1) Vector length: The vector with smaller length carries more descriptive information than the longer one. That is, the information load per feature is high. 2) Feature position: Features early positioned in the vector are possibly more informative than the late ones. People usually tend to refer to important features at the earliest while describing products. So at the second phase, features in each vector ordered according to their weights and vector lengths are trimmed to cluster average so that features not contributing to the products identification are eliminated. In the final phase the vectors are just trimmed to target-dimension. The key property of the algorithm is that any modification to

overlapped vectors can be seen by other clusters and this causes information flow among clusters achieving better identification of informative features.

In order to present the proposed algorithm, let's define the following terminology.

$VS = \{ V_i \mid i = 1..N \}$  represents feature vector space where  $V_i$  represents the feature vector of the  $i^{th}$  product among  $N$  products.

$C_i$  represents  $i^{th}$  cluster which contains indices of the vectors in  $VS$  that are similar to  $V_i$  according to similarity metric and the threshold as system parameters, and formulated in (1).

$$C_i = \{ j \in 1..N \mid \text{similarity}(V_i, V_j) \geq \text{threshold} \}$$

$$\text{where } \text{similarity}(V_i, V_j) = \frac{|V_i \cap V_j|}{\min(|V_i|, |V_j|)} \quad (1)$$

$AD_i$  represents the average dimension of the vectors in  $C_i$ , and formulated in (2).

$$AD_i = \left\lfloor \frac{\sum_{k \in C_i} |V_k|}{|C_i|} \right\rfloor \quad (2)$$

$FR(f, C)$  denotes the frequency of feature  $f$  in vectors of cluster  $C$ , and formulated at (3).

$$FR(f, C) = \sum_{k \in C} \begin{cases} 1, & \text{If } f \in V_k \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

$AL(f, C)$  denotes the average length of the vectors in cluster  $C$  that includes feature  $f$  and formulated in (4) and (5). This is used for the first criterion in (7).

$$SL(f, C) = \sum_{k \in C} \begin{cases} |V_k|, & \text{If } f \in V_k \\ 0, & \text{Otherwise} \end{cases} \quad (4)$$

$$AL(f, C) = \frac{SL(f, C)}{FR(f, C)} \quad (5)$$

where  $SL(f, C)$  denotes the summation of the vector lengths in cluster  $C$  that includes feature  $f$ .

$ML(f, C)$  denotes the minimum length of the vectors in cluster  $C$  that includes feature  $f$  and is formulated in (6). This is used for the first criterion in (7).

$$ML(f, C) = \min \{ |V_k| \mid k \in C \text{ and } f \in V_k \} \quad (6)$$

$WL(f, C)$  denotes the weight of the feature  $f$  in  $C$  according to the first criterion (vector length criterion), and is formulated in (7). Once a feature is found in a minimum length vector ( $ML$ ), this should be highly emphasized (most informative state). The more the feature  $f$  repeats in cluster  $C$ , the more the weight approaches to its most informative value ( $ML$ ). The less the feature  $f$  repeats in cluster  $C$ , the more the weight approaches to cluster average ( $AL$ ).

$$WL(f, C) = ML(f, C) + \frac{AL(f, C) - ML(f, C)}{FR(f, C)^2} \quad (7)$$

$AP(f, C)$  denotes the average position of the feature  $f$  in vectors of cluster  $C$ , and is formulated in (8). This is used for the second criterion at (9).

$$AP(f, C) = \frac{\sum_{k \in C} pos(f, V_k)}{FR(f, C)} \quad (8)$$

where  $pos(f, V)$  returns the position (index) of feature  $f$  in vector  $V$ , or return zero when  $f$  is absent in  $V$ .

$WP(f, C)$  denotes the weight of the feature  $f$  in  $C$  according to the second criterion (feature position criterion), and is formulated at (9) to give a little bit more significance to the second criterion than the first criterion and also to escape from zero division when they are combined in (10).

$$WP(f, C) = \log_e^2(AP(f, C)) + 1 \quad (9)$$

$Weight(f, C)$  denotes the overall weight of the feature  $f$  in  $C$  according to our two criteria for vector length and feature position, and is formulated in (10).

$$Weight(f, C) = \frac{1}{WP(f, C) * WL(f, C)} \quad (10)$$

The proposed algorithm sorts the vector  $V$  in each cluster  $C$  in descending order of  $Weight(f, C)$  where  $f$  scans the features of vector  $V$  and then attempts to reduce the dimension of  $V$  to cluster average if possible or to target dimension otherwise.

The proposed feature selection algorithm is shown below:

---

### Feature Selection Algorithm

---

**Input:** initial vector space (VS), threshold, target\_dimension

**Output:** resulted vector space

**For**  $i=1$  to  $N$

$C_i \leftarrow$  create cluster  $C_i$  according to (1)

**For**  $i=1$  to  $N$

$CS \leftarrow$  extract all distinct features in vectors of  $C_i$

$AD_i \leftarrow$  compute average dimension of  $C_i$  according to (2)

**If**  $AD_i < target\_dimension$  Then  $AD_i = target\_dimension$

**For** each feature  $f$  in  $CS$

$Weight(f, C_i) \leftarrow$  calculate weight according to (10)

**For** each  $k$  in  $C_i$

$V_k \leftarrow$  sort  $V_k$  in descending order of  $Weight(f, C_i)$   
where  $f$  scans features of  $V_k$

**If**  $|V_k| \geq AD_i$  Then

$V_k \leftarrow$  select the first  $AD_i$  number of features of  $V_k$

**For**  $i=1$  to  $N$

**If**  $|V_i| \geq target\_dimension$  Then

$V_i \leftarrow$  select the first  $target\_dimension$  number  
of features of  $V_i$

---

### B. Clustering Model and Performance Metrics

In this section, we describe the clustering model for product labeling on which we applied our proposed feature selection algorithm. We have considered hierarchical types of clustering to demonstrate how our feature selection algorithm achieves well in labeling identical products. Hierarchical clustering algorithms use similarity metrics and linkage metrics. The similarity metric determines the degree of similarity between any pair of vectors. This paper considers the similarity metrics recently proposed in [4] for non-uniform feature vectors. So we demonstrate the performance of our algorithm using four similarity metrics proposed in [4], namely minimally-normalized intersection similarity (MNI), globally-normalized locally weighted similarity (GNLW), globally-normalized-indexed similarity (GNI) and globally-normalized globally weighted similarity (GNGW).

Hierarchical clustering algorithm needs also linkage metrics, which use the underlying similarity metrics to measure the similarity among sub clusters to merge them to form a bigger cluster in the hierarchy. With the four similarity metrics mentioned earlier and the five following linkage metrics, different clustering algorithms are considered in the paper. These linkage metrics are single (nearest distance), complete (furthest distance), average (unweighted average distance), weighted (weighted average distance), and median (weighted center of mass distance) linkage clustering. The single method considers the smallest distance between the points in two clusters for the decision of merging whereas the complete method considers the furthest distance between two clusters. The other methods behaves similarly as their names indicate.

We used the performance measurement metrics recently proposed in [4] where three metrics, namely False-Positive (FP), False-Negative (FN), and Total Error (TE) are defined to assess the performance of the product labeling when the original cluster labels are available.

The metric considers a space of product pairs where the labels of pairs that are detected as identical or non-identical by the algorithm are compared with the original true labels that are priori available. The metrics are defined in [4] as follows:

1) False-Negative (FN) indicates the number of the product pairs that are classified as non-identical by the algorithm, although they are actually identical.

2) False-Positive (FP) indicates the number of the product pairs that are classified as identical by the algorithm, although they are actually non-identical.

3) Total Error (TE) indicates the total number of decision errors caused by either False-Negatives or False-Positives.

## IV. RESULTS AND DISCUSSION

### A. Datasets

The dataset is acquired as two text files from [4] where one file is for product description (product-file) and the other is for error-free product labels (label-file).

Each line of the product-file describes one product without using any predefined structure (see Table 1) and the corresponding product label is given at the same line of the label-file. The dataset includes one thousand products, selected randomly from one million products, which are blindly crawled from 20 most popular online Turkish sellers (see Table 2). The products are from many different categories including computers, home appliance, smart phones, etc. So each line of the product-file serves as initial feature vector and our proposed feature selection algorithm generates final feature vectors that are input to the hierarchical clustering algorithms.

TABLE. II. SOME ONLINE SELLERS THAT ARE CRAWLED TO COLLECT PRODUCT INFORMATION [4]

Web Site	Number of Products Crawled	Number of Pages Crawled
hepsiburada.com	177.310	313.946
hizlial.com	84.046	166.197
webdenal.com	69.979	121.853
ereyon.com.tr	68.960	92.076
pratikev.com	63.170	69.275
netsiparis.com	40.525	59.294

### B. Experimental Results

In this section, the results of the proposed feature selection algorithm are shown and the performance of its success in product labeling problem is demonstrated using the hierarchical clustering algorithms where four similarity metrics (MNI, GNLW, GNI, GNGW) and five linkage metrics (single, complete, average, weighted, median) are used. Exhaustive experiments are conducted and optimum values of performance metrics are provided in tables. All the algorithms in this paper are implemented using MATLAB®. The results of the proposed method are given in Tables 3 to 8 for different target dimensions where the optimum threshold is given at the table title. These tables only provide some example sets for demonstration purposes.

The results in Tables 4, 6 and 8 show that the algorithm selects and orders the informative features successfully in general. For example, the feature 'st1500dl003' as an informative word isn't initially contained in top 3 features in Table 3; however our algorithm succeeds in bringing it to the second position as shown in Table 4. Similarly, in Table 5, the most descriptive features 'kingston', '16', 'gb', 'dtig3' are successfully selected in top 4 features in Table 6. Similar success can be seen in Table 8 where the most 6 informative features are selected. These results are qualitative examples where one must further analyze the selected features.

In order to demonstrate the success of the proposed feature selection algorithm quantitatively we can compare the product labeling performance of the hierarchical clustering algorithm with and without applying the proposed feature selection. The evaluation of each clustering experiment is done by the performance metrics proposed in [4].

TABLE. III. ORIGINAL SET 1

'seagate' 'barracuda' 'green' '5tb' '5900rpm' 'sata' 'gb' 'sn' 'ncq' 'sabit' 'disk' 'st1500dl003'
'seagate' 'barracuda' 'green' 'st1500dl003' '5tb' 'sata' 'sabit' 'disk'
'seagate' '5tb' '6gb' 'barracuda' 'green' 'st1500dl003'
'seagate' 'st1500dl003' '5tb' '5900rpm' '64mb' 'sata3' '6gb' 'barracuda' 'green'

TABLE. IV. RESULTED SET 1 WITH THRESHOLD=0.74, TARGET\_DIMENSION = 3

'seagate' 'st1500dl003' '5tb'
'seagate' 'st1500dl003' '5tb'
'seagate' 'st1500dl003' '5tb'
'seagate' 'st1500dl003' '5tb'

TABLE. V. ORIGINAL SET 2

'kingston' '16' 'gb' 'usb' 'memory' 'dtig3' '16gb'
'16' 'gb' 'usb' 'dtig3' 'kingston'
'kingston' 'datatraveler' 'g3' '16' 'gb' 'usb' 'bellek' 'dtig3' '16gbz'
'kingston' 'dtig3' '16gbz' '16gb' 'datatraveler' 'g3' 'usb' 'flash' 'disk'

TABLE. VI. RESULTED SET 2 WITH THRESHOLD=0.72, TARGET\_DIMENSION = 4

'kingston' '16' 'gb' 'dtig3'
'kingston' '16' 'gb' 'dtig3'
'kingston' '16' 'dtig3' 'usb'
'kingston' 'dtig3' '16gb' 'usb'

TABLE. VII. ORIGINAL SET 3

'samsung' 'intel' 'atom' 'n570' '66ghz' '2gb' '320gb' '10' 'beyaz' 'netbook' 'n150' 'jp0xtr'
'samsung' 'np' 'n150' 'jp0xtr' 'beyaz' 'atom' 'n570' '2gb' '320gb' 'payla?ml?' 'vga' 'gma3150' '10' 'win' 'starter'
'samsung' 'n150' 'jp0xtr' 'atom' 'n570' '66ghz' '2gb' '320gb' '10' 'netbook' 'w7s' 'beyaz'
'samsung' '320gb' 'beyaz' 'n570' 'netbook' '10' '2gb' 'jp0xtr' 'n150'

TABLE. VIII. RESULTED SET 3 WITH THRESHOLD=0.61, TARGET\_DIMENSION = 6

'samsung' 'n150' 'jp0xtr' 'n570' '2gb' '320gb'
'samsung' 'n150' 'jp0xtr' 'np' 'n570' '2gb'
'samsung' 'n150' 'jp0xtr' 'n570' '2gb' '320gb'
'samsung' 'n150' 'jp0xtr' 'n570' '2gb' '320gb'

TABLE. IX. PRODUCT LABELING PERFORMANCE WITHOUT THE PROPOSED FEATURE SELECTION (LEGACY METHOD)

Similarity	MNI			GNLW			GNI			GNGW		
	FN	FP	TE									
single	0.29	0.26	<b>0.28</b>	0.26	0.09	<b>0.19</b>	0.39	0.42	0.40	0.21	0.18	<b>0.20</b>
complete	0.58	0.00	0.44	0.57	0.00	0.43	0.35	0.16	0.27	0.54	0.03	0.41
average	0.48	0.11	0.35	0.43	0.20	0.35	0.23	0.20	<b>0.22</b>	0.40	0.24	0.33
weighted	0.52	0.05	0.38	0.46	0.13	0.34	0.25	0.22	0.24	0.37	0.13	0.28
median	0.54	0.16	0.43	0.40	0.40	0.40	0.55	0.13	0.43	0.35	0.36	0.35

TABLE. X. PRODUCT LABELING PERFORMANCE WITH THE PROPOSED FEATURE SELECTION WITH TARGET\_DIMENSION=3

Similarity	MNI			GNLW			GNI			GNGW		
	FN	FP	TE									
single	0.27	0.02	<b>0.17</b>	0.23	0.03	<b>0.15</b>	0.13	0.08	<b>0.10</b>	0.12	0.10	<b>0.11</b>
complete	0.37	0.07	0.26	0.36	0.08	0.26	0.15	0.07	0.11	0.29	0.17	0.23
average	0.31	0.09	0.27	0.33	0.09	0.23	0.11	0.16	0.13	0.20	0.12	0.16
weighted	0.36	0.07	0.26	0.36	0.07	0.26	0.11	0.13	0.12	0.21	0.11	0.17
median	0.33	0.06	0.23	0.32	0.07	0.22	0.09	0.19	0.14	0.27	0.07	0.19

The results are given in Tables 9 and 10 for target dimension = 3 and summarized in Fig. 1 to 3 for target dimension of 3, 4 and 6. Tables 9 and 10 show the performance of our feature selection algorithm in product labeling is almost 50% better than the legacy approach when the best TE of the legacy (0.19) and the best of the proposed method (0.10) are compared. Furthermore, our feature selection algorithm performs better than the legacy one at all linkage and similarity metrics when best TEs of each metric, denoted in bold in Tables 9 and 10, are compared.

Generally, single linkage performs better than the other linkages. Thus, we analyzed the single linkage further for other dimensions (dimension = 4 and 6) and the results are summarized in Fig. 1 to 3. Figures show that our proposed method improves the success of the product labeling for all dimensions. The success of the product labeling with the proposed feature selection is better for smaller dimensions. That is, the proposed method successfully selects the informative features. As the dimension increases the performance of all methods gets worse due to the fact that resulting feature vectors tend towards original feature vectors. That is, the more features are selected, the more unnecessary details are taken into account. As the dimension increases the GNGW performs better than the other metrics whereas GNI gets worse. For small dimensions GNI and GNGW are preferable. Thus, GNGW provides better performance for the average dimension.

Depending on the problem domain, some linkage or similarity metrics could be preferable. A hierarchical clustering algorithm with a particular linkage and a similarity metric defines the behavior of the algorithm. For product labeling considered here the results show that single linkage

is favorable. Similarly, tolerance to decision errors is also dependent on the problem domain. For instance, some problems may have tolerance to the FN errors but not to the FPs. The results show that product labeling with the proposed feature selection method has more tolerance to FPs in general. That is, FNs errors contribute more to the total decision errors than FPs. The only exception is the GNGW where the total errors are mainly caused by both FNs and FPs errors.

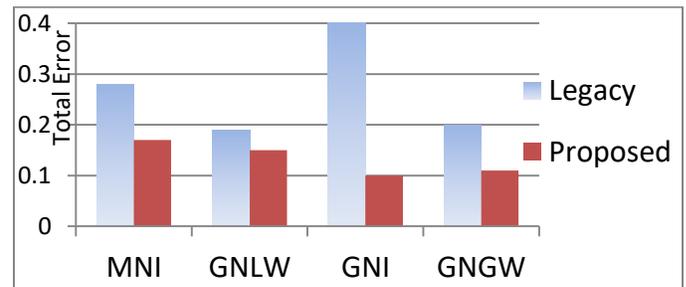


Fig. 1. Performance comparison in terms of TEs for the legacy and the proposed methods with single linkage and target\_dimension = 3.

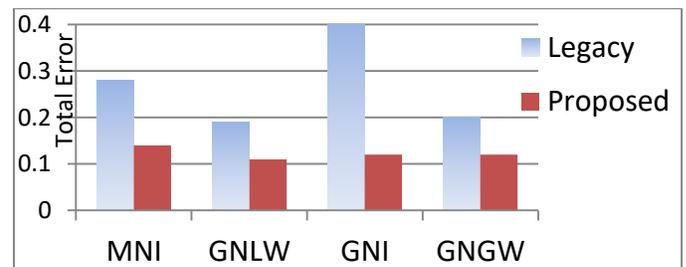


Fig. 2. Performance comparison in terms of TEs for the legacy and the proposed methods with single linkage and target\_dimension = 4.

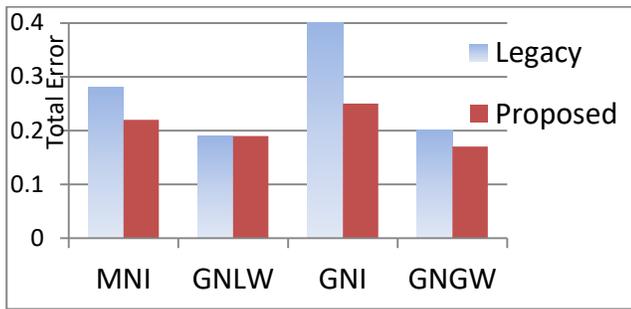


Fig. 3. Performance comparison in terms of TEs for the legacy and the proposed methods with single linkage and target\_dimension = 6.

The proposed feature selection algorithm improves the performance of the product labeling by reducing the total error for all possible cases of linkage, similarity and dimension. The proposed algorithm can also be used in product clustering generally to enhance the clustering performance. However, the algorithm is tested on only available dataset due to absence of such datasets in the internet. Further study is still needed to test and improve the algorithm according to its success in new datasets.

## V. CONCLUSION

A new feature selection algorithm is introduced for unstructured product data. The performance of the proposed algorithm is demonstrated by applying it into the product labeling problem where our algorithm selects most informative features before product labeling. The proposed algorithm can be used in feature selection phase of any product clustering algorithms. The performance comparison of the proposed algorithm is done by the state of the art performance metrics recently developed for the product labeling problem. The results show that the proposed algorithm provides almost 50% better performance in term of total error when compared to the legacy approach. The proposed algorithm successfully selects the brand names and major descriptive words such as category and model names. However, future works are needed to test the success of the feature selection algorithm on different datasets and improve the algorithm to cope with imperfect nature of data, such as using natural language processing, which is not addressed here.

## REFERENCES

[1] L. S. Chen, F. H. Hsu, M. C. Chen, and Y. C. Hsu, "Developing recommender systems with the consideration of product profitability for sellers," *Information Sciences* 2008; 178: 1032–1048.  
[2] M. Bilenkoil, S. Basil, and M. Sahami. "Adaptive product normalization: Using online learning for record linkage in comparison shopping" Fifth IEEE International Conference on Data Mining, IEEE, 2005.

[3] R. B. Doorenbos, O. Etzioni, and D. S. Weld. "A scalable comparison-shopping agent for the world-wide web." *Proceedings of the first international conference on Autonomous agents*. ACM, 1997.  
[4] Z. Yetgin, and F. Gözükar. "New metrics for clustering of identical products over imperfect data." *Turkish Journal of Electrical Engineering & Computer Sciences* 23.4 (2015): 1195-1208.  
[5] S. Park, W. Kim, S. Lee, and S. Bang, "Product matching through ontology mapping in comparison shopping", In: *Proceedings of IJWAS*; 4–6 December 2006; Yogyakarta, Indonesia: ACS. pp. 39–49.  
[6] M. Walther, N. Jackel, D. Schuster, and A. Schill, "Enabling product comparisons on unstructured information using ontology matching", *Advances in Intelligent and Soft Computing* 2011; 86: 183–193.  
[7] B. Galitsky, and J. L. Rosa, "Concept-based learning of human behavior for customer relationship management", *Information Sciences* 2011; 181: 2016–2035.  
[8] H. Almagrabi, A. Malibari, and J. McNaught. "A Survey of Quality Prediction of Product Reviews", *International Journal of Advanced Computer Science & Applications (IJACSA)* 1.6 (2015): 49-58.  
[9] H. Jeong, D. Shin, and J. Choi, "Ferom: feature extraction and refinement for opinion mining", *ETRI Journal* 2011; 33: 720–730.  
[10] Y. Cao, P. Zhang, and A. Xiong. "Sentiment analysis based on expanded aspect and polarity-ambiguous word lexicon." *International Journal of Advanced Computer Science and Applications (IJACSA)* 6.2 (2015): 97-103.  
[11] M. Al-Ayyoub, A. Nuseir, G.Kanaan, and R. Al-Shalabi, "Hierarchical classifiers for multi-way sentiment analysis of arabic reviews." *International Journal of Advanced Computer Science and Applications (IJACSA)* 7.2 (2016): 531-539.  
[12] D. Marcu, and A. Popescu, "Extracting product features and opinions from reviews". In: *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*; October 2005; Stroudsburg, USA: ACL. pp. 339–346.  
[13] Z. Shu, J. Wenjie, X. Yingju, M. Yao, and Y. Hao, "Morpheme-based product features categorization in chinese reviews mining", In: *Proceedings of the 6th International Conference on Advanced Information Management and Service*; December 2010; Seoul, China: IEEE. pp. 324–329.  
[14] G. Somprasertsri, and P. Lalitrojwong, "A maximum entropy model for product feature extraction in online customer reviews". In: *Proceedings of IEEE Conference on Cybernetics and Intelligent Systems*; July 2008; Las Vegas, USA: IEEE. pp. 575–580.  
[15] Z. Zhongwu, L. Bing, X. Hua, and J. Peifa, "Clustering product features for opinion mining". In: *Proceedings of the 4th ACM International Conference on Web Search and Data Mining*; February 2011; Hong Kong, China: ACM. pp. 347–354.  
[16] S. S. Sadidpour, et al. "Context-Sensitive Opinion Mining using Polarity Patterns" *International Journal of Advanced Computer Science and Applications (IJACSA)* 7.9 (2016): 145-150.  
[17] P. Christen, *Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection*. Springer Science & Business Media, 2012.  
[18] S. Basu, M. Bilenko, and M. Sahami. "Method and system to produce and train composite similarity functions for product normalization." U.S. Patent No. 7,702,631. 20 Apr. 2010.  
[19] F. Gözükar, and S. A. Özel, "An Incremental Hierarchical Clustering System for Record Linkage in Ecommerce Domain", unpublished.

# Intelligent System for Detection of Micro-Calcification in Breast Cancer

M. Abdul Rehman, Jamil Ahmed, Ahmed Waqas, Ajmal Sawand

Department of Computer Science, Sukkur IBA University  
Sukkur, Pakistan

**Abstract**—Recently; medical image mining has become one of the well-recognized research area(s) of machine learning and artificial intelligence techniques have been vastly used in various computer added diagnostic systems. Specifically; breast cancer classification problem is considered as one of the most significant problems. For instance, complex, diverse and heterogamous malignant features of micro-calcification in DICOM (Digital Communication in Medicine) images of mammography are very difficult to classify because the persistence of noise in mammogram images creates lots of confusions for doctors. In order to reduce the chances of misdiagnosis and to discernment the difference between malignant and benign lesions of micro-calcification this paper proposes a system so called “Intelligent System For Detection of Micro-Calcification in Breast Cancer” by considering all above stated problems. Overall our system comprises over three main stages. In first stage, adaptive threshold algorithm is used to reduce the noise, and canny edge detection algorithm is used to detect the edges of every macro or micro classification. In second stage, deginated as feature selection is done by using auto-crop algorithm, which crops all types of calcifications and lesions by proposed algorithm so called CFEDNN (Calcification Feature Extraction Deep Neural Networks) which is designed to avoid the manual ROIs (Region of Interest). Decision model is constructed by using DNN (Deep Neural Networks) and the best classification accuracy is measured as 95.6%.

**Keywords**—Medical image mining; machine learning; feature extraction; classification; Digital Communication in Medicine (DICOM)

## I. INTRODUCTION

Recently artificial intelligence (AI) techniques have been frequently applied in healthcare industry and computer added diagnostics systems have been witnessed with significant impact to mold the traditional procedures into computerized DSS (decision support systems) for diagnosis and prognosis of various diseases such as colorectal cancer, lung cancer, breast cancer and so on. According to recent statics [1]-[3] breast cancer accounts for 1.7 million yearly deaths in world population and many people are badly suffering from such type of cancers. Breast cancer is one of the common diseases found in female gender but male gender is also facing casualties related to the breast cancer. Traditionally mammography [4]-[6] is one of the non-invasive techniques used to diagnose the breast cancer but there are some limitations associated with X-Ray based technologies. It is really challenging to differentiate between the malignant and non-malignant behaviors from the film base masses, lesions, micro and macro calcifications of human breast tissues due to

some cultural, regional and socio-economic problems which are likely to be found to form the dense and tender tissues because race, gender, occupation, geographic conditions, social system and other contributing factors create [7], [8] lots of variations in the composition of human breast tissues. On the other side there are lots of limitations associated with mammogram image accusation process due to improper acquisition processes caused by unavoidable technological noise seen in medical images [9], [10]. This create lots of confusions for radiologist to interpret the particular type of malignant behaviors and sometime ultrasound guided biopsies are recommended for deep analysis of the disease and results may reveal no malignancy. In recent past some of the related works [11], [13], [15] have been seen and these approaches have attempted to resolve the classification problem of mammograms but some of approaches consider the malignant masses accumulatively set of pixels, since detection of micro-calcification in massive masses may enhance the diagnostic accuracy of breast cancer. For example micro-calcifications may exists between the dense or soft margins of breast but due to diffuse shapes, sizes and malfunctioning of mammogram x-ray technology, there are several chances of miss-diagnosis because fibro adenoma cancer persists with complex patterns of massive masses along with set of calcifications. In-order to reduce the chances of misdiagnosis and to identify the difference between malignant and benign lesions of connected micro-calcifications, this paper proposes a system so called “Intelligent System for Detection of Micro-Calcification in Breast Cancer” by considering all above stated problems. Overall proposed system comprises over three main stages. In first stage noise reduction techniques are applied and each macro or micro calcification is detected by using canny edge detection algorithm. In second stage feature selection is done by using proposed algorithm so called CFEDNNs (Calcification Feature Extraction Deep Neural Networks) which is designed to avoid the manual ROIs (Region of Interest). Classification model is built by using the DNNs (Deep Neural Networks) and best epochs were measured as 95.60%.

The paper is organized in several sections. Firstly introduction is presented, secondly literature review is described, thirdly methodology is explained, fourthly results are shown and finally conclusion & discussion with future dimensions is described.

## II. RELATED WORKS

Basically this paper proposes an approach, which deals as predictive modeling in the domain of machine learning and

proposed system consider the detection of malignant behaviors of breast cancer masses, lesions and macro or micro classifications. Some of the related works have been seen in recent past are cited as below.

A system [11] was proposed to detect the abnormalities of breast cancer using mammograms. AI base Swarm Optimized Wavelet Neural Network algorithms were used and the reported classification accuracy is 92.10%. This paper proposes a system to classify the malignant and benign behaviors by selecting the automated ROIs (Region of Interest) features, which helps to avoid the manual feature selection from breast images.

A system [12] was proposed to classify malignant and benign images of mammograms. The extracted statistical features were classified by using ANN (Artificial Neural Networks) and reported classification accuracy of the system is 94%. Proposed algorithm so called CFEDNN to select the features from several parts of breast neighbored regions such as axilla and breast regions.

A comparison [13] of various machine learning techniques (SVM, AdaBoost) was proposed for classification of breast

cancer. Reported best accuracy for AdaBoost measured as 87.42%. Since the proposed system of this paper detects micro calcifications in the environment of complex and diffuse shape of masses at any sight of mammogram images.

A comparison [14] different machine learning techniques was proposed and best accuracy was measured 95.34% for breast cancer detection using cellular tissues RBF. ConvNet layers provide better feature extraction because final layer is fully connected with all previous layers.

A system [15] for diagnosis of breast cancer was proposed by using machine learning techniques. The reported classification accuracy using SVM was measured as 95.23%. Proposed approach of this paper offers contributions from two perspectives. Firstly from the image preprocessing perspective, in which an algorithm is proposed to reduce the noise by considering unnecessary noise and selects the most significant set of regions such as margins of axilla and breast along with the malignant behaviors consideration at pixel levels. Secondly building of ConvNet layers are for mammography lesions and classification of malignant/benign behaviors. The classification accuracy of the system was measured about 95.60%.

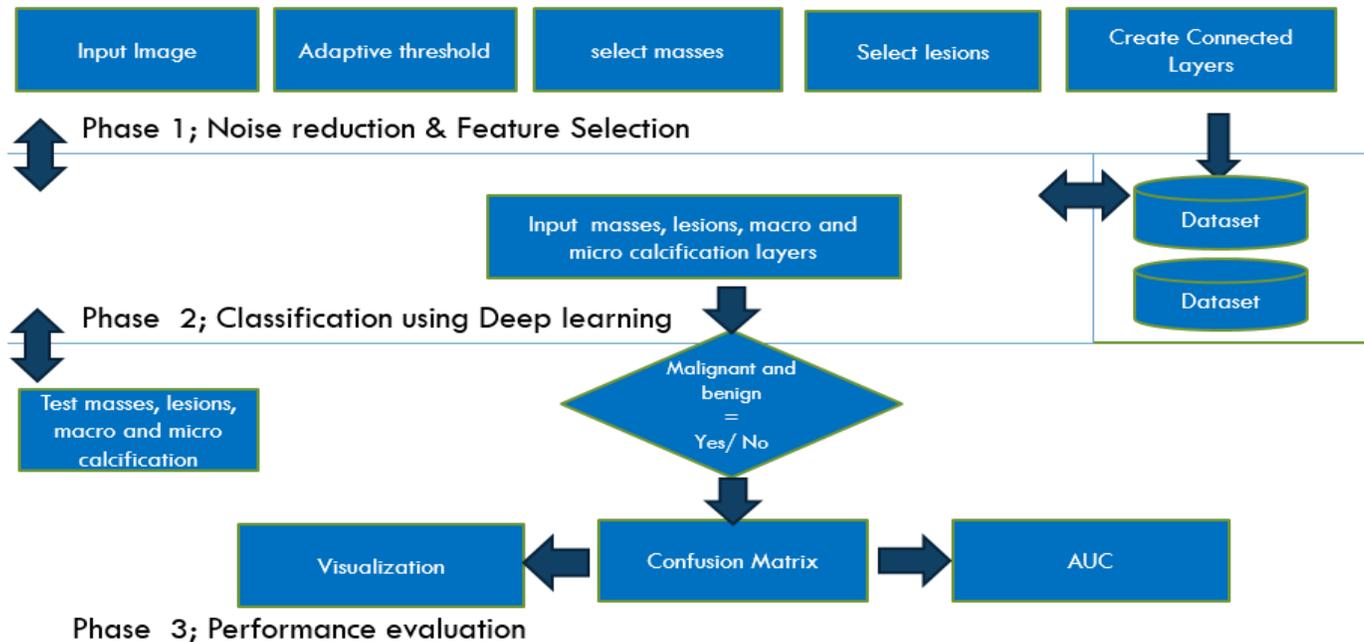


Fig. 1. Intelligent system for detection of micro-calcification in breast cancer workflow.

### III. METHODOLOGY

Basically proposed approach of this paper falls into the category of predictive modeling and machine learning techniques have been used to construct the decision model based upon the deep learning techniques as per following systematic procedure (Fig. 1). A mammogram consists upon thickly populated set of pixels. This paper presents novel algorithm so called CFEDNN (Calcification Feature Extraction Deep Neural Networks) which deals with the object detection, feature selection and classification. Since the appropriate object detection and feature selection is highly

desirable for good results to solve the classification of DICOM (Digital Communication in Medicine) of mammography. The proposed system is divided into three major phases and each phase is assigned interconnected tasks. The first phase deals with the data preparation such as noise reduction and feature selection, whereas second phase is assigned the task to build the classification model where associated quantities selected from the related pixels based upon information in the form of tensor data type. The classification model has to classify the micro and macro classifications to assist the doctors during the diagnostic phase.

Algorithm 1: (CFEDNN)

Calcification\_Feature\_Extraction\_Deep\_Neural\_Networks

Input: Mammogram image as D

Output: Axilla margins, breast tissues and Class Label Malignant =  
Yes/No

$D \leftarrow HCV\ Breast_{Tissues} components\ for\ Binarization\ B$

Visit = each pixel as  $q_n(x_i, y_i)$

for each  $\sigma^2 w(t) \leftarrow q1(t)\sigma_1^2 + q2(t)\sigma_2^2(t)$

$$\sigma_1^2(t) \leftarrow \sum_{i=1}^i [i - \mu_1(t)]^2 \frac{IP(i)}{q_1(t)} \wedge \mu_1(t) = \sum_{i=t+1}^i [i - \mu_1(t)]^2 \frac{P(i)}{q_2(t)}$$

for each  $P_i \in q_i(t)$  do

$$K \leftarrow g(x, y)^{q1(t)} \leq 0_n(k) = 1$$

if  $g(x, y)^{q1(t)} = 1$

Count  $\leftarrow p_i 1 + + ||axilla\ regions||$

if  $g(x, y) = 0$

Count  $\leftarrow p_i 0 + + ||breast\ regions||$

end if

Return  $\leftarrow Breast_{Tissues}$

create ConvNet Layers  $\leftarrow g_h(\tau) = h_b(x) * g(x)$

for each  $g_a(\tau) \in T$  do

$$h_a(x) = \sigma(\theta^T) = + \frac{1}{1 + \exp(-\theta^T)}$$

Where  $a(z) = 1/(1 + \exp(-z))$   
end if

### Stage 1: Noise Reduction and Feature Selection:

#### A. Calcification feature extraction Deep Neural Networks algorithm:

In DICOM (Digital communication in medicine) [16] mammogram images have different contrasts and entropies at each set of pixels in complex, dense and diffuse set of pixels because mostly in all types of breast tissues are found with varying behaviors from many perspectives in terms of shapes, sizes and other properties (such as malignant masses, lesions, macro and micro calcifications). Thus these types of circumstances produce lots of confusions for the radiologists during the interpretation of mammograms. Proposed system comprises upon the main three stages. In first phase of image preprocessing the noise reduction techniques are used, in second phase a decision model is constructed. In third stage performance evaluation and result visualization is shown. Let's consider a dataset D of mammogram images and the digital information is scattered between  $(x_i, y_i)$  spatial locations. The first object of our algorithm CFEDNN (Calcification feature extraction + deep neural networks) is to classify the breast cancer using mammogram images. Adaptive threshold algorithm [17] is used to reduce the noise. The technique calculates accumulated values of neighborhood pixels that represent the small set of objects in medical images by considering mean values of concerned regions and the results of segmentation can be transformed into better illumination despite of varying requirements of different thresholds prerequisite for dissimilar regions of image designated as objects such like malignant behaviors. Since Otsu's algorithm tries to find a threshold value (t) which minimizes the weighted within-class variance given by the relation in a medical image.

The second objective of proposed algorithm is to avoid unnecessary information from the mammograms and to select the most significant feature without using manual ROIs (region of interest) [18]. ROI is a method to select the different regions of image to measure the entropies and other operations. This technique is also compatible to manual

cropping of various objects from image. The proposed algorithm partition the image into three regions, in first step it selects the regions of axilla, in second step different regions of breast masses are detected and finally irrelevant regions such as noise is determined and omitted (Fig. 2). Proposed system creates the three layers using ConvNet based features [19]-[22] as observation collected from selected regions. Let's consider a dataset D of mammogram consisting upon the MXN pixels scattered over the size of in  $(x_i, y_i)$  and the regions of interest persist as  $q_n(x_i, y_i)$  (Fig. 2). These regions  $q_n$  are also often known as several number of image partitions based upon the clustering set of homogenous neighborhood pixels. In human breast there are two partitions having most significant importance number one any existence of nodular quantity or breast is affected due to involvement of calcifications.

#### B. ConvNet Architectures

Proposed preprocessing algorithm selects automated regions as defined in Section 3. This paper uses Deep Neural Networks and builds the dataset with three common layers, in first two layers CONV and POOL layer are created as max pool where each pixel value is represented into tensor data type. In third layer which is fully-connected layer of related pixel representation and a doctor defined class label attribute for malignant and benign classes is included RELU activation function into the training datasets of mammograms as layer.

#### C. Pooling Layer

The successive Conv Layer architecture includes a pooling layer from time to time. The function is responsible increasingly to reduce the spatial size of image pixel interpretation. It not only controls the overfitting of parameters but also regulates computation throughout the computational activities of related layers and incorporates the in-depth inputs received from the pooling layer operations by applying consideration to the set of depth slice observations during the interpretation of image interpreted inputs in Max operations in such way that pooling filters consisting upon the number of max over 4. Since the depth height and width can be reduced up to 75% and builds small regions comprising upon the 2 x 2 at the low levels of depth since these depth features remain with constant quantities.

#### D. Fully-connected layer

In fully connected layer, the last layer (class label attribute) is final output layer which fully connected set of observations related to particular class along with the consideration of biased and weighted values of all previous activated connections exists between the different layers.

### Stage 2: Classification using Deep Learning:

The DNN (Deep Neural Network) uses weighted matrix W and bias vector b and the inputs of set of observation which collects inter connected input quantities consisting upon the image features as neurons. Hidden layers of neurons depend upon the complexity of problem. Weighted and bias values has to construct hyperplane to decrement the behaviors of image data by considering vector  $x \in class\ i$  and variable Y is a stochastic variable.

ReLU (Rectified linear units) layer is used as activation function  $f(x) = \max(0, x)$  and receptive fields of Convolution layer promotes to increase the properties related to nonlinear decision by considering no effect to the above stated layer. Since some other functions may also be used to enhance into the incremental support for nonlinearity. Hyperbolic tangent  $f(x) = \tanh(x)$ ,  $f(x) = |\tanh(x)|$  and to boost the training phase of neural network the sigmoid function  $f(x) = (1 + e^{-x})^{-1}$  offers very nice services at ReLU layer. Since the *softmax* loss can be used to classify the k classes by using the sigmoid cross entropy loss by considering the prediction of class label k as independent probability based quantities as [0,1] and to regress the real label as Euclidean loss  $[-\infty, \infty]$ .

Deep Neural Networks performs well by using the standard backpropagation algorithm and can be discriminated as nested discrete connected networks. The efficiency could also be find by using spars networks. Stochastic gradient descent can be expressed in weighted backpropagation by this equation  $w_{ij}(t + 1) = w_{ij}(t) + n \frac{\partial C}{\partial w_{ij}} + \epsilon(t)$  where hidden layers of neural network perform like human brain neurons. Since learning rate is represented by cost function designated as C and  $\epsilon(t)$  term is derived as scholastic theorem. The reinforcements refers to the choice of cost functions where

supervised or unsupervised learning is applied specially is multiclass problem can be learned with the functions of softmax and cross entropies where  $P_j = \frac{\exp(z_j)}{\sum_k \exp z_k}$  is used to denote the representation of the class probability as output by considering the quantities of  $x_j$  and  $z_k$ . Since the cross entropies could be interpreted as  $C = -\sum_j d_j \log(p_j)$ . A binary mask can be bounded with deep neural networks based regression to increase the localization precision and to learn the features by using regression because some of the layers are convolutional, pooling and fully connected where each object has to be reflected linear unit for transformation perspectives of nonlinear as well.

**Stage 3: Performance Evaluation and result visualization:**

Total number of 1273 images (Table 1) of mammogram were used out of these 362 images were belonging to the category of malignant and 911 were belonging to the benign class. The 327 true positive and 35 true negative and 890 benign number of images were classified as false negative and 21 false positive. The comparison with literature is presented in (Table 2) and the measured classification accuracy (Table 3) is about 95.60%.

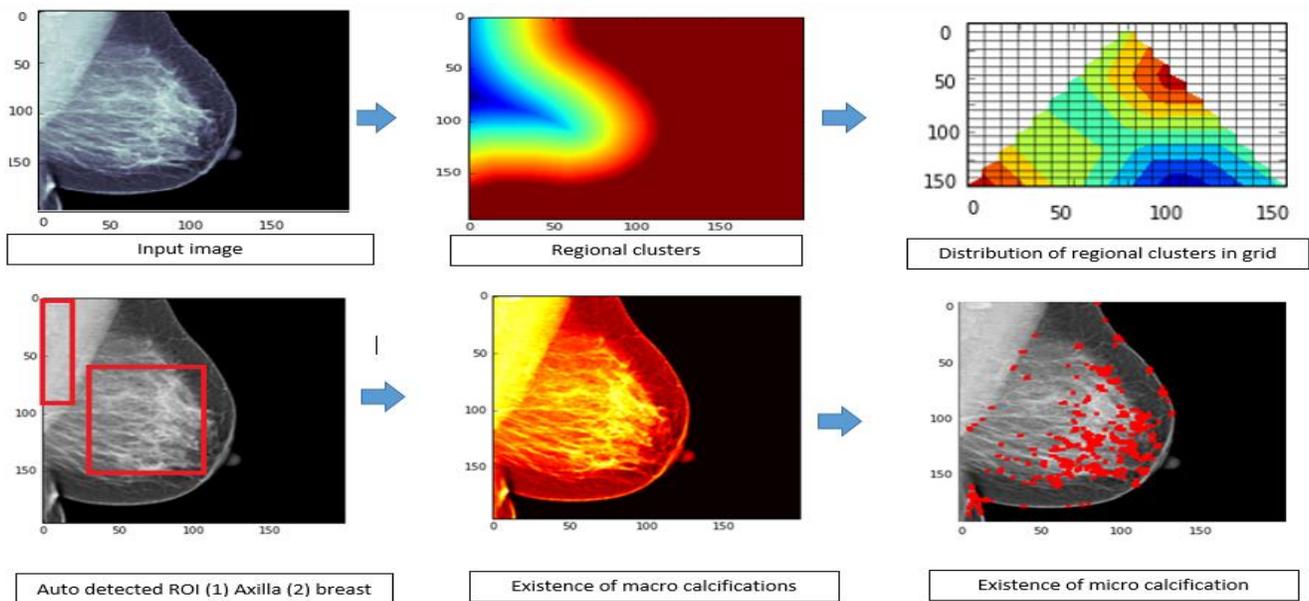


Fig. 2. Preprocess Method, Input image, Regional clusters, Distribution of clusters in grid, Auto detected ROI (1) Axilla (2) Breast, Existence of Macro-Calcification, Existence of Micro - Calcification

TABLE. I. CONFUSION MATRIX

	MALIGNANT	NON-MALIGNANT
MALIGNANT	327	35
NON-MALIGNANT	21	890
MEASURED CLASSIFICATION ACCURACY	95.60%	

TABLE. II. COMPARISON OF PROPOSED APPROACH WITH LITERATURE

APPROACH	TECHNIQUE	ACCURACY
DHEEBA, J (2014)	SWARM OPTIMIZED WAVELET NEURAL NETWORK	92.10%
LASHKARI (2016)	ANN (ARTIFICIAL NEURAL NETWORKS)	94.00%
RAJESH KUMAR(2014)	SVM, ADABOOST	87.42%
HAMID H (2016)	RBF	95.34%
GRAHAM (2005)	SVM	95.23%
OUR APPROACH	CFENNS (CALCIFICATION FEATURE EXTRACTION CONVOLUTIONAL NEURAL NETWORKS)	95.60%

TABLE. III. OVERALL PERFORMANCE OF PROPOSED METHODOLOGY

	RAW IMAGES	NO OF EXTRACTED NUCLEI	NO OF CLASSIFIED NUCLEI	NO OF MISS-CLASSIFIED NUCLEI	PRECISION	RECALL
MALIGNANT	10	362	327	35	90.33%	93.96%
NON-MALIGNANT	10	911	890	21	97.69%	96.21%

#### IV. RESULTS

Overall proposed method for preprocessing is shown in (Fig. 2) and first input image have been shown than its regional clusters have been transformed. These clusters have been further divided into further sub regions in a grid to estimate the ROI (region of interest) where two level regions have been selected, firstly the axilla margins and secondly breast regions. The involvement of axilla margins helps doctors to investigate about the potential presence of lymph nodes, since these nodes may be removed in case of applying surgical procedures such mastectomy or lumpectomy to cure the life of breast cancer patients. High contrast levels have been used by increasing the green color spectrum where massive malignant masses exist and micro-calcification level features have been recorded with the assistance of top edge level behaviors of classifications which would become more useful features to diagnose the fibro adenoma. In Fig. 3, ROC is plotted and measured classification accuracy of the system is approximated about 95.60%.

#### V. CONCLUSION AND DISCUSSION WITH FUTURE DIMENSIONS

Mammogram X-Rays are very difficult to interpret because complex, diverse and heterogynous malignant behaviors of micro-calcifications are very difficult to visualize without using machine learning techniques. Since presence of massive masses accumulated as huge noise in DICOM (Digital Communication in Medicine) images of mammography and micro calcifications may be hidden between these masses and the persistence of non-palpable lumps and nodular structure in breast is an alarming situation which may assist the radiologist because initial changes of malignant and benign diagnosis may enhance the survival rate of life. This paper proposes a system so called "Intelligent System for Detection of Micro-Calcification in Breast Cancer"

to differentiate the malignant and benign calcifications. Proposed system overall comprises over three main stages. In first stage noise reduction techniques are used by using adaptive threshold algorithm and detection of each macro and micro classification is done. In second stage of feature selection auto-crop technique have been used to crop the axilla margins and breast lesions by using proposed algorithm CFEDNN (Calcification Feature Extraction Deep Neural Networks) which is designed to avoid the manual ROIs (Region of Interest). Decision model is constructed by using the DNNs (Deep Neural Networks) and best classification accuracy was measured as 95.6%. As a future work there are some very important dimensions have been observed: How to measure the size of breast malignant masses? How to classify the histopathological images by correlating the mammogram? Another very significant problem is to estimate the growth pattern of malignant and non-malignant masses because it would provide more precise assistance to doctors to assess the survival rate of patients.

#### ACKNOWLEDGMENTS

We are highly thankful to all officials of Sukkur IBA University and medical partners for providing an opportunity to conduct the research in applied sciences.

#### REFERENCES

- [1] Torre, L. A., Bray, F., Siegel, et. al. (2015), Global cancer statistics. CA:A Cancer Journal for Clinicians, 65: (2012) pp 87–108.
- [2] Parkin, D. Max, et al. "Global cancer statistics." CA: a cancer journal for clinicians 55.2 (2005) pp 74-108.
- [3] Jemal, Ahmedin, et al. "Global cancer statistics." CA: a cancer journal for clinicians 61.2 (2011) pp 69-90.
- [4] Bleyer, Archie, and H. Gilbert Welch. "Effect of three decades of screening mammography on breast-cancer incidence." New England Journal of Medicine 1998-2005. Vol 367.21 (2012).
- [5] Kriege, Mieke, et al. "Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic

- predisposition.” New England Journal of Medicine Vol. 351.5 (2004). pp 427-437.
- [6] Fletcher, Suzanne W., and Joann G. Elmore. “Mammographic screening for breast cancer.” New England Journal of Medicine Vol. 348.17 (2003). pp 1672-1680.
- [7] Harris, Randall E., and Ernst L. Wynder. “Breast cancer and alcohol consumption: a study in weak associations.” *JaMa* Vol. 259.19 (1988). pp 2867-2871.
- [8] McGinnis, J. Michael, Pamela Williams-Russo, and James R. Knickman. “The case for more active policy attention to health promotion.” *Health affairs* Vol. 21.2 (2002). pp 78-93.
- [9] Ren, Baorui, et al. “Design and performance of the prototype full field breast tom synthesis system with selenium based flat panel detector.” *Medical Imaging. International Society for Optics and Photonics*, 2005.
- [10] Lee, Charmaine Pei Ling, et al. “3D Mammography in Diagnostic Environment: A Subjective Review.” *Journal of Medical Imaging and Health Informatics* 6.1 (2016). pp 215-220.
- [11] Dheeba, J., N. Albert Singh, and S. Tamil Selvi. “Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach.” *Journal of biomedical informatics* Vol. 49 (2014). pp 45-52.
- [12] G. Durgadevi, Himanshu Shekhar, “An Intelligent Classification of Breast Cancer Images” *Indian Journal of Science and Technology*, Volume 9(28)(2016). pp 20-30.
- [13] Lashkari, Amir Ehsan, and Mohammad Firouzmand. “Early Breast Cancer Detection in Thermogram Images using Supervised and Unsupervised Algorithms.” *The Journal of Engineering* Vol. 1.1 (2016).
- [14] Rajesh Kumar Tripathy, Sailendra Mahanta and Subhankar Paul, “Artificial intelligence-based classification of breast cancer using cellular images”, *An international journal to further the chemical sciences*, Issue 18 (2014)
- [15] Hamid Hassanpour & et. al. “Fully automatic classification of breast cancer microarray images”, *Journal of Electrical Systems and Information Technology*, 2016. pp 2314-7172
- [16] Graham, R. N. J., R. W. Perriss, and A. F. Scarsbrook. “DICOM demystified: a review of digital file formats and their use in radiological practice.” *Clinical radiology* 60.11 (2005). pp. 1133-1140.
- [17] Hao, Ying-ming, and Feng Zhu. “Fast Algorithm for Two-dimensional Otsu Adaptive Threshold Algorithm [J].” *Journal of Image and Graphics* Vol. 4 (2005).
- [18] Verma, Brijesh, Peter McLeod, and Alan Klevansky. “Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer.” *Expert systems with applications* Vol. 37.4 (2010). pp 3344-3351.
- [19] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. “Imagenet classification with deep convolutional neural networks.” *Advances in neural information processing systems*. 2012.
- [20] Sermanet, Pierre, Soumith Chintala, and Yann LeCun. “Convolutional neural networks applied to house numbers digit classification.” *Pattern Recognition (ICPR)*, 21st International Conference on. IEEE, 2012.
- [21] Jia, Yangqing, et al. “Caffe: Convolutional architecture for fast feature embedding.” *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 2014.
- [22] Simonyan, Karen, and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition.” *arXiv preprint arXiv* (2014). pp 1409.155

# Financial Market Prediction using Google Trends

Farrukh Ahmed

Research Student, Department of Computer Science &  
Software Engineering  
NED University of Engineering & Technology, Karachi,  
Pakistan, 75270

Dr. Raheela Asif

Assistant Professor, Department of Computer Science &  
Software Engineering  
NED University of Engineering & Technology, Karachi,  
Pakistan, 75270

Dr. Saman Hina

Assistant Professor, Department of Computer Science &  
Software Engineering  
NED University of Engineering & Technology, Karachi,  
Pakistan, 75270

Muhammad Muzammil

Research Student, Department of Computer Science &  
Software Engineering  
NED University of Engineering & Technology, Karachi,  
Pakistan, 75270

**Abstract**—Financial decisions are among the most significant life-changing decisions that individuals make. There is a strong correlation between financial decision making and human behavior. In this research the relationship between what people think and how stock market moves is investigated. The data from 2010 to 2015 of some of business, political and financial events which directly impact the local stock market in Pakistan is analyzed. The data was collected from search engine Google via Google trends. The association between internet searches regarding the political or business events and how the subsequent stock market moves is established. It was found that increase in search of these topics may lead to stock market fall or rise. The overall objective of this research is to predict Karachi Stock Exchange (now known as Pakistan stock exchange) 100 index by quantifying the semantics of international market. In addition to that, the relation between what an individual thinks while searching on Google which affects the local market is also investigated. The collected data has been mined by Multiclass Neural Network and Multiclass Decision Trees. The result shows that Multiclass Decision Trees performed best with an accuracy of 94%.

**Keywords**—Google trends; financial market; stock market; Karachi stock market; multiclass neural network; multiclass decision trees

## I. INTRODUCTION

We are living in a world where data is generated from all domains of life. For example, from every social media interaction do, from every computer, every mobile, every sensor and now even from watches and other wearable gadgets. The real question is how we can convert this data into meaningful information for decision making such as to predict stock market behavior. Stock market prediction is a domain of challenging factors which is based on many important aspects and collective thinking of the financial experts. Stock Market data can be acquired from different sources. Its impact has generated considerable scientific attention due to its complexity and size. Despite its huge size, such data sets capture only the final action taken at the end of a decision-making process. No insight is provided into earlier stages of this process, where traders may use this information to

determine what consequences of various actions and factors may be.

The aim of this study is to predict the behavior of local stock market in Pakistan based on available historical data and International market factors such as International Gold Rates, US dollar Rates, International stock markets and foreign exchange reserves etc. For today's world, data gathering frequently comprises of seeking on the web sources. Few years back Google has given access to cumulative information on the volume of queries for different search terms and how these volumes change over time, via publicly available service named Google Trends [1]. The gathered data was pre-processed using data cleaning and data filtering. The preprocessed data was then analyzed using Machine learning algorithms.

## II. LITERATURE REVIEW

The advent of Internet has seen people have used it as a main source of Information gathering and search engines like Google have become a gateway to this information. This fine grained data available on internet has opened up new options for researchers. Studies have found that large volumes of search queries for a specific word can linked up to real life events, such as forecasting the housing prices and sales [2]; popularity of films, games, and music on their release [3]; unemployment rate [4]; This openly available digital data also help researchers to find how the stock market moves, a recent study has found that increase in search volumes of some topics tends to precede stock market falls [5]. Some researchers have successfully found the relationship between behavior of people through social media (like twitter) and prediction of the stock market [6].

Karachi Stock Market (KSM) is one of the top 10 markets in the world. There are dozens of factors which impacts stock exchange directly or indirectly. That's why this research was intended to work on unique factors which impacts stock exchange. The objective of this research was to predict complex behavior of Karachi stock market (KSM) using historical data of KSM in combination with International economic factors such as US Dollar rates, gold prices, Oil

price, NYE 100 index data; Shanghai 100 index data are few factors which influence KSM.

### III. DATA GATHERING AND PREPARATION

#### A. Data Gathering

The datasets used in this study was gathered from different sources. These sources are described below:

##### 1) Data Generation from Google Trends Analytics

Research was primarily concentrated on gaining insight knowledge of what is going through people's mind and the most right source for it is what they search on Google. As we are aware of the fact that different events put a huge impact on financial markets. Therefore, the focus is on finding out the impact of these events via what people searched about it on Google.

The first task was to collect data about major events in Pakistan from 2010 to 2014. It includes finding major events from Wikipedia and other sources. Then these events are searched in Google trends and their *csv* files were downloaded. The major events from 2010-2015 are listed in Table 1.

TABLE I. YEAR WISE MAJOR EVENTS IN PAKISTAN

Year	Major Events
2010	Air blue airline crash, Imran Farooq murder, Aman ki asha.
2011	Karachi bombing, Karachi target killing, Lahore bomb blast, Salman Taseer assassination, Osama bin laden killing, Raymond Davis scandal, Resignation of Hussain haqani.
2012	Karachi Factory fire case, Malala Yousuf Zai attempted killing, Memo gate scandal.
2013	Finance bill 2013, General Elections 2013, IMF Loan 2013, Meer Hazar Khoso become takecarer prime minister, Nawaz Shareef becomes prime minister of Pakistan.
2014	Ban on Geo TV, Attack on Hamid Meer, Karachi Airport Attack, Model Town Incident, Peshawar School Attack, Imran Khan long march, Operation Zarb-e-Azab.

##### 2) KSM Dataset

The past 15 years of KSM data was collected. This includes Date, Index points, Volume, High and Low. Data was filtered from 2010 to March 2015 for research purposes.

##### 3) Dollar Rate Dataset

According to financial analysts, the International Gold Rate impacts KSM. Therefore, it was decided to put dollar rate prices, up and down measure in dataset for predicting Stock Market. The Dollar Rate data was obtained from Quandl [7]. It

includes Date, Rate on that date, highest and lowest price and percent change in Rate. This data was obtained from year 2010 onwards.

##### 4) Gold Rate Dataset

Gold rate is another factor which impacts stock markets. Therefore, it was decided to put international market gold rate in this research. It was obtained from Quandl from 2010 and onwards. Parameters includes Date, Gold rate in USD and change in percentage.

##### 5) Newyork Stock Exchange Dataset

According to financial analysts, International stock markets impact KSM greatly. Therefore it was decided to use Newyork stock market index data set for building prediction model. Data was obtained from Quandl which includes date, volume, opening and closing index parameters. In addition to that, US International Index, composite Index and 100 Index were also used.

##### 6) Shanghai Stock Exchange Dataset

Another international market which impact KSM exchange according to financial analysts is the Shanghai Stock Exchange. It was decided to also use shanghai stock exchange parameters in the dataset. The data of Shanghai stock exchange was obtained from year 2010 onwards. Parameters includes data, opening, closing, volume and change in percentage.

#### B. Data Pre-Processing/Cleaning

This section is focused on cleaning the data which was obtained by the process described above. Data was cleaned by going through series of steps each of which are illustrated below. Trends data obtained from Google Trends was not in proper format. Data was in three formats i.e weekly, daily and monthly. But the desired final *csv* was intended to be in day format from year 2010 to 2015. The data transformation was done from weekly to daily basis and is described in the following section.

##### 1) Weekly/Monthly to Daily Transformation

Weekly/monthly data was binded by date index column. Hence the string is aplitted in two parts that are start and end date. Thus a range was developed and then this range is converted into days by using *DateTime library* in python pandas package. The repeating values for the whole week/month generates a data file for each event .

##### 2) Merging of Data

This includes the following steps:

- Gather all event files year wise.
- Calculate mean of each file and saving the results in their respective year file. In this way, the five years of trend data in five different files can be acquired.

##### a) Merging of Data other than Trend data.

This step includes merging the data of Karachi stock exchange dataset, Newyork stock exchange dataset, Gold rates data set, Dollar rate data set, and Shanghai stock exchange data set.

b) Developing the complete dataset

The final dataset was formed by merging Newyork stock exchange dataset, Gold rates data set, Dollar rate data set, Shanghai stock exchange data set, Karachi stock exchange dataset and Trends dataset. The attributes of these datasets are shown in Table 2.

TABLE. II. ATTRIBUTES AND THEIR DATASET

Dataset	Attributes
Newyork stock exchange dataset	NYSEInternational_100Index(Open), NYSEInternational_100Index(High), NYSEInternational_100Index(Low), NYSEInternational_100Index(Close), NYSEUS_100Index(Open), NYSEUS_100Index(High), NYSEUS_100Index(Low), NYSEUS_100Index(Close), NYSEComposite(Open), NYSEComposite(High), NYSEComposite(Low), NYSEComposite(close)
Shanghai stock exchange dataset	ShanghaiStockExchange(Open), ShanghaiStockExchange(High), ShanghaiStockExchange(Low), ShanghaiStockExchange(Close)
Dollar Rates dataset	Dollar Rates, Dollar (High), Dollar ( Low )
Trends dataset	Trend
Karachi stock exchange dataset	Index(KSE)

IV. EXPERIMENTS AND RESULTS

A new column is added in the final dataset named Index-Difference-Class which is the label class. This label is obtained by the difference of subsequent observation of Index (KSE) column. If the difference is less than 100 observations than the value for label class becomes -1 and if the difference is greater than 100 observations than the value of label turns out to be 1, and 0 if the difference is between

-100 and 100.

Once the label class is obtained, the final dataset is uploaded in Azure Machine learning studio. The cleaning missing data module is used in order to remove observations which have missing data. After that Meta editor module is used to define the label class, which is index-difference-class in this case. Then split data module is used to divide data into training and evaluating data. The default setting that is used is 70:30 in Azure machine learning studio<sup>1</sup>. The training data percentage was 70% of the real data and the testing data is 30% of the actual data.

The following modules of Azure machine learning studio are used to obtain the results.

- 1) Cleaning Missing data module<sup>2</sup>.
- 2) Project column module<sup>2</sup>.

- 3) Meta data editor<sup>2</sup>.
- 4) Split data module<sup>2</sup>.
- 5) Training model module<sup>2</sup>.
- 6) Score model module<sup>2</sup>.
- 7) Evaluate model module<sup>2</sup>.

A. Applying Multiclass Neural Network

Multiclass Neural Network is constructed by using training data and evaluated by using testing data. The trainer mode is single parameter with 100 numbers of hidden rows and learning rate of 0.1. The initial learning weight diameter was set to be 0.1 with momentum 0.

The steps for applying Multiclass Neural Network for training and testing data are as follows:

- *Training Model Module* is used for training data on applied algorithm.
- *Split Data module* was connected with Multiclass Neural Network module and Training Model Module.
- *Score Model Module* is used for testing a trained classification or regression Module.
- Split data module and Training Data Module are connected with Score Model.

The resultant confusion matrix for this experiment is shown in Fig. 1.

B. Applying Multiclass Decision Forest

Secondly, the Multiclass Decision Forest is applied to the final dataset for evaluating results. The Resampling method is Replicate and trainer mode is single parameter. Number of decision trees are 8 and minimum depth of decision trees are set to be 8. Maximum depths of the decision trees are 32 and number of random splits per node is set to be 128. The minimum number of samples per leaf node is 100.

After applying all these settings, similar steps are executed as performed in the previous experiment. The confusion matrix for this experiment is shown in Fig. 2.

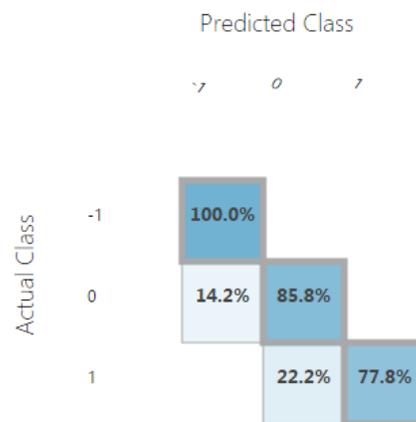


Fig. 1. Confusion matrix of multiclass neural network algorithm.

<sup>1</sup> <https://studio.azureml.net/>

<sup>2</sup> <https://msdn.microsoft.com/en-us/library/azure/dn906033.aspx>

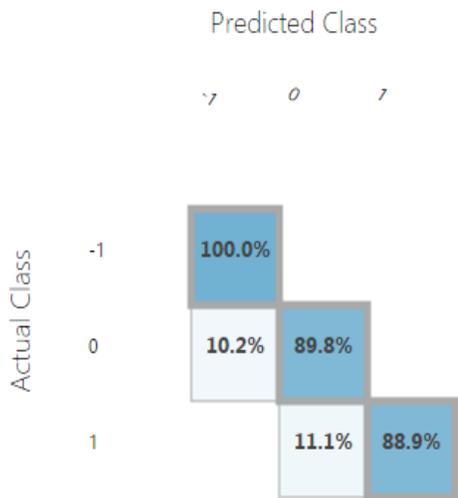


Fig. 2. Confusion matrix of multiclass decision forest.

### C. Evaluating the results.

The Precision, Recall and overall results of Multiclass neural network and Multiclass decision forest are shown in Table 3.

As it can see from Table 3, that multiclass decision forest has a higher accuracy than multiclass neural network.

TABLE. III. PREDICTING STOCK MARKET

Algorithms	Overall Precision (%)	Recall (%)	Average Accuracy (%)
Multiclass neural network	85.3	87.8	90.7
Multiclass decision forest	89.3	92.8	94.1

## V. CONCLUSION AND FUTURE WORK

The research was intended to predict KSM behavior by quantifying the semantics of people with the help of Google Trends Analytics. In addition to the datasets from Google Trends, this work also involves international factors which impact financial markets.

The research work can be expanded by introducing additional features in dataset. Other factors like inflation rate, interest rate, tax figures etc. can be used as supporting factors for improving the overall accuracy of algorithms. Once a stable model is established, work on a data product can be done which will be beneficial for investors. The data available on Google Trends or other resource can be utilized as an input to web service and getting the analysis and the prediction about how financial market moves.

The product will directly analyze people behavior and it will help investors in decision making process for buying or selling stocks. Investors will know about the overall geo political situation and will act upon it to get better financial outcomes.

### REFERENCES

- [1] T. Preis, T. H. S. Moat, H. E. Stanley, "Quantifying Trading Behavior in Financial Markets Using Google Trends", SCIENTIFIC REPORTS 3 Article number: 1684, DOI: 10.1038/srep01684J., 2013.
- [2] W. Lynn, E. B. Jolfsson, "The Future of Prediction: How Google Searches Foreshadow Housing Prices and Quantities", ICIS 2009 Proceedings. 147. <http://aisel.aisnet.org/icis2009/147,2009>.
- [3] S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, D. J. Watts, "Predicting consumer behavior with Web search". Proc Natl Acad Sci USA 107(41):17486–17490, 2010.
- [4] N. Askitas, K. F. Zimmermann, "Google econometrics and unemployment forecasting", Discussion Paper No. 4201, 2009.
- [5] C. Chester, P. Tobias, H. E. Stanley and H. S. Moatb, "Quantifying the semantics of search behavior before stock market moves", 2013.
- [6] J. Bollen, H. Mao, X. Zeng, "Twitter mood predicts the stock market", 2011.
- [7] <https://www.quandl.com/data/CURRFX/USDPKR-Currency-Exchange-Rates-USD-vs-PKR>.

# Ultra-Wideband Antenna Design for GPR Applications: A Review

Jawad Ali\*<sup>†</sup>, Noorsaliza Abdullah\*, Muhammad Yusof Ismail\*, Ezri Mohd\*, Shaharil Mohd Shah\*

\*Department of Communication Engineering, Faculty of Electrical and Electronic Engineering,  
Universiti Tun Hussein Onn Malaysia,  
Johor, Malaysia

<sup>†</sup> Department of Electrical Engineering,  
COMSATS Institute of Information Technology,  
Lahore, Pakistan

**Abstract**—This paper presents a comparative review study on ultra-wideband (UWB) antenna technology for Ground Penetrating Radar (GPR) applications. The proposed antenna designs for UWB ground penetrating radar include a bow-tie antennas, Vivaldi antennas, horn antennas, planar antennas, tapered slot antennas, dipole antennas, and spiral antennas. Furthermore a comprehensive study in terms of operating frequency range, gain and impedance bandwidth on each antenna is performed in order to select a suitable antenna structure to analyze it for GPR systems. Based on the design comparison, the antenna with a significant gain and enhanced bandwidth has been selected for future perspective to examine the penetration depth and resolution imaging, simultaneously suitable for GPR detection applications. Three different types of antennas are chosen to be more suitable from the final comparison which includes Vivaldi, horn and tapered slot antennas. On further analysis a tapered slot antenna is a promising candidate as it has the ability to address the problems such as penetration depth and resolution imaging in GPR system due to its directional property, high gain and greater bandwidth operation, both in the lower and higher frequency range.

**Keywords**—Ultra-wideband antennas; ground penetrating radar; antennas; antenna review

## I. INTRODUCTION

The research interest on ultra-wideband (UWB) systems has gained popularity mainly after the year 2002 when the US department of Federal Communications Commission (FCC) allocated a license-free spectrum for Industrial and Scientific purposes. This greater step of FCC has opened new doors of researches for UWB in the field of wireless communications and microwave imaging [1], [2]. The UWB covers a frequency band ranges from 3.1 to 10.6 GHz that has foreseen the applications in the field of Wireless Local Area Networks (WLAN), Wireless Body Area Networks (WBAN), Wireless Interoperability for Microwave Access (WiMAX), Wireless Personal Area Networks (WPAN) and Ground Penetrating Radar (GPR) technology where wide bandwidth is required [3], [4].

GPR is one of the major applications of UWB technology, which is widely used in military and civilian applications such as the detection of land-mines [5]. In addition, GPR is also used in remote sensing techniques such as nondestructive

testing of concrete and detection of trapped people under-debris or in opaque environment [6]. For the implementation of UWB GPR systems, the performance of various antenna designs, such as bow-tie antenna [7], [8], spiral antenna [7], loaded dipole antenna [9], TEM horn antenna [10], [11], tapered slot antenna [12], [13] and Vivaldi antenna [14], [15] have been evaluated.

This paper examines the technical and methodological aspects involved during the design of ultra-wideband antennas for ground penetrating radar detection applications. The study has been performed based on antenna gain, directivity, complexity of the design as well as the frequency bandwidth. Later the comparison of results to address the fundamentals of GPR applications, such as penetration depth and resolution has been performed. Finally one of the best suitable antenna design is chosen for further research and future directions. The organization of this paper is as follows: Section II discusses the UWB antenna designs and methodologies suitable for GPR application. Section III presents the results and analysis of the antennas discussed in the previous section. Finally, the paper is concluded in Section IV.

## II. ANTENNA DESIGN FOR GPR

A number of ultra-wideband antennas have been designed for GPR applications. The study based on the lower frequency band is conducted mainly to improve the penetration depth while the designing in the higher frequency band is performed to achieve better resolution imaging for GPR systems. Some of the work focus on the entire UWB frequency range to further enhance the bandwidth while other focused on gain enhancement.

Based on these previously mentioned requirements, different types of antennas such as bow-tie antennas, Vivaldi antennas horn antennas and few more antenna designs have been studied and implemented for GPR applications.

### A. Bow-tie Antenna

A bow-tie UWB antenna is widely used in the design for GPR applications, as it has the ability to reduce ground susceptibility during GPR operations [16]. Fig. 1 shows a simple bow-tie antenna design which consists of two flares connected to a common feed.

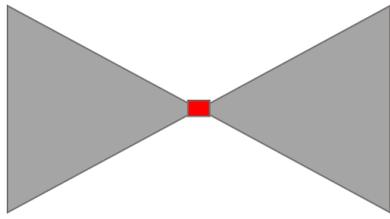


Fig. 1. Bow-tie antenna design [24].

The length,  $l$  and width,  $w$  of the flares in bow-tie structure can be determined by (1) and (2) [16-19]:

$$l = \frac{1.6\lambda_o}{\sqrt{\epsilon_r}} \quad (1)$$

$$w = \frac{0.5\lambda_o}{\sqrt{\epsilon_r}} \quad (2)$$

Where  $\lambda_o$  is the wavelength of the low frequency range in free space. The flare angle is totally dependent on impedance [19].

In [17], a monolithic bow-tie UWB antenna for GPR applications is designed by using a total geometric morphing approach with small pads to form an array. The antenna has been fabricated on a glass epoxy FR-4 substrate with a dielectric constant value,  $\epsilon_r$  of 20 and it is fed by a coaxial cable for impedance matching to the balanced current balun.

A slot and reflector have been introduced into the antenna structure to achieve a wide beam width and unidirectional radiation pattern [18]. The antenna is fed by a coplanar waveguide (CPW).

Furthermore, a modified bow-tie antenna with a shielded back cavity for omnidirectional radiation pattern is also proposed for GPR application [19]. The antenna is fabricated on a FR-4 substrate with a dielectric constant,  $\epsilon_r$  of 4.6 and centered at 900 MHz. The edge of the antenna is cut to make the antenna to be more compact and study the effect of edges on reflection coefficient.

As the research progresses, new designs have been developed with modifications in the shapes and dimensions of the bow-tie antenna. In addition, one of the antennas has been designed by using the finite difference time domain approach [20]. CPW is used as the feeding method while metal stubs is introduced for impedance matching. The work in [7] introduces the structure of a bow-tie antenna into its dipole antenna to improve the gain for GPR. The ideal performance of a bow-tie dipole antenna can be achieved by feeding the antenna with 100-200 ohm impedance. In [8], a resistive-loaded UWB bow-tie antenna is proposed to improve the forward gain from the employment of metamaterial lens.

As the signal of the GPR system required propagating through inhomogeneous media, the efficiency of the antenna must also be taken into account. Thus, in order to enhance the efficiency, a compact shape slotted bow-tie antenna is designed [21]. The compactness in the shape of antenna is achieved by rounding the sharp corners of the bow-tie structure and triangular stubs with extended arms. The antenna is fed by a CPW and a thin graphite sheet is used to reduce the end-fire reflection of the UWB antenna for GPR application.

In [22], an elliptical shape of bow-tie antenna is proposed to broaden the bandwidth. A semicircular slotted-tuned half-ellipse antenna in a bow-tie formation is designed in [23]. The proposed antenna has four semicircular slots in the bow-tie ellipse to further improve its penetration properties for vital signs detection.

A low cost UWB bow-tie antenna for GPR is presented in [24]. A reflector cavity design technique is used where the model is initially tested with only a bottom reflector followed by a side reflector and both reflectors at the same time. There is also a gap between both flares of the bow-tie where a balun is deployed to match the impedance. Finally, the antenna is equipped with a four-sided reflector to study the gain parameters.

### B. Vivaldi Antenna

Vivaldi antennas are commonly used for the applications such as GPR which require a greater bandwidth usually with a ratio 10:1 [16]. Therefore, various designs of Vivaldi antennas have been studied and discussed. Fig. 2 shows the design of a Vivaldi antenna.

A Vivaldi antenna consists of a radiating and ground planes of a similar shape placed in the opposite of each other separated by a substrate. The exponential curves in the Vivaldi antenna as shown in in the figure can be described by (3) and (4) [15]:

$$x = \pm 0.14e^{0.27y} \quad 0.1 \leq y \leq 11.4 \quad (3)$$

$$x = \pm 3.04e^{0.425y-4.845} \quad 11.5 \leq y \leq 14.2 \quad (4)$$

In [14], a Vivaldi antenna array is designed for GPR measurement application. An experiment is conducted to observe the penetration through a concrete structure. Five array elements are used in the receiving antenna to gather all the information.

In order to achieve the compactness, small size, balanced bandwidth and a reasonable gain for GPR applications, an antipodal structure semi-empirical Vivaldi antenna is proposed [25]. The antenna is fed by a microstrip line with PTFE board as the substrate. The antenna is designed based on two aspects: the transition aspect and radiation aspect, to meet the UWB system requirements.

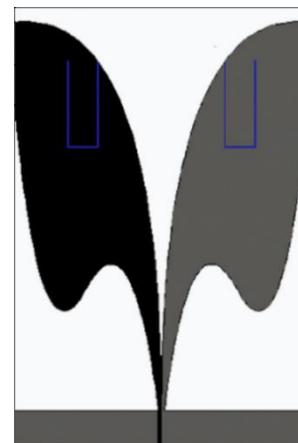


Fig. 2. A Vivaldi antenna [27].

Another balanced antipodal Vivaldi antenna structure has been proposed in which an L-shaped slot has been introduced at the edge of a radiating patch [26]. In the design, a structural increase in the radiating patch has increased the electrical length of the antenna while the L-shaped slot has enabled a radiation from the outer edge of the antenna. A comparison is then made with a conventional Vivaldi antenna with various substrates such as a FR-4, Rogers, PTFE and TP-2. In [27], two antipodal-shaped Vivaldi antennas are proposed. The first antenna consists of a simple antipodal shape to achieve compactness and comparatively small in size with a wider bandwidth and good gain. U shaped slots have been introduced into the structure to enhance the gain.

A novel Vivaldi antenna with an exponentially tapered slot edge (TSE) is designed in [15]. Seven pairs of electromagnetic band gaps in either a loop or square shapes have been introduced in the ground plane. The purpose of these band gaps is to extend the lower end bandwidth of the UWB antenna and improve the impedance matching over the same band. The proposed antenna also helps in the electrical length reduction in comparison with the original antenna without tapered slots.

A double-slot antenna structure is proposed in [28] to further improve the gain and directivity of the Vivaldi antenna. The antenna is also suitable for UWB GPR applications. The slots are excited by using a T-junction power divider to generate plane waves in the E-plane at the aperture of the antenna and it is compared with a conventional Vivaldi antenna that uses an exponentially tapered method. Another directional Vivaldi antenna with eye-shaped slots is proposed in [29] to reduce the side lobes and increase the efficiency to approximately 80%. The proposed antenna also has a greater gain which makes it suitable for GPR systems.

### C. Horn Antenna

In a GPR system, better depth penetration and ease of scanning the shallow targets often require elevated antennas because the energy must radiates into the ground for detection. Thus, it makes a horn antenna as one of the best candidates as it is less susceptible to the effects of ground [16]. A typical horn antenna can be seen in Fig. 3.

The construction of a horn antenna mainly consists of two parts: An aperture and waveguide transition. The dimensions of an aperture can be approximated by (5) and (6) [11]:

$$\Delta = \frac{w^2}{8R} \quad (5)$$

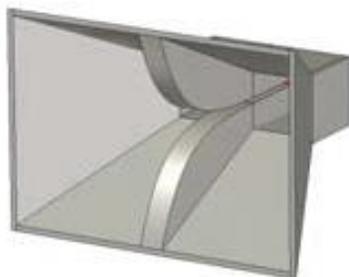


Fig. 3. A horn antenna [30].

$$S = \frac{\Delta}{\lambda} = \frac{w^2}{8\lambda R} \quad (6)$$

Where S is a dimensionless quantity, w is the distance between the horn apertures, R is the length of aperture and  $\lambda$  is the wavelength.

A double ridged horn (DRH) antenna is presented in [30]. The antenna has a wider bandwidth and is proposed for GPR applications. The reduction in operating frequency will compromise the size of the antenna. Thus, in order to achieve a small size but with lower operating frequency, the ridges have been extended from the aperture plane and the gap is filled with dielectric material. In [7], another DRH antenna is also designed to increase the bandwidth.

An UWB quad ridged horn (QRH) antenna is proposed in GPR application for deep penetration by modifying a conventional triangular transverse electromagnetic (TEM) horn [11]. First modification is performed by tapering the horn plates with a curvature and the second modification is the addition of double ridges in between the horn aperture [30]. A substrate with a high dielectric constant is used to fill the gap. The substrate used is the Rogers RT-3010 with  $\epsilon_r$  of 10.2. The fabricated antenna is then compared with the SH-68 Satimo horn antenna that is commercially available.

There are some reflections and ringing effects in the conventional TEM horn antenna that are not suitable in GPR application. Therefore, an optimized design of TEM horn antenna for UWB is proposed [10]. The antenna is designed by carving an arc in the two exponentially tapered plates of horn and perpendicular plates are connected at the lower end.

In [31], another TEM horn with a flare shape is designed with a balun for proper current balancing. The design consists of two flaring conductors constructed by cutting a brass sheet for the top and bottom flare and combined together with a triangular slab from Styrofoam to secure the conductor in a flare. The antenna is experimentally tested with a concrete to observe the gain when in contact with other materials.

In [9], an UWB horn antenna consists of a coaxial feed line and a proper waveguide with a round and tapered-shaped aperture. The antenna is designed by using the D-angle and W-angle methods for the waveguide. The aperture has rounded corners to improve the performance.

Various feeding techniques are also discussed in [6] other than proposing a new feeding technique. In order to realize the new technique, the waveguide of the antenna is fixed and a screw is added on the opposite side of the feeder to increase the gain.

### D. Planar Antenna

A planar antenna, as shown in Fig. 4, is a popular candidate for UWB due to its simplicity, conformity in design, cost effectiveness and light weight properties [32], [33].

A planar antenna mainly consists of a radiating patch and either a full or partial ground plane with a defected ground structure (DGS) [15] which can be varied according to the

design requirement. The length, L and width, W of the patch can be determined from (7) and (8) [32], [37]:

$$W = \frac{c}{2f} \left[ \sqrt{\frac{c}{(\epsilon_r + 1)}} \right] \quad (7)$$

$$L = \frac{c}{2f} - \Delta L \quad (8)$$

Where:

$$\frac{\Delta L}{h} = \frac{0.412 \left[ (\epsilon_{\text{reff}} + 0.3) \left( \frac{w}{h} + 0.264 \right) \right]}{\left[ (\epsilon_{\text{reff}} - 0.258) \left( \frac{w}{h} + 0.8 \right) \right]} \quad (9)$$

$$\epsilon_{\text{reff}} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left\{ 1 + 12 \frac{h}{w} \right\}^{-\frac{1}{2}} \quad (10)$$

In [34], a hexagonal fractal patch antenna is proposed to achieve a wideband characteristic. The design is modified by iteration of hexagonal fractals with slits in the ground plane to create a notch that affects the impedance bandwidth. Hexagonal slots are also introduced in the patch and slots in the ground plane alongside the slits.

A modified circular patch antenna for UWB GPR applications is proposed in [35]. The conventional circular disc is modified to produce a mickey-shaped patch radiator with a CPW feed. The substrate used is FR-4 with  $\epsilon_r$  of 4.3. A rectangular copper reflector is inserted below the antenna to make it directional.

Another circular disc antenna is designed in [36] to improve the impedance bandwidth and efficiency of the planar antenna. A rectangular slot is introduced at the edge of the ground.

A stepped feeding and two level notched stairs in the patch with a partial ground plane over Rogers Duroid RT-5880 with a dielectric constant,  $\epsilon_r$  of 2.2 are included in the design of an UWB planar antenna in [37]. The antenna is tested in step-by-step basis. Firstly, the antenna is simulated and tested only with a simple rectangular patch and a partial ground followed by a stair patch and a partial ground. Finally, the stubs are introduced in the stairs feed patch for band notch. A multioctave frequency selective antenna with a reflector is proposed in [38] to study the gain performance.

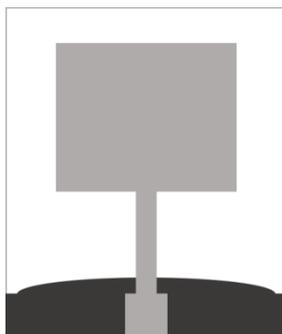


Fig. 4. A planar antenna [37].

An UWB quasi-planar antenna in [39] is designed for gain enhancement. The antenna consists of a CPW-fed semicircular disc monopole antenna with a short horn. The shape of this quasi antenna makes it a potential candidate for GPR technology.

A quadruple-band planar antenna is developed in [40]. This is a simple microstrip patch antenna with four slots of different geometries namely, square, rectangle, circle and ellipse. Similarly, another slotted monopole planar antenna with a key shape is designed in [41] where the shape is achieved by etching two symmetrical curved slots in the circular patch which in return, increases the efficiency of the antenna. A spanner-shaped antenna is proposed in [42] which is achieved by cutting a rectangular slot on the upper side of minor axis in an elliptical-shaped antenna to improve the gain and bandwidth for GPR applications.

### E. Tapered Slot Antenna

Gain is one important parameter that is required in a GPR system as it can enhance the signal appearance which can be affected by signal attenuation and other losses as well [43]. In order to improve the gain factor in an antenna design, different methods are used. One of them is the tapered slot method. Fig. 5 shows the UWB tapered slot antenna which consists of a coplanar patch and ground that is suitable for GPR applications.

In [12], one of the tapered slot compact antennas with a high gain for GPR applications is proposed. This compact planar antenna is fed by a CPW with one slot line to supplement the taper slot. A resistive loading is introduced with some discontinuities in the design to avoid any strong reflections. The antenna is designed by introducing two main slots: one slot is represented by a Vivaldi shape and another slot is a triangular slot.

A double exponentially tapered slot antenna (DETTSA) is designed in [13] in which the exponential arms of the antenna are rolled back to improve the bandwidth. Another modification in the design is the introduction of a coupled-strip line (CPS) instead of a CPW for feeding purpose. The measured losses of a CPS are less than that of a CPW. The antenna is designed on a FR-4 dielectric material with  $\epsilon_r$  of 4.4.

In [44], another modified slotted antenna with a backed absorber is investigated for broadband antennas. The proposed design is modified by cutting a slot in the patch and then tapering the slot edges so that the antenna can perform better in GPR applications. This model is designed on a lossy TMM-10 substrate with  $\epsilon_r$  of 9.2.

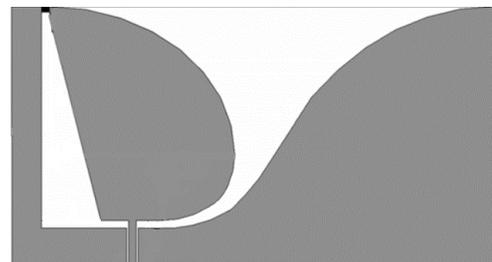


Fig. 5. Tapered slot antenna [12].

A semicircular slotted antenna is introduced in [5] to enhance the gain factor. In the design, semicircular slots are introduced with tapered transitions for a CPW feed. A frequency selective surface (FSS) method with a dual-layer reflector is used which helps in the gain enhancement and impedance matching. The distance between the antenna and reflector can be approximated by (11) [45]:

$$\phi_{FSS} - 2\beta h_{FSS} = 2n\pi; \quad n = \dots, -2, -1, 0, 1, 2, \dots \quad (11)$$

Where:

$\phi_{FSS}$  = Reflection phase

$h_{FSS}$  = Distance between reflector and antenna

$\beta$  = Propagation constant of free space

Another high gain antenna with FSS has been designed by etching two elliptical structures for the slot [45]. The gain of the antenna without the FSS is very low and has been further enhanced with the presence of FSS. A star-shaped antenna is designed with an asymmetric slot in the ground plane with an open ended CPW feed [46] for gain enhancement from the employment of FSS which makes it suitable for GPR applications.

#### F. Dipole, Cone, and Spiral Antenna

Dipole, canonical and spiral antennas are also the potential candidates for GPR applications and are discussed in this section. A dipole antenna can be constructed on either a planar structure or from a wire [7]. Various dipole antennas have been designed such as wired dipole, elliptical dipole and short dipole [47] for UWB applications. The length, L of a dipole from a simple wire can be calculated as follows [52]:

$$L = \frac{1}{4\lambda} \quad (12)$$

Where:

$$\lambda = \frac{c}{f} \quad (13)$$

The advantage of using a wired dipole is that it ensures high microwave power which is preferable in GPR technology [7]. Similarly, cone and spiral antennas are also proposed for UWB applications. To increase directivity, a radiation stub cone or disk cone [48] has been introduced into the antenna with a fractional bandwidth of 70-80% [7]. Spiral antennas are also favorable as the antennas have a balanced feed over the entire frequency band [49]. The arms of spiral antenna can be defined by using a polar function which is given by (14) [16]:

$$r = R_0 e^{a\phi} \quad (14)$$

Where  $R_0$  controls the radius of the spiral as it grows exponentially, while  $a$  controls the flare rate.

In [50], an elliptical-shaped structure is introduced into each of the dipole's arm as shown in Fig. 6. The proposed design is capable to improve the gain and reflection coefficient for UWB applications. The elliptical slots are used for time domain analysis. Another antenna consists of a 16-port shared-arm dipole array is proposed for radar imaging [51].

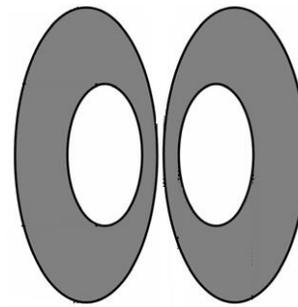


Fig. 6. A dipole antenna [50].

A magneto-electric dipole is investigated in [52]. In the design, a greater bandwidth is achieved by introducing two slots in the magneto-electric dipole by using a rectangular box reflector to fix the broadside direction of main beam.

Other antenna designs for UWB GPR applications are constructed from the cone and spiral shapes [7]. A cone antenna, as shown in Fig. 7, is an elementary antenna in 3-D structure to improve the impedance bandwidth for GPR systems. The angle of the cone is always related to the impedance, which can be calculated from (15) [49]:

$$Z = 120 \ln \left( \cot \left( \frac{\phi_k}{2} \right) \right) \quad (15)$$

On the other hand, spiral antennas with two uniform width arms with gain are also suitable for GPR applications. Fig. 8 shows the spiral antenna.

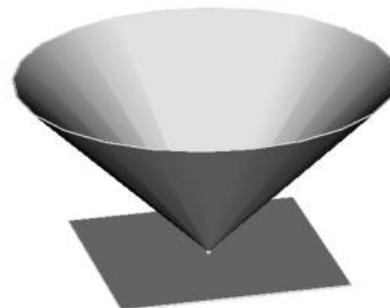


Fig. 7. Cone antenna design [7].

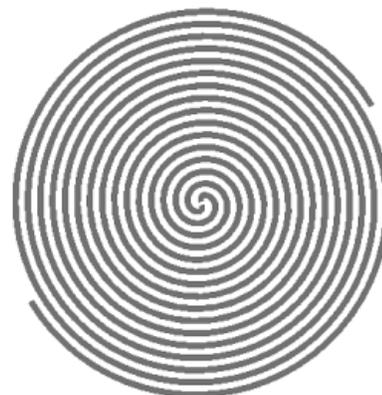


Fig. 8. A spiral antenna [7].

III. RESULTS AND DISCUSSIONS

Ultra-wideband antenna designs that are suitable for ground penetrating radar applications mainly require greater frequency bandwidth along with the parameter of high gain. Based on the above mentioned parameters, results of antenna designs discussed in Section 2 are presented for the analysis during this section.

The operating frequency of the designed antenna, its percentage bandwidth and the maximum value of obtained gain over the entire UWB band are summarized in tabular form. This summarization helps us to develop a better understanding regarding antenna performance for GPR applications. The measured results for bow-tie antennas are shown in Table 1. From these results, it can be observed that the bow-tie antenna designed by cutting edges has sufficient percentage bandwidth, which is suitable for an UWB GPR system. But there are some significant differences shown in the gain obtained for these antennas. The results of a backed cavity bow-tie antenna designed in [24] and bow-tie slotted antenna in [21] have a comparatively better gain with the value around 7 dB. But the bow-tie antenna designed using metamaterial lens has the highest gain with the value of 11.52 dB. In addition to this, these designed bow-tie antennas also have a sufficient flares length which is preferable to match the lower frequency band. Because of this reason these antennas are mainly capable of targeting the penetration depth properties for the GPR detection applications. However, the discussion about the ability of these antennas to distinguish among different detectable objects through imaging is difficult because of low resolution properties, as lower frequencies are not suitable for imaging characteristics.

In a similar manner, the measured gain and bandwidth obtained for Vivaldi antenna designs are shown in Table 2. The percentage bandwidth of all the proposed UWB antennas is sufficient to consider Vivaldi shape design for GPR applications. This is due to the fact that Vivaldi antennas offer dimensionality as well as ground optimization to achieve the desired bandwidth. The exponentially tapered slotted edge antenna designed in [15] has the least percentage bandwidth and the value of gain for this antenna is 8 dB, due to which this design is also considerable for GPR systems. From the table, it can be deduced that the gain of Vivaldi antennas have considerable values and also they covered the complete UWB frequency band defined from 3.1 to 10.6 GHz. The antipodal Vivaldi antenna designed in [27] performs better with greater bandwidth of approximately 15.9 GHz and the gain of 9.6 dB.

TABLE. I. PERFORMANCE OF BOW-TIE ANTENNAS

Table with 4 columns: Ref. #, Frequency (GHz), Percentage Bandwidth (%), Maximum Gain (dB). Rows include data for various bow-tie antenna designs with frequencies ranging from 0.012 to 4.8 GHz and gains from 2.5 to 11.52 dB.

TABLE. II. PERFORMANCE OF VIVALDI ANTENNAS

Table with 4 columns: Ref. #, Frequency (GHz), Percentage Bandwidth (%), Maximum Gain (dB). Rows include data for various Vivaldi antenna designs with frequencies ranging from 0.05 to 20 GHz and gains from 7.7 to 15.2 dB.

Another antipodal antenna is also proposed in [27] with U-shaped slots and a mouth opening at the order of λ/4 also gives significant results with a gain of 15.2 dB and a greater bandwidth that covers UWB band. These results shown in Table 2 provides a wider opportunity for the Vivaldi shape antenna designs that makes it one of the suitable candidate for GPR applications with deep penetration and better resolution imaging.

Horn antenna designs, discussed in the earlier section possess high gain value mainly because of its directional properties. Measured results of horn antennas in the form of gain and percentage bandwidth are shown in Table 3. From the table, the performance of the antennas in terms of gain and bandwidth is analyzed and the results showed a significant increase in gain. As such, a TEM flare horn antenna designed in [31] can be used in a GPR system for detection as it addresses the requirements of deep wave penetration and better resolution imaging to distinguish among the detected objects due to its high gain as well as greater bandwidth compared to remaining horn designs. While other designed antennas seemed to target penetration depth as their operating bandwidth lies within the lower frequency band. The horn antenna designed in [6] has the least percentage of bandwidth even though its gain value is significantly higher due to which it can penetrates up to few feet under the ground. Nevertheless, a horn antenna design is still one of the best suitable candidates for GPR applications because of its highly directional properties, suitable gain, proper impedance bandwidth and less susceptibility to ground effect the performance of antenna.

TABLE. III. PERFORMANCE OF HORN ANTENNAS

Table with 4 columns: Ref. #, Frequency (GHz), Percentage Bandwidth (%), Maximum Gain (dB). Rows include data for various horn antenna designs with frequencies ranging from 0.05 to 12 GHz and gains from 9.9 to 15 dB.

The results obtained from planar antenna designs are shown in Table 4. From the given table, printed circular patch antennas with the slotted ground are designed in [36], [38]. These antenna have sufficient gain value which is approximately 7.5 dB and also have greater bandwidth. This significant increase in the bandwidth of planar antennas can be attributed to the extra electromagnetic coupling between the ground and the radiating patch [36]. The performance of a circular patch antenna designed in [36], makes planar antenna a suitable candidate for GPR applications. The performance of the other designed planar antennas is also considerable only for deep penetration as the bandwidth operates in lower frequencies.

Tables 5 and 6 shows the results obtained from a tapered slotted antennas, dipole, spiral and cone antennas. From the results shown in Table 5, it can be observed that the tapered slot antenna (TSA) designed in [12], double exponentially tapered slot antenna (DETTSA) designed in [13] and eye-shaped slotted antenna designed in [45] have comparatively high gain with greater percentage of bandwidth. The good performance of tapered slot antennas is mainly due to the coplanar properties of radiator and ground, and tapering characteristics, which is helpful in order to improve the performance of the antennas. Therefore, these tapered slot antennas have the ability to address the problem of deep penetration properties and high resolution imaging in GPR applications. From the results obtained in Table 6, it can be observed that the magneto-electric dipole antenna designed in [52] has significantly high gain value over the entire operating bandwidth. This magneto-electric antenna also cover the complete UWB band which makes this antenna design as one of the suitable antenna that can be designed for GPR detection applications.

TABLE IV. PERFORMANCE OF PLANAR ANTENNAS

Table with 4 columns: Ref. #, Frequency (GHz), Percentage Bandwidth (%), Maximum Gain (dB). Rows include data for various antenna designs with their respective frequency ranges, bandwidth percentages, and gain values.

TABLE V. PERFORMANCE OF TAPERED SLOT ANTENNAS

Table with 4 columns: Ref. #, Frequency (GHz), Percentage Bandwidth (%), Maximum Gain (dB). Rows include data for various tapered slot antenna designs with their respective frequency ranges, bandwidth percentages, and gain values.

TABLE VI. PERFORMANCE OF DIPOLE, CONE AND SPIRAL ANTENNAS

Table with 4 columns: Ref. #, Frequency (GHz), Percentage Bandwidth (%), Maximum Gain (dB). Rows include data for dipole, cone, and spiral antenna designs with their respective frequency ranges, bandwidth percentages, and gain values.

From the results obtained for different types of antennas, designed specifically for GPR applications for detection purpose, it can be observed that the Vivaldi, horn and tapered slot antenna designs have significantly better performance. However, based on further comparison between these designed antennas, a tapered slot antenna is the best suitable candidate for UWB ground penetrating radar system and it has the potential to be further developed due to significant gain and greater bandwidth.

IV. CONCLUSION

A comparative review based study into a potential ultra-wideband antenna for GPR applications has been performed in this paper. The detailed summary of a bow-tie antenna, Vivaldi antenna, horn antenna, planar antenna and tapered slot antenna along with different design methodologies have been presented. The demonstration of results and the discussion about it, mainly focused on the gain as well as the bandwidth because these two parameters shows a greatest interest to design UWB antennas for GPR systems. Based on the measured results, three different types of antennas have been selected which includes Vivaldi, horn and tapered slot antennas. But on further comparison it is concluded from the study that a tapered slot antenna can be considered for future research. These antennas have the potential to address the issues of deep penetration under the surface of ground and better resolution imaging for GPR systems mainly because of its directional properties, high gain and greater operational bandwidth, both in the lower as well as higher frequency range.

REFERENCES

List of 5 references including authors like M. M. Islam, P. Sarkar, S. Shi, J. Zhang, and Y. Ranga, detailing various antenna designs and their performance characteristics in GPR applications.

- [6] C. Ozdemir, B. Yilmaz, S. I. Keceli, H. Lezki and O. Sutcuoglu, "Ultra-Wide Band horn antenna design for Ground Penetrating Radar: A feeder practice," 15th Int. Radar Symposium (IRS), Gdansk, pp. 1-4, 2014.
- [7] I. Hertl, and M. Strycek, "UWB antennas for ground penetrating radar application," 19th Int. Conference on Applied Electromagnetic and Communication, Dubrovnik, pp. 1-4, 2007.
- [8] K. K. Ajith, and A. Bhattacharya, "Improved ultra-wide bandwidth bow-tie antenna with metamaterial lens for GPR applications," Proceeding of the 15th Int. Conference on Ground Penetrating Radar, Brussels, pp. 739-744, 2014.
- [9] A. Ahmed, Y. Zhang, D. Burns, D. Huston and T. Xia, "Design of UWB Antenna for Air-Coupled Impulse Ground-Penetrating Radar," IEEE Geoscience and Remote Sensing Letters, vol. 13, pp. 92-96, 2016.
- [10] A. A. Jamali, and R. Marklein, "Design and optimization of ultra-wideband TEM horn antennas for GPR applications," XXXth URSI General Assembly and Scientific Symposium, Istanbul, pp. 1-4, 2011.
- [11] H. A. Mohamed, H. Elsadek and E. A. Abdallah, "Quad ridged UWB TEM horn antenna for GPR applications," IEEE Radar Conference, Cincinnati, pp. 79-82, 2014.
- [12] J. Shao, G. Fang, Y. Ji, K. Tan and H. Yin, "A Novel Compact Tapered-Slot Antenna for GPR Applications," IEEE Antenna and Wireless Propagation Letters, vol. 12, pp. 972-975, 2013.
- [13] F. Zhang, G. Y. Fang, Y. C. Ji, H. J. Ju and J. j. Shao, "A Novel Compact Double Exponentially Tapered Slot Antenna (DETSA) for GPR Applications," IEEE Antenna and Wireless Propagation Letters, vol. 10, pp. 195-198, 2011.
- [14] H. Liu, and M. Sato, "In situ measurement of pavement thickness and dielectric permittivity by GPR using an antenna array," NDT & E International, vol. 64, pp. 65-71, 2014.
- [15] D. M. Elsheakh, and E. A. Abdallah, "Novel shapes of Vivaldi antenna for ground penetrating radar (GPR)," 7th European Conference on Antenna and Propagation (EuCAP), Gothenburg, pp. 2886-2889, 2013.
- [16] F. B. Gross, Frontiers in antennas: Next generation design & engineering, New York: McGraw-Hill Professional, 2011.
- [17] R. Persico, N. Romano, and F. Soldovieri, "Design of a balun for a Bow tie antenna in reconfigurable Ground Penetrating Radar systems," Progress In Electromagnetic Research C, vol. 18, pp. 123-135, 2011.
- [18] Q. Lu, L. Zhou, C. Tan and L. Guanghua, "A novel wide beam UWB antenna design for Through-the-Wall radar," Int. Conference on Microwave and Millimeter Wave Technology, Chengdu, pp. 1912-1915, 2010.
- [19] G. Chen, and R. C. Liu, "A 900MHz shielded bow-tie antenna system for ground penetrating radar," Proceedings of the XIII Int. Conference on Ground Penetrating Radar, Lecce, pp. 1-6, 2010.
- [20] F. Sagnard, and F. Rejiba, "Wide band coplanar waveguide-fed bowtie slot antenna for a large range of ground penetrating radar applications," IET Microwaves, Antenna & Propagation, vol. 5, pp. 734-739, 2011.
- [21] R. Nayak, S. Maiti and S. K. Patra, "Design and simulation of compact UWB Bow-tie antenna with reduced end-fire reflections for GPR applications," Int. Conference on Wireless Communication, Signal Processing and Networking (WiSPNET), Chennai, pp. 1786-1790, 2016.
- [22] Y. C. Wang, D. Yin and J.W. Liu, "A novel ellipse back-cavity and slot-tuned bow-tie antenna for ground penetrating radar," 16th Int. Conference on Ground Penetrating Radar (GPR), Hong Kong, pp. 1-4, 2016.
- [23] J. Shao, G. Fang, Y. Ji, and H. Yin, "Semicircular slot-tuned planar Half-Ellipse antenna with a shallow Vee-Cavity in vital sign detection," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, pp. 767-774, 2014.
- [24] H. M. P. B. Ranasinghe, S. M. P. Senanayake, U. I. P. Senarathne, A. U. A. W. Gunawardena and D. N. Uduwawala, "Design of a low-cost cavity backed wideband bow-tie antenna for ground penetrating radar systems" IEEE 8th Int. Conference on Industrial and Information System, Peradeniya, pp. 370-375, 2013.
- [25] L. Yang, H. Guo, X. Liu, H. Du and G. Ji, "An antipodal Vivaldi antenna for ultra-wideband system," IEEE Int. Conference on Ultra-Wideband, Nanjing, pp. 1-4, 2010.
- [26] R. Natarajan, J. V. George, M. Kanagasabai and A. Kumar Shrivastav, "A Compact Antipodal Vivaldi Antenna for UWB Applications," IEEE Antenna and Wireless Propagation Letter, vol. 14, pp. 1557-1560, 2015.
- [27] X. Q. Yang, Y. Y. Zhai, C. Xu and G. Z. Jia, "Design of antipodal Vivaldi antenna with better performances for ultra-wideband applications," Int. Journal of Applied Electromagnetics and Mechanics, vol. 46, pp. 527-536, 2014.
- [28] Y. W. Wang, G. M. Wang and B. F. Zong, "Directivity Improvement of Vivaldi Antenna Using Double-Slot Structure," IEEE Antenna and Wireless Propagation Letters, vol. 12, pp. 1380-1383, 2013.
- [29] K. Ma, Z. Zhao, J. Wu, M. S. Ellis, and Z.-P. Nie, "A printed vivaldi antenna with improved radiation patterns by using two pairs of eye-shaped slots for UWB applications," Progress In Electromagnetic Research, vol. 148, pp. 63-71, 2014.
- [30] B. Panzner, A. Jöstingmeier and A. Omar, "A compact double-ridged horn antenna for ground penetrating radar applications," 18-th Int. Conference On Microwaves, Radar And Wireless Communication, Vilnius, pp. 1-4, 2010.
- [31] A. E. C. Tan, K. Jhamb and K. Rambabu, "Design of Transverse Electromagnetic Horn for Concrete Penetrating Ultrawideband Radar," IEEE Transactions on Antenna and Propagation, vol. 60, pp. 1736-1743, 2012.
- [32] C. A. Balanis, Antenna theory: Analysis and design. UK: John Wiley & Sons Ltd, 2005.
- [33] L. Chang, C. Liao, L. lu Chen, and X. Zheng, "Compact planar Ultra-wideband antenna with wireless local area network and X band notches," IETE Technical Review, vol. 29, pp. 76, 2012.
- [34] S. Agrawal, R. D. Gupta and S. K. Behera, "A hexagonal shaped fractal antenna for UWB application," Int. Conference on Communication, Dev. and Intelligent System (CODIS), Kolkata, pp. 535-538, 2012.
- [35] P. Cao, Y. Huang and J. Zhang, "A UWB monopole antenna for GPR application," 6th European Conference on Antenna and Propagation (EuCAP), Prague, pp. 2837-2840, 2012.
- [36] R. Azim, M. T. Islam, and N. Misran, "Printed circular disc compact planar antenna for UWB applications," Telecommunication System, vol. 52, pp. 1171-1177, 2013.
- [37] R. Kr, S. Gurjar, and M. Sharma, "Design of stair and slotted UWB antenna using stepped-feed with modified slotted ground plane," IICA Proceedings on National Seminar on Recent Advances in Wireless Networks and Communication, NWNC, pp. 22-25, 2014.
- [38] Y. Ranga, L. Matekovits, K. P. Esselle, and A. R. Weily, "Multioctave frequency selective surface reflector for Ultrawideband antennas," IEEE Antenna and Wireless Propagations Letters, vol. 10, pp. 219-222, 2011.
- [39] Y. Ranga, A. K. Verma, K. P. Esselle, and S. G. Hay, "An ultra-wideband quasi-planar antenna with enhanced gain," Progress In Electromagnetic Research C, vol. 49, pp. 59-65, 2014.
- [40] A. Kumar, and P. Kumar, Design of a quadruple band microstrip antenna for application in UWB region," Int. Conference on Innovation in Information, Embedded and Communication System (ICIIECS), Coimbatore, pp. 1-3, 2015.
- [41] N. A. Kumar, and A. S. Gandhi, "Small size planar monopole antenna for high speed UWB applications," Twenty Second National Conference on Communication (NCC), Guwahati, pp. 1-5, 2016.
- [42] C. Waghmare, and A. Kothari, "Spanner shaped ultra-wideband patch antenna," Students Conference on Engineering and Systems, Allahabad, pp. 1-4, 2014.
- [43] H. M. Jol, Ground penetrating radar: Theory and applications, Amsterdam: Elsevier Science, 2009.
- [44] W. O. Coburn, "An ultra-wideband absorber backed planar slot antenna," 31st Int. Review of Progress in Applied Computation Electromagnetics (ACES), Williamsburg, pp. 1-2, 2015.
- [45] N. Kushwaha, R. Kumar, and T. Oli, "Design of a high-gain ultra-wideband slot antenna using frequency selective surface," Microwave and Optical Technology Letters, vol. 56, pp. 1498-1502, 2014.
- [46] N. Kushwaha, and R. Kumar, "High gain UWB antenna using compact multilayer FSS," IEEE Int. Microwave and RF Conference (IMaRC), Bangalore, pp. 100-103, 2014.

- [47] A. J. Fenn, and P. T. Hurst, *Ultrawideband phased array antenna technology for sensing and communications systems*, United States: MIT Press, 2015.
- [48] J. D. D. Kraus, and R. J., Marhefka, *Antennas: For all applications*, Boston: McGraw-Hill Higher Education, 2001.
- [49] D. Valderas, X. Chen, C. Ling, J. I. Sancho, and D. Puente, *Ultrawideband antennas: Design and applications*, London: Imperial College Press, 2010.
- [50] H. Nazli, E. Bicak, B. Turetken and M. Sezgin, "An Improved Design of Planar Elliptical Dipole Antenna for UWB Applications," *IEEE Antenna and Wireless Propagation Letters*, vol. 9, pp. 264-267, 2010.
- [51] W. Kang, K. Kim, and W. Kim, "Design of a 16-Port shared-arm dipole array for Monostatic imaging radar," *IEEE Transactions on Antenna and Propagation*, vol. 62, pp. 454-459, 2014.
- [52] M. Li, and K. M. Luk, "A Differential-Fed Magneto-Electric Dipole Antenna for UWB Applications," *IEEE Transactions on Antenna and Propagation*, vol. 61, pp. 92-99, 2013.

# Introducing Time based Competitive Advantage in IT Sector with Simulation

Rida Maryam

Department of Computer Science  
COMSATS Institute of Information  
Technology,  
Islamabad, Pakistan

Junaid Haseeb

Department of Computer Science  
University of Management and  
Technology,  
Sialkot, Pakistan

Muhammad Tayyab

Department of Computer Science  
University of Management and  
Technology,  
Sialkot, Pakistan

Adnan Naseem

Department of Computer Science  
COMSATS Institute of Information  
Technology,  
Islamabad, Pakistan

Khizar Hameed

Department of Computer Science  
University of Management and  
Technology,  
Sialkot, Pakistan

Babar Shahzaad

Department of Computer Science  
University of Management and  
Technology,  
Sialkot, Pakistan

**Abstract**—Incompletion of projects in time leads to project failure which is the major dilemma of the software industry. Different strategies are used to gain a competitive advantage over competitors in business. In software perspective, time is an incredibly critical factor, software products should be delivered in time to gain competitive advantage. However, at a halt, there is no such strategy that covers time perspective. In this paper, a time-based strategy for software products is introduced. More specifically, the importance of time-based strategy by analyzing its associated factors is highlighted using simulations.

**Keywords**—Business strategy; competitive advantage; time-base; a competitor; simulation; software industry

## I. INTRODUCTION

A strategy is responsible for designing plan of actions and assigning required resources to achieve long-term goals of an organization [1]. The strategy is viewed as the process of creating a unique and valuable position by means of a set of activities in a way that creates synergistic pursuit of the objectives of a firm [2]. In terms of its importance, strategy helps to gain substantial advantages and it is considered as a vital source for generating favorable situations between the firm and its competitors [3], [4]. The strategy is a pattern of resource allocation that enables firms to maintain or improve their performances [5]. Identify trends and opportunities in the future. The firm can strive to gain competitive advantage over its competitors only when it maintains a difference with competitors [3]. Entire visualization of the firm is created by business strategy. The business strategy describes the internal and external condition of firm required for competing with competitors. It is crucial that goals and missions of organization would be clear to everyone. The strategy provides help to stable firm's goal.

Since the 1980s, competitive advantage is the most significant concerns of business administration. In a business perspective, competitive advantage is described as an attribute and unique features through which an organization outperforms its competitors for the targeted market [6]. Researchers have

different opinions about the concept of competitive advantage and widely studied in [3], [7]-[9] to analyze the firm's performance. Competitive advantages can take a number of perspectives; these can include organizational structure and process [10], knowledge and capital derived from employees [11], all of which constitute resources residing internally within the firm. There are three basic categories of competitive strategies [12] that can be applied by companies in order to achieve sustainable competitive advantages: low-cost leadership, Differentiation, and focus. The differentiation and cost leadership are two major strategies to compete with opponent firms [6], [13]-[15]. Finally, by forming a business strategy a firm can achieve competitive advantage competitive advantage and eventually realize more about their current and future situations.

In time-based competition, time is a critical source and most important factor for gaining competitive advantage in a worldwide context. "Time-based competition will be the rule of the day" [16]. Strategic timing is the primary choice of the firms, to become the first, second or last move to the market [17]. The purpose of the time-based strategy is to trim down the time for the completion of the task. "Time-based Competitors are offering greater varieties of products and services, at lower costs and in less time than are their more pedestrian competitors" [18]. Researchers, practitioners, and companies demonstrated through case studies, surveys and empirical approaches that the business and IT (Information Technology) performances are tightly coupled [19]-[27] and enterprises cannot be competitive if their business and IT strategies are not aligned. Time consideration in IT firms is even more curious. According to the Chaos report, only 16.2% projects can complete on time, remaining may fail due to the delay in completion time [28]. As time is a very critical factor in the production of software, in this paper, a time-based strategy for software products is introduced. Multiple factors associated with time perspective are identified and finally, the positive and negative effects of these factors are analyzed using simulations so that importance of time-based factors can be

highlighted. Time base simulations will provide a systematic view, when and how to launch a software product to get the maximum competitive advantage either it should be in time, pre-time or post time launch.

Remaining paper is structured as follows: Section II comprises extensive background knowledge. Section III discusses the importance of time-based competitive advantage. Simulation results are described in Section IV. Section V offers a conclusion and future work.

## II. BACKGROUND KNOWLEDGE

Business competitive strategies were used to improve the business performance and to gain a competitive advantage for the firm. A survey of the literature was conducted comprising business strategies, author and year along with their main purpose as shown in Table 1.

The researchers proposed the different competitive strategies that can be used globally in a different context of business. There were three basic strategies introduced by Porter in 1980 i.e. Cost leadership, Differentiation and Focus. With the passage of time, some other strategies were added into and derived from basics ones. Customer oriented and market-oriented strategies were used in 2006 for building strategies to fulfill customer and market need. Market differentiation is a sub-strategy of differentiation used in 2007 for producing a unique product in term of marketing. Similarly, quality differentiation, service differentiation, innovation differentiation was used in 2007, 2014, 2015 and onward for building service wise and quality wise unique products. Then in 2016, innovation strategy was further extended as product innovation and process innovation for the introduction and implementation of innovative product and process respectively.

Contemporary studies proved that time is a critical factor but in existing studies, time-based factors are merged with other strategies as sub factors. So, there is a need of time-based strategy for gaining a competitive advantage against competitors.

## III. IMPORTANCE OF TIME-BASED COMPETITIVE ADVANTAGE

Traditional business strategies i.e. Differentiation, Cost leadership and/or focus consider “time” as a subfactor having less influence in competitive advantage. It may or may not be a subfactor in businesses other than IT but here it is a critical one. Based on the literature, it is found that time is a critical factor so a trinity of factors that can influence performance is presented in Fig. 1. A company can gain its competitive advantage against its competitor by focusing on differentiation, time, and cost. A company can use a strategy to manage low cost to get competitive advantage, it may invest more money to make a unique product, if it is investing money to deliver a unique product by utilizing Differentiation strategy and launching it without focusing on time then there is a chance that its competitor might take the competitive advantage by launching a product in adequate time i.e. pre-time, in time or post time.

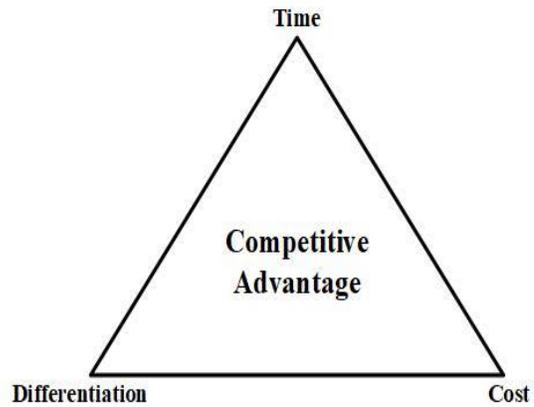


Fig. 1. Trinity of time, cost, and differentiation to gain competitive advantages.

Fig. 2 shows that individual competitive strategies have a positive impact on competitive advantage for example if a company used Differentiation as a business strategy, then it will enhance the Competitive advantage for itself. Same is the case for cost leadership (to minimize overall cost). If we talk about Time leadership or time-based strategy then it will impact on the competitive advantage positively by completing product within planned time.

The introduction to the market of the product depends upon competitor’s launch. This is because both competitors are aware of each other’s commercialization, launching time and features of the product by different means. The most common way is to track the events of commercialization organized by competitors to get updates about being a launching product.

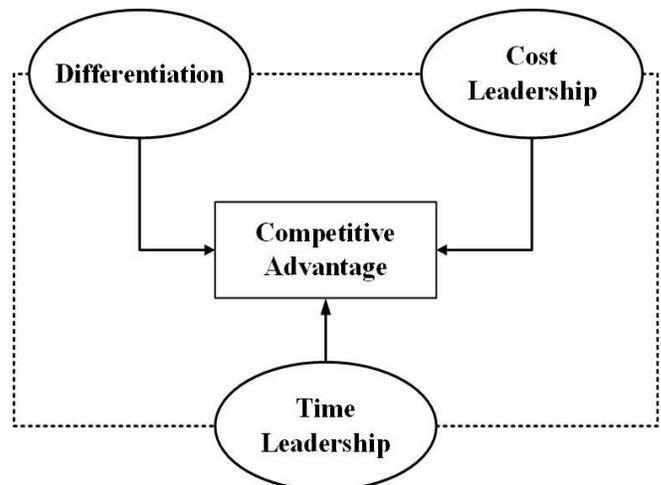


Fig. 2. Relation of strategies with competitive advantage.

Competitors add more features if needed to make the product more ideal to launch. If the features take more time than post-time is focused otherwise pre-time is suitable. In time product launch is ideal in the case of unique features are added in the product or the competitor is not launching its product at the same time.

TABLE. I. STRATEGIES WITH THEIR MAIN PURPOSE

<u>Author &amp; Year</u>	<u>Strategy</u>	<u>Main purpose</u>
John et al. 2006 [29], Amir et al. 2014 [30], Johnny 2008 [31], Audhesh et al. 2012 [32], Nadia & Lassaad 2015[33], Amir et al. 2015 [34], David et al. 2015 [35], Bulent et al. 2007 [36].	<b>Cost Leadership</b>	About reducing the cost to the organization of delivering products or services.
John et al. 2006 [29], Johnny 2008 [31], Audhesh et al. 2012 [32], Nadia & Lassaad 2015 [33], Devakumar & Barani 2016 [37], Jermias (2008) [38].	<b>Differentiation</b>	Making your products or services unique from or more striking than those of your rivals.
Audhesh et al. 2012 [32], Devakumar & Barani 2016 [37].	<b>Focus</b>	Focus on a narrow fragment and within that fragment try to accomplish either cost advantage or differentiation.
Gaurav & Himanshu 2016 [39], Wa'el 2015 [40].	<b>Customer Relation Management (CRM)</b>	Organizing a company's relations with existing and outlook customers for improving business relationships and retaining customers
John et al. 2006 [29]	<b>Customer Oriented</b>	Concentrate on fulfilling a customer's needs except only increasing profit.
Devakumar & Barani 2016 [37]	<b>Market Penetration</b>	Practical for successfully using your product, when company enters a new market Helpful to increase product demand and raise market outline.
John et al. 2006 [29], Bulent et al. 2007 [36]	<b>Marketing and Market Oriented</b>	Focus on how you increase sales by getting and keeping customers. Focus on the needs of the market.
Devakumar & Barani 2016 [37]	<b>Sale Service Support</b>	Provide maintenance after delivering the services.
John et al 2006 [39]	<b>Staff Development</b>	The Staff Development Strategy is an information strategy with supporting process documents for Staff Development topics
Jeffrey & Joohyung 2016 [31]	<b>Family Owned</b>	In which two or more family members are concerned and the mainstream of rights or control lay within a family. By using this strategy creativity, human resource efficiency, structural R&D factors, and cost in form of return and business growth may increases.
Daniel 2016 [41]	<b>Process Innovation</b>	Implementation of a new or considerably enhanced production or delivery processes (including significant changes in techniques, equipment and/or software).
Daniel 2016 [41], Prajogo (2016) [42]	<b>Product Innovation</b>	The introduction of a product or service that is innovative or considerably enhanced with respect to its uniqueness or future uses.
Amir et al.2014 [30], Amir et al. 2015 [34], David et al. 2015 [35], Bulent et al. 2007 [36], Martie & Denny 2011 [43].	<b>Innovation Differentiation</b>	Finding ways to optimize a precise set of differentiators that are most relevant to a specific set of needs.
Amir et al. 2014 [30], Amir et al 2015 [34].	<b>Quality Differentiation</b>	Used to differentiate product in terms of quality
Amir et al. 2014 [30], Amir et al 2015 [34].	<b>Service Differentiation</b>	Used to differentiate product in terms of service
David et al. 2015 [35], Bulent et al. 2007 [36], Martie & Denny 2011 [43].	<b>Marketing Differentiation</b>	Used to differentiate product in terms of marketing

Fig. 3. shows some common factors that affect all three business strategies to gain competitive advantage. If all these three strategies would apply to any product development than the result will be an ideal product i.e. a unique product is developing at low cost with minimum time.

A unique product with innovation will impact positively on Differentiation but it will take the time to market, more

maintenance will require, time is important here because uniqueness will take more time. A flexible process claims low cost while commercialization will increase the cost. A more suitable and concrete planning will aid time leadership along with time to market and market needs those also impact positively on it. Here, time to market has a positive impact on time leadership; therefore, we cannot neglect individuality of both.

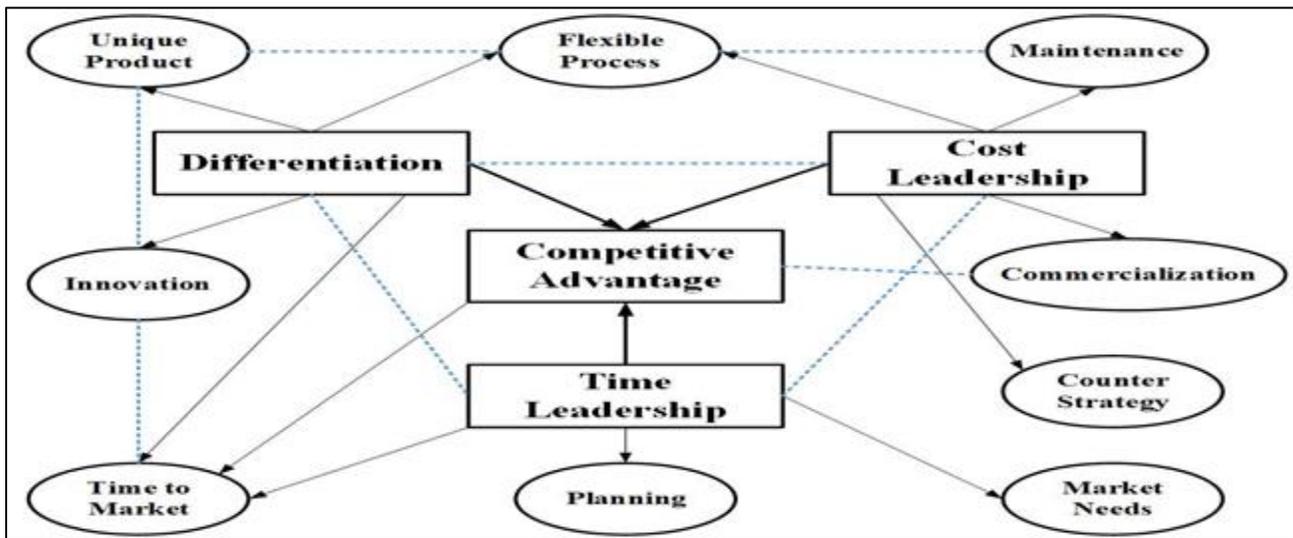


Fig. 3. Factors affecting strategies to gain competitive advantage.

#### IV. SIMULATION RESULTS

We only present simulation for limited scenarios here. Simulation results for more concrete and real scenarios will be investigated in the upcoming paper. According to the simulation results, right decision can be made by foreseeing the impact of the result in the future. To investigate the factors that have an impact on the competitive advantage with respect to time, an extensive literature was carried out. Most influencing factors are extracted and mapped with respect to time. Time-to-market of own versus a competitor, resources, HR skills, Project management, features of the own product, competitors and maintenance were the factors that affect competitive advantage.

The Likert scale was initially used for all factors to get the idea of winning competitive advantage of company A against its competitors more simply. Time to market was calculated for 2 companies to start the simulation; i.e. company A and its competitor. Then the difference was analyzed. If the company A has more week's spare to launch than its competitor and number of features it has built are more and/or unique than company A has more chances of winning competitive advantage by launching the product in time. On contrast, if it has low features and then it has to apply a sub-strategy to launch the product in post-time after completing the features, adding some more features and improving the limitations found in the recently launched product by the competitor. It is obvious that all the competitors keep track of each other's products updated by different means. If the competitor has more chances of launching the product in-time as it has a more positive value of time to market as compared to company A than company A has no other option unless it launches at post-time. Here, to apply another strategy of increasing resources, by simulating time difference and features differences but it will increase the cost. Now, let us simulate the impact of competitive advantage dependency. After performing simulations, generated results are presented in graphs to

represent the trends how different factors are influencing when a time-based strategy is introduced. Fig. 4. shows, when the time difference is greater than 0 and feature difference is equal or greater than 0, then competitive advantage will be greater than zero and product could be launched. Competitive Advantage could be negative as well if the three differences go overall negative. In that case, Time based sub-strategies are applied as shown in Fig. 5.

#### V. CONCLUSION AND FUTURE WORK

We tried to highlight time-based competitive advantage with more focusing on time to complete and launch the product on pre-time, in time or post time to get the maximum advantage from the competitor. In the literature, time is used as a sub factor, however, intentions were to highlight its primary importance for example, in Information Technology time, cost and quality are most critical factors for a product to complete and majority of the software does not meet deadline and hence, their competitor take the advantage by launching the product at adequate time. They attract the market towards them. The simulation was performed to check the influence of different factors on time to get a competitive advantage. Small strategies were applied to get the competitive advantage with respect to the competitors. Simple scaling and raw data were used which will be replaced by the more concrete scale and empirical study in coming paper. We are claiming on this result that time-based competitive advantage needs primary importance as a strategy along with cost leadership and differentiation. In future work, we also intended to introduce time-based competitive advantage strategy in local as well as in global software development context (GSD). In GSD, time zone difference is the main reason for increasing problem in synchronous communication and not completing the project with-in time that leads to the failure of the projects or causing a delay in the launching of the product. By introducing time-based strategy in global context, we may gain a competitive advantage in round the clock development as compared to our competitors.

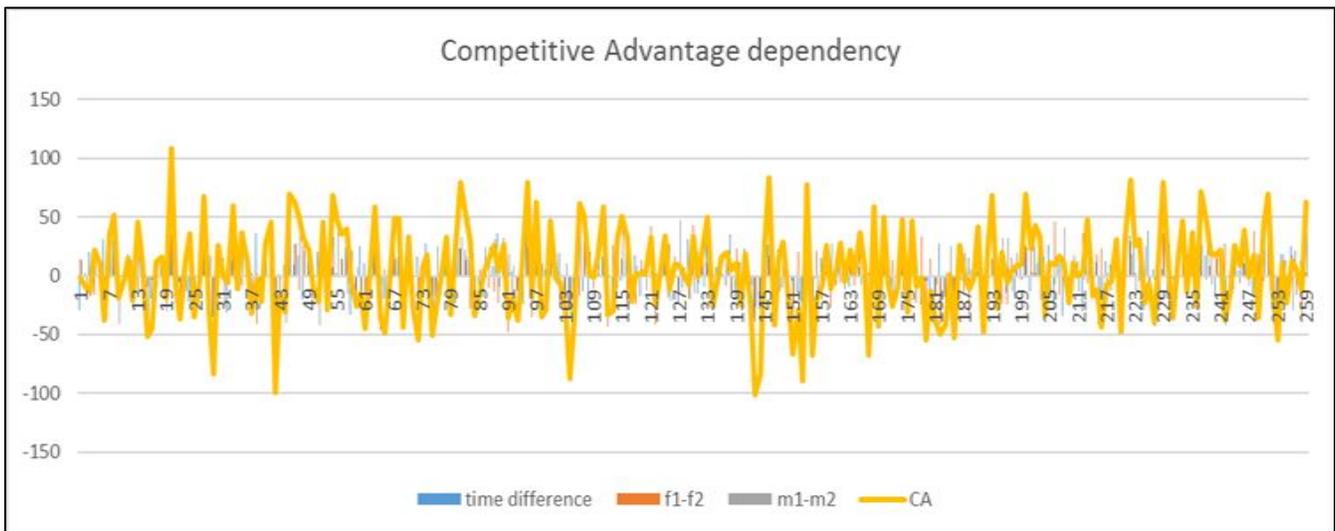


Fig. 4. Dependency of competitive advantage.

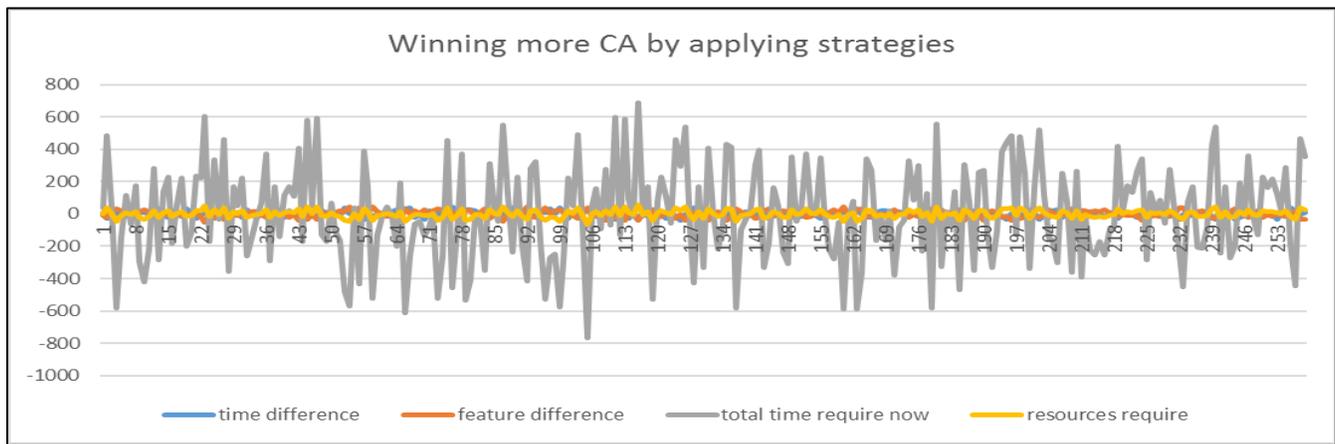


Fig. 5. Sub-strategies for winning competitive advantage.

REFERENCES

[1] Chandler et al., "Strategy and Structures: Chapters in the History of the Industrial Enterprise," MIT Press, Vol. 120, 1990.

[2] M.E. Porter, "What is Strategy?," Harvard Business Review, pp. 61-78, 1996.

[3] M.E.Porter, "Competitive Advantage: Creating and Sustaining Superior Performance," New York, The Free Press, 1985.

[4] A.Brandenburger and B. J. Nalebuff, Co-opetition, New York: Doubleday, 1996.

[5] J.B.Barney, "Gaining and Sustaining Competitive Advantage," U.S.A., Addison-Wesley, 1997.

[6] Wang et al., "Types of Competitive Advantage and Analysis," International Journal of Business and Management, vol. 6, no. 5, pp.100-104, 2011.

[7] A.C.Cooper et al., "Strategies of high performing new and small firms: A reexamination of the niche concept." Journal of Business Venturing, vol. 1, no. 3, pp.247-260, 1986.

[8] E.Mosakowski, "A resource-based perspective on the dynamic strategy-performance relationship: An empirical examination of the focus and differentiation strategies," Journal of Management, vol. 19, no. 4, pp. 819-839, 1993.

[9] D.Smallbone et al., "The Characteristics and Strategies of Fast Growth SMEs", International Journal of Entrepreneurial Behavior & Research, vol. 1, no.3, pp. 44-62, 1995.

[10] J. Barney, "Firm Resources and Sustained Competitive Advantage," Journal of management, vol. 17, no.1, pp. 99-120, 1991.

[11] N.Bontis, "Intellectual capital: an exploratory study that develops measures and models," Manage.Decision, vol. 36, no. 2, pp. 63-76, 1998.

[12] M.E.Porter, "Competitive Advantage". Ch. 1, pp 11-15, New York, The Free Press, 1985.

[13] R.Panwar et al., " The effect of small firms' competitive strategies on their community and environmental engagement," Journal of Cleaner Production, vol. 129, pp. 578-585, 2016.

[14] R.M.Beal and Y.Ardekani, "Performance implications of aligning CEO functional experiences with competitive strategies," Journal of Management, vol. 26, no. 4, pp. 733-762, 2000.

[15] P.S.Gopalakrishna, "Revisiting the pure versus hybrid dilemma: Porter's generic strategies in a developing economy," Journal of Global Marketing, vol.15, pp. 61-79, 2001.

[16] K.W.M.Tyson, "Competition in the 21st Century. New York, St Lucie Press, pp. 64, 1997

[17] M.A.Maidique and P.Patch, "Corporate strategy and technological policy". Management of innovation, pp. 273-285, 1982.

[18] G.Stalk and T.Hout, Competing against time. New York, Free press, 1990.

- [19] D.Avison et al., "Using and validating the strategic alignment model," *The Journal of Strategic Information Systems.*, vol. 13, no. 3, pp. 223–246, 2004.
- [20] S.J.Bleistein et al., "Validating strategic alignment of organizational IT requirements using goal modeling and problem diagrams," *Journal of Systems and Software*, vol. 79, pp. 362–378, 2006.
- [21] Society For Information Management, *IT Management Concerns Survey, What keeps CIO awakeatnight?*, 2006.
- [22] Y.E.Chan et al., "Business strategic orientation, information systems strategic orientation, and strategic alignment," *Information systems research*, vol. 8, no. 2, pp. 125–150, 1997.
- [23] J.Coughlan et al., "Understanding the business–IT relationship", *International Journal of Information Management*, vol. 25, no.4, pp. 303–319, 2005.
- [24] J.DeLeede et al., "Innovation, improvement and operations: an exploration of the management of alignment," *International Journal of Technology Management*, vol. 23, no. 4, pp. 353–368, 2002.
- [25] V.A.Hooper et al., "The impact of IS-marketing alignment on marketing performance and business performance," *The DATA Base for Advances in Information Systems, ACM SIGMIS Database*, vol. 41, pp. 36–55, 2010.
- [26] G.S.Kearns and A.L.Lederer, "Are source based view of strategic IT alignment: how knowledge sharing creates competitive advantage," *Decision sciences*, vol. 34, no. 1, pp. 1–29, 2003.
- [27] C.Pereira, P.Sousa, "Getting into the misalignment between Business and Infor-mation Systems": 10th European Conference on Information Technology Evaluation, Madrid, Spain. 2003.
- [28] Chaos, *The Standish Group Report*. 2014.
- [29] J. a Parnell, N. O'Regan, and A. Ghobadian, "Measuring Performance in Competitive Strategy Research," *International. Journal of Management and. Decision Making.*, vol. 7, no. 4, pp. 408–417, 2006.
- [30] A Samarrokh, K. Jenab, V. C. Arumugam, and P. D. Weinsier, "A new model for achieving sustainable competitive advantage through operations strategies in manufacturing companies," *International Journal of Logistics Systems and Management*, vol. 19, no. 1, pp. 115–130, 2014.
- [31] J. M. Campbell and J. Park, "Internal and external resources of competitive advantage for small business success: validation across family ownership," *International Journal of Entrepreneurship and Small Business*, vol. 27, no. 4, pp. 505–524, 2016.
- [32] A. K. Paswan, F. Guzm??n, and C. Blankson, "Business to business governance structure and marketing strategy," *International Journal of Business and Management*, vol. 41, no. 6, pp. 908–918, 2012.
- [33] N. Ben Romdhane Ladib and L. Lakhali, "Alignment between business model and business strategy and contribution to the performance: Empirical evidence from ICT Tunisian venture," *The Journal of High Technology Management Research*, vol. 26, no. 2, pp. 168–176, 2015.
- [34] A. Samarrokh, K. Jenab, V. C. Arumugam, and P. D. Weinsier, "Analysis of the effects of operations strategies on sustainable competitive advantage in manufacturing systems," *International Journal of Industrial and Systems Engineering*, vol. 19, no. 1, p. 34, 2015.
- [35] M.S. David, C. Devece, C.L. Albert. "How information systems strategy moderates the relationship between business strategy and performance." *Journal of Business Research* vol. 68, no. 7, pp. 1592–1594, 2015.
- [36] M. Bulent, S. Auh, and E. Shih. "Transformational leadership and market orientation: Implications for the implementation of competitive strategies and business unit performance." *Journal of Business Research*, vol. 60, no. 4, pp. 314–321, 2007.
- [37] G. Devakumar and G. Barani, "Marketing strategies for competitive advantage: structural equation modelling approach on agricultural sector industry in South India," *International Journal of Business Excellence*, vol. 9, no. 2, pp. 225–239, 2016.
- [38] J. Jermias, "The relative influence of competitive intensity and business strategy on the relationship between financial leverage and performance," *The British Accounting Review.*, vol. 40, no. 1, pp. 71–86, 2008.
- [39] G. Gupta and H. Aggarwal, "Analysing customer responses to migrate strategies in making retailing and CRM effective," *International Journal of Indian Culture and Business Management* vol. 12, no. 1, pp. 92–127, 2016.
- [40] W. Hadi, "The Relationship between CRM Strategies Stage on Competitive Advantage: An Analytical Perspective," *International Journal of Business and Management*, vol. 10, no. 8, pp. 245–252, 2015.
- [41] P. I. Daneial "The strategic fit between innovation strategies and business environment in delivering business performance." *International Journal of Production Economics*, vol. 171, pp. 241–249, 2016.
- [42] D. I. Prajogo, "The strategic fit between innovation strategies and business environment in delivering business performance," *International Journal of Production Economics*, vol. 171, pp. 241–249, 2014.
- [43] M. L. Verreynne and D. Meyer, "Differentiation strategies in mature small firms – the impact of uncertain environments," *International Journal of Entrepreneurship and Small Business*, vol. 12, no. 3, p. 327, 2011.

# New Deep Kernel Learning based Models for Image Classification

Rabha O.Abd-elsalam\*, Yasser F.Hassan\*, Mohamed W.Saleh\*

\*Department of Mathematics and computer science  
Faculty of science, Alexandria University  
Alexandria, Egypt

**Abstract**—Deep learning system is used for solving many problems in different domains but it gives an over-fitting risk when richer representations are increased. In this paper, three different models with different deep multiple kernel learning architectures are proposed and evaluated for the breast cancer classification problem. Discrete Wavelet transform and edge histogram descriptor are used to extract the image features. For image classification purpose, support vector machine with the proposed deep multiple kernel models are used. Also, the span bound is employed for optimizing these models over the dual objective function. Furthermore, the comparison between the performance of the traditional support vector machine which uses only single kernel and the introduced models is worked out that show the efficiency of the experimental results of the proposed models.

**Keywords**—Deep learning; multiple kernel; support vector machine; image classification

## I. INTRODUCTION

Recently, deep learning techniques are used for solving many problems in different domains as a result of performing well when training the regression model in high-dimensional data. Deep learning techniques succeed in both, machine learning and traditional computer vision. But, the identification of the application condition is needed for the deep learning. Many researchers make different studies to discover the pros and cons of deep learning over other machine learning methods. The most direct form is making a comparison between deep learning architectures and support vector machine (SVM) in processing audio, images and videos. However, there are not enough studies to choose the parameters once using deep learning technique for regression and classification tasks [1]-[5].

In machine learning filed, kernel learning technique is an active research subject [6], [7]. Kernel principal component analysis (KPCA) and SVM are the most common methods relay on kernel techniques. These kernel approaches have been applied to different applications due to their good performance. Unfortunately, the performance of those approaches depends on the selected kernel [8-10]. Thus, different studies have been introduced to learn the best kernel for these approaches [11], [12].

Multiple kernel learning (MKL) has been suggested to state the limits of single fixed kernel techniques. Bach et al. introduced the first MKL formulation [11]. Recently, MKL has been developed for automated kernel parameter tuning. Its goal

is to learn a linear or convex combination of multiple regular kernels to define the best target kernel for the given application [13], [14].

Many algorithms for extended MKL methods have been introduced to enhance the performance of the regular MKL method. In some real applications, MKL methods do not always yield better experimental performance once they compared with the regular techniques. Therefore, the deep learning architectures [15]-[17] are very promising choices than the shallow one. Furthermore, they can be used for feature extraction and in kernel applications as classifier (multilayer of multiple kernels learning (MLMKL)) [10].

The authors in [18] introduced a novel kernel which mimics the deep learning structure. They obtained static network where fixed kernels are used without learning the optimal kernels combination. A general framework for MLMKL is proposed in [10]. The authors had some problems in network optimization beyond two layers. The second layer only contains a single radial basis function (RBF) kernel.

The authors in [19] optimized the MLMKL with several layers and they used the leave-one-out error estimation algorithm. Unfortunately, their method is not evaluated over the MKL. Furthermore, no enhancements were achieved when using more than two layers.

In this paper, three models for deep kernel learning (DKL) are proposed and evaluated in the breast cancer classification problem. Additionally, span bound is exploited for the sake of optimizing the proposed models over the dual objective function. A comparison between the performance of the regular SVM using single kernel and the proposed DKL models is introduced.

The paper is organized as follows. The multi-layer multiple kernel deep learning is briefly described in Section 2. While Sections 3 and 4, introduce the methodology and the proposed deep kernel models. The experimental results are explained in Section 5. Section 6 concludes the work and presents the future work.

## II. MULTI-LAYER MULTIPLE KERNEL DEEP LEARNING

### A. Multiple Kernel Learning

Suppose that  $\{(a_1, b_1), \dots, (a_m, b_m)\}$  are  $m$  training samples where  $a_j \in \mathbb{R}^d$  is the feature vector of the sample and  $b_j$  is the sample label. The problem of MKL is generally described as the constrained optimization problem [10], [11]:

$$\begin{aligned} \min \quad & \lambda \|f\|_{H_k} + \sum_{j=1}^m \ell(b_j f(a_j)), \\ \text{s.t} \quad & k \in \mathcal{K}, f \in H_k \end{aligned} \quad (1)$$

Where  $\ell(\cdot)$  refers to some loss function like  $\ell(u) = \max(0, 1 - u)$  that used in SVM,  $\lambda$  is the regularization parameter,  $\mathcal{K}$  represents the candidate kernels optimization domain, and  $H_k$  is the reproducing kernel Hilbert space related to the  $k$  kernel.

In (1), the decision function  $f(a)$  is in the form of linear expansion of kernel evaluation on the training samples  $a_j$ ,

$$f(a) = \sum_{j=1}^m \alpha_j k(a_j, a) \quad (2)$$

Where  $\alpha_j$  are the coefficients referred to in [10].

In [10], the kernel is a set of convex combination of predefined base kernels:

$$\mathcal{K}_{conv} := \left\{ k(\cdot, \cdot) = \sum_{i=1}^n \rho_i k_i(\cdot, \cdot): \sum_{i=1}^n \rho_i = 1, \rho_i \geq 0, i = 1, \dots, n \right\}, \quad (3)$$

where each candidate  $k$  is the summation of the  $n$  base kernels  $\{k_1, \dots, k_n\}$ , and  $\rho_i$  is the coefficient of the  $i^{th}$  base kernel. So, the decision function can be expanded with the multiple kernels as:

$$\begin{aligned} f(a) &= \sum_{j=1}^m \alpha_j \sum_{i=1}^n \rho_i k_i(a_j, a) \\ &= \sum_{j=1}^m \sum_{i=1}^n \alpha_j \rho_i k_i(a_j, a), \end{aligned} \quad (4)$$

and the last kernel will be a linear summation of  $n$  base kernels.

### B. Deep Kernel Learning

Recently, many studies show that there is a limitation in conventional learning methods concerning their learning structural design. The deep structural design is often better than the shallow ones. The idea of deep learning of kernel methods that introduced in [19], [20] can be applied either in shallow structures such as SVM or in deep architectures.

The  $l$ -layer kernel is the inner product after several feature mapping of inputs:

$$k^{(l)}(a_i, a_j) = \underbrace{\langle \varphi(\varphi(\dots(\varphi(a_i)))) \rangle}_{l\text{-times}}, \underbrace{\langle \varphi(\varphi(\dots(\varphi(a_j)))) \rangle}_{l\text{-times}} \quad (5)$$

Here  $\varphi$  is the essential feature mapping function of  $k$  and  $\langle \cdot, \cdot \rangle$  represents the inner product.

Polynomial kernel is considered as an example of two-layer kernel, such as:

$$k^{(1)}(a, b) = (\delta(a, b) + \gamma)^d$$

$$k^{(2)}(a, b) = (\delta(k^{(1)}(a, b) + \gamma)^d, \quad (6)$$

Where  $\delta, \gamma$  and  $d$  refer to the free parameters of the polynomial kernel. The Gaussian RBF kernel can be approximated as:

$$\begin{aligned} k^{(1)}(a, b) &\approx k^{(2)}(a, b) = \varphi^{(2)}(\varphi^{(1)}(a)) \cdot \varphi^{(2)}(\varphi^{(1)}(b)) \\ &= e^{-2\lambda(1-k(a,b))}. \end{aligned} \quad (7)$$

The DKL has been suggested to use the deep learning idea for improving the MKL task.

A domain of  $l$ -level multi-layer kernels is defined as follows:

$$k^{(l)} = \{k^{(l)}(\cdot, \cdot) = \varphi^{(l)}([k_1^{(l-1)}(\cdot, \cdot), \dots, k_n^{(l-1)}(\cdot, \cdot)])\} \quad (8)$$

Where  $\varphi^{(l)}$  is a function to pool multiple  $(l-1)$  level kernels that should guarantee the valid resulting kernel.

The optimization problem of  $l$ -level MLMKL is described as:

$$\min_{k \in k^{(l)}} \min_{f \in H_k} \lambda \|f\|_{H_k} + \sum_{j=1}^m \ell(b_j f(a_j)) \quad (9)$$

## III. METHODOLOGY

The design of the image recognition system generally involves collection data, feature extraction, model selection or training, and evaluation. This part describes the design of the recognition system for the breast cancer classification problem in the digital image.

### A. Data Collection

The breast cancer databases are sets of mammograms images. This work used BCDR-F01 (Film Mammography dataset number 1) which is the first dataset of BCDR. The BCDR-F01 is a binary class dataset which composed by biopsy (Benign vs. Malign) [21].

### B. Features Extraction

Feature descriptors play an important role in recognition system. Really, they permit a mapping from visual information to a numerical vector which returns the semantic contents of the images. Regarding features extraction, this work used MPEG-7 edge histogram descriptor (EHD) [22] as input to train, evaluate and compare the proposed models and the traditional SVM classifier. EHD is used to refer the frequency and directionality of edges within each image region. Initially, simple edge detector operator is used for identifying edges and grouping them into five categories: horizontal, vertical, diagonal, anti-diagonal and non-edge. Then, global, local and semi-local edge histograms are calculated. The EHD features are represented by a vector of dimension 150.

Additionally, this work used discrete wavelet transform (DWT) to decompose an input digital image into four sub-bands of different frequencies [23]. The four sub-bands are generally denoted as approximation image (LL), horizontal (HL), vertical (LH) and diagonal (HH) detail components. The LL sub-band is used in this experiment to hold the most useful information of the input image.

### C. Classification

This work employed SVM for breast cancer classification, which is a two-class problem. SVM is a machine learning method that involves training and testing steps. With the two-class problem, training samples  $(a_j, b_j)$  are given, where  $a_j \in \mathbb{R}^d$  is the feature vector of the given sample and  $b_j$  is the label of its class, (+1 and -1 point to the two classes which are benign and malign classes respectively). SVM builds an optimal hyper-plane that maximizes the margin to classify the samples [24].

Traditionally, the margin is maximized through the gradient of the dual objective function with respect to the kernel hyper-parameters. But, the structures of deep learning give an over-fitting risk when richer representations are increased. So, looking for a tight bound of the leave-one-out error is needed. This paper used the span bound due to its promising results in single layer multiple kernel learning. The span bound is defined as:

$$T_{span} := L((a, b), \dots, (a_n, b_n)) \leq \sum_{p=1}^n \varphi(\alpha_p^* S_p^2 - 1) \quad (10)$$

Where,  $L$  points to the leave-one-out error and  $S_p$  refers to the distance between the support vector and the set  $A_p$  [25], [26] where:

$$A_p = \left\{ \sum_{i \neq p, \alpha_i > 0} \lambda_i \varphi_{k_\theta}(\mathbf{a}_i) \mid \sum_{i \neq p} \lambda_i = 1 \right\}. \quad (11)$$

### IV. PROPOSED MODEL

It is usually agreed that SVM is highly depend on the selected kernel function. In the regular SVM, the kernel function maps the input data, and then, the SVM is trained using this input data for the classification task. MKL is one probable structure, which designs the multiple kernels as linear combinations of base functions. Instead of using a single kernel function, a set of kernel functions can be organized in a particular structure to transform the original data over a number of layers of kernels. Then, the final kernel is used to learn the SVM decision function. The gradient descent presented in [19] is adopted in this work for optimizing the weights of the proposed deep kernels.

In this paper, three different models with different framework are considered for deep kernel learning architecture as shown in Fig. 1, 2 and 3 where the lines represent the weights for each kernel. Every model tries to optimize the weights of its architecture. The number of kernels in each layer is two in the three models. The first and third models have three layers, while the second model has only two layers.

The first model in Fig. 1 explores the combination of multiple kernels (two kernels). The elementary kernels in the first layer are computed from the given data and fed as input to the deep structure. The final kernel is learned as a three-multi-layered linear combination of functions where each one takes in a combination of two basic or two intermediate functions on multiple features.

In the second model shown in Fig. 2, the first and the second kernels ( $k_1^1, k_2^1$ ) in the first layer and the first kernel ( $k_1^2$ ) in the second layer transform the given input data. On the other hand, the second kernel ( $k_2^2$ ) in the second layer takes the linear combination of the output of the first and second kernels ( $k_1^1, k_2^1$ ) in the first layer. The final kernel is learned as a linear combination of the output of the two kernels in the second layer ( $k_1^2, k_2^2$ ).

In the third model shown in Fig. 3, the two kernels of the first layer map the original given data. While the first kernel ( $k_1^2$ ) in the second layer map the output of the second kernel ( $k_2^1$ ) in the first layer. The second kernel ( $k_2^2$ ) in the second layer maps the output of the first kernel ( $k_1^1$ ) in the first layer. The first kernel ( $k_1^3$ ) in the third layer maps the combination of the output of the two kernels ( $k_1^2, k_2^2$ ) in the second layer. While the second kernel ( $k_2^3$ ) in the third layer maps the combination of the original data and the output of the first kernel ( $k_1^2$ ) in the second layer. The final kernel is a combination of the output of the kernels ( $k_1^3, k_2^3$ ) in the third layer.

### V. EXPERIMENTAL RESULTS

The proposed DKL system has been implemented in MATLAB® 2015b with Windows 7 enterprise edition environment. The BCDR-F01 dataset is used to test and evaluate the performance of the DKL system in breast cancer classification problem. BCDR-F01 is a binary class dataset which composed by biopsy (Benign vs. Malign) [21].

The proposed models are tested on 86 images (29 benign images and 57 malignant images). The label of the benign class is +1 while the label of the malign class is -1. In the classification operation, 50% of images are used for training the classifier and 50% for testing the trained classifier; the images are randomly selected for training and testing stages. For the sake of comparing the performance of the proposed DKL models and the regular SVM, the performance is given in terms of accuracy which is the proportion of the correct classified samples to the total number of samples.

$$Accuracy = \frac{TP + TN}{N} \quad (12)$$

Where, TP is the number of true positive, TN is the number of true negative, and N is the total number of instances in the test set.

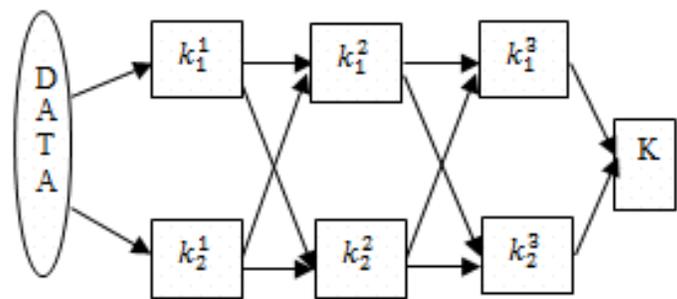


Fig. 1. DKL architecture for three layers with two kernels in each layer.

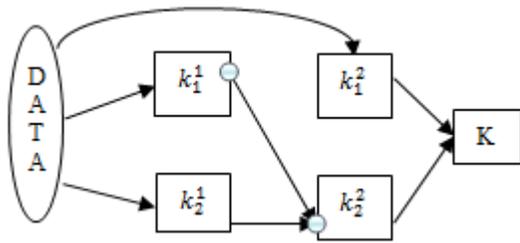


Fig. 2. DKL architecture for two layers with two kernels in each layer.

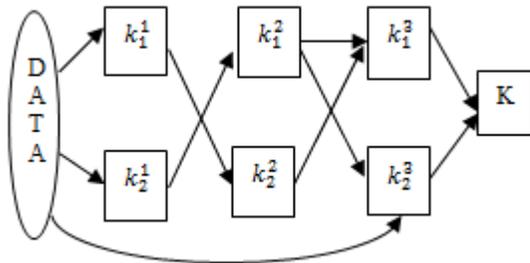


Fig. 3. DKL architecture for three layers with two kernels in each layer.

The same type of kernels in the architecture is used for those models (either all are RBF kernels or polynomial kernels). Multiple Kernels are considered for generating diverse representation of the data with basic functions. Deep learning structures present an over-fitting risk when richer representations are increased [19]. The over-fitting can be avoided by selecting a small number of base kernels. So, two kernels are used in each layer. Furthermore, the span bound is employed for finding a tight bound of the leave-one-out error. Span bound, presented promising results in [19], [26] with single layer multiple kernel learning over the dual objective function. In this paper, the gradient descent is used on the span bound for 100 iterations to DKL structure. The SVM penalty constant is fixed to 10 and the value of the learning rate is  $10^{-4}$ .

Table 1 illustrates the accuracy of the proposed three DKL models and the regular SVM, which uses only single kernel, using the feature extraction methods (EHD and DWT). DWT achieves better accuracy than EHD with the second and third DKL models when RBF kernel is used. But, EHD descriptor gives better accuracy than DWT with all models when polynomial (POLY) kernel is used. The third model achieves better results than other models with the two feature extraction methods.

POLY kernel gives better accuracy than RBF kernel in all models. When the DKL system has been tested with the POLY kernel, the first kernel in each layer is the POLY kernel with the degree of 2 while the second kernel is the POLY kernel with the degree of 5. The POLY kernel with the degree of 2 is already flexible to discriminate between the two classes with a good margin. Also, the POLY kernel with the degree of 5 yields a similar decision boundary. Model 3 achieves the best results among all models due to its deep architecture which can help to boost accuracy as shown in shaded cell in Table 1 (88%).

TABLE. I. A COMPARISON AMONG OUR PROPOSED DKL MODELS AND REGULAR SVM

Kernel Type	Feature Extraction	Accuracy			
		Classification Method			
		Model1	Model2	Model3	SVM
RBF Kernel	EHD	66.67	64.29	64.29	69.05
	DWT	66.67	76.19	76.19	66.67
Polynomial Kernel	EHD	76.19	85.71	<b>88.10</b>	73.81
	DWT	71.43	80.95	83.33	71.43

## VI. CONCLUSION AND FUTURE WORK

In this paper, three DKL models for breast cancer classification problem are introduced. Span bound is used for optimizing the proposed models over the dual objective function. A comparison between the performance of the regular SVM which uses only single kernel and the proposed models is introduced. The experimental results show that the proposed models overcome the traditional SVM. Furthermore, model 3 gives the best results among the other models due to its deep architecture that can help boost accuracy.

New features sets with another deep kernel structures will be explored on bigger dataset for the sake of determining which features set is the most discriminative with respect to breast cancer classification problem. Since the DKL has a bigger adaptability to data (because it's based on the creation of an optimal kernel to fit that data). These orientations will be the ultimate subject of the future work.

### REFERENCES

- [1] P. Liu, K.-K.R. Choo, L. Wang, F. Huang, Soft Computing (2016) 1-13.
- [2] G. Omer, O. Mutanga, E.M. Abdel-Rahman, E. Adam, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8 (2015) , pp. 4825-4840.
- [3] Y. LeCun, Y. Bengio and G. Hinton. "Deep Learning", Nature 521 (2015) , pp. 436-444.
- [4] W. Zhao, S. Du. "Spectral-spatial feature extraction for hyper spectral image classification: A dimension reduction and deep learning approach." IEEE Transactions on Geoscience and Remote Sensing 54 (2016) , pp. 4544-4554.
- [5] J.J. Tompson, A. Jain, Y. LeCun, C. Bregler, Joint training of a convolutional network and a graphical model for human pose estimation, Advances in neural information processing systems, 2014, pp. 1799-1807.
- [6] I. Rebai and Y. Belayed and W. Mahdi. "Deep multilayer multiple kernel learning". Neural Computing and Applications 27 (2016) , pp. 2305-2314.
- [7] J. Shawe-Taylor, N. Cristianini, Kernel methods for pattern analysis, Cambridge university press, 2004.
- [8] M. Wiering, M. Van der Ree, M. Embrechts, M. Stollenga, A. Meijster, A. Nolte, L. Schomaker, The neural support vector machine, BNAIC 2013: Proceedings of the 25th Benelux Conference on Artificial

- Intelligence, Delft, The Netherlands, November 7-8, 2013, Delft University of Technology (TU Delft); under the auspices of the Benelux Association for Artificial Intelligence (BNVKI) and the Dutch Research School for Information and Knowledge Systems (SIKS), 2013.
- [9] X. Xu, I. W. Tsang and D. Xu. "Soft margin multiple kernel learning". *IEEE Trans Neural Netw Learn Syst* 24 (2013) 749-761.
- [10] J. Zhuang, I.W. Tsang, S.C. Hoi, Two-Layer Multiple Kernel Learning, *AISTATS*, 2011, pp. 909-917.
- [11] F.R. Bach, G.R. Lanckriet, M.I. Jordan, Multiple kernel learning, conic duality, and the SMO algorithm, *Proceedings of the twenty-first international conference on Machine learning*, ACM, 2004, p. 6.
- [12] M. Hu, Y. Chen and J. T.-Y. Kwok, "Building sparse multiple kernel SVM classifiers," *IEEE Transactions on Neural Networks* 20 (2009) 827-839.
- [13] M. Jiu, H. Sahbi, Deep kernel map networks for image annotation, *Acoustics, Speech and Signal Processing (ICASSP)*, 2016 *IEEE International Conference on*, IEEE, 2016, pp. 1571-1575.
- [14] Z. Xu, R. Jin, I. King, M. Lyu, An extended level method for efficient multiple kernel learning, *Advances in neural information processing systems*, 2009, pp. 1825-1832.
- [15] G. Chen, C. Parada, G. Heigold, Small-footprint keyword spotting using deep neural networks, *Acoustics, Speech and Signal Processing (ICASSP)*, 2014 *IEEE International Conference on*, IEEE, 2014, pp. 4087-4091.
- [16] D. C. Cireşan, U. Meier, L. M. Gambardella and J. Schmidhuber, "Deep, big, simple neural nets for handwritten digit recognition", *Neural Computation* 22 (2010) 3207-3220.
- [17] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [18] Y. Cho, L.K. Saul, Kernel methods for deep learning, *Advances in neural information processing systems*, 2009, pp. 342-350.
- [19] E.V. Strobl, S. Visweswaran, Deep multiple kernel learning, *Machine Learning and Applications (ICMLA)*, 2013 12th International Conference on, IEEE, 2013, pp. 414-417.
- [20] Z. Sun, N. Ampornpunt, M. Varma, S. Vishwanathan, Multiple kernel learning and the SMO algorithm, *Advances in neural information processing systems*, 2010, pp. 2361-2369.
- [21] BCDR . "Breast Cancer Digital Repository," <http://bcdr.inegi.up.pt>, April 4, 2016.
- [22] N. Prajapati and A. K. Nandanwar. "Edge Histogram Descriptor Geometric Moment and Sobel Edge Detector Combined Features Based Object Recognition and Retrieval System," *(IJCSIT) International Journal of Computer Science and Information Technologies*, Vol. 7 (1) (2016) 407-412.
- [23] A. Goel. "Discrete Wavelet Transform (DWT) with Two Channel Filter Bank and Decoding in Image Texture Analysis." *International Journal of Science and Research (IJSR) Volume 3 (April-2014)*.
- [24] P. Liu, K.-K. R. Choo, L. Wang et al., "SVM or deep learning? A comparative study on remote sensing image classification," *Soft Computing*, 2016, p. 1-13.
- [25] B. Ghattas, A. B. Ishak, "An Efficient Method for Variables Selection Using SVM-Based Criteria.", (2005).
- [26] Y. Liu, S. Liao, Y. Hou, Learning kernels with upper bounds of leave-one-out error, *Proceedings of the 20th ACM international conference on Information and knowledge management*, ACM, 2011, pp. 2205-2208.

# Low-Power Hardware Design of Binary Arithmetic Encoder in H.264

Ben Hamida Asma<sup>1</sup>, Nedra Jarray<sup>1</sup>

<sup>1</sup> Laboratory of Electronic and Micro-Electronic  
(LAB-IT06)  
Faculty of Sciences of Monastir  
5019, Tunisia

Zitouni Abdelkrim<sup>2</sup>

<sup>2</sup> College of Education in Jubail,  
University of Dammam,  
Saudi Arabia

**Abstract**—Context-Based Adaptive Binary Arithmetic Coding (CABAC) is a well-known bottleneck in H.264/AVC, owing to the highly serialized calculation and high data dependency of the binary arithmetic encoder. This work presents a hardware architecture for the sub-module binary arithmetic encoder of the CABAC. Moreover, a clock gating technique is inserted into our design for power saving. An FPGA design of the proposed architecture can work at a frequency up to 268 MHz on Virtex 5. The suggested design can achieve 17% of power consumption saving, which allows it to be applied for low power video coding applications.

**Keywords**—H.264; Binary Arithmetic Encoder (BAE); Context-based Adaptive Binary Arithmetic Coding (CABAC); clock gating

## I. INTRODUCTION

In the H.264/AVC standard, two entropy encoders are defined: Context-based Adaptive Variable Length Coding (CAVLC) and Context-based Adaptive Binary Arithmetic Coding (CABAC). The CAVLC is a low-complexity entropy coding technique based on the use of switched context-adaptively sets of variable-length codes. Compared to CABAC, The compression efficiency improvement is obtained at the cost of an inevitable complexity overhead. Software-based complexity analysis results show that switching from CAVLC to CABAC usually leads to complexity increasing by 25–30% for encoding and 12% for decoding. As an average, 30–40 cycles are required to encode a single bit on digital signal processors, so it takes thousands of cycles to encode one macroblock, which is unacceptable for real-time video coding applications [2]. Therefore, a hardware implementation of CABAC encoder is always required. However, the bit-serial nature of the CABAC algorithm and the strong data dependency between contiguous bits make it hard to improve the throughput and to parallelize the encoding process.

Hence, a lot of work has been proposed to improve the throughput of the CABAC by processing more than one bin in a single cycle. Yuan Li et al. put forward in [3] a high-throughput low-latency arithmetic encoder (AE) design suitable for HD real-time applications, utilizing a macroblock level pipeline. This design could achieve a throughput of 2–4 bins per cycle sufficient for real-time encoding. In [4], a software-hardware codesign for a whole entropy coder was

suggested, which took Binary Arithmetic Encoder (BAE) module for the H.264/AVC CABAC entropy encoder as a hardware accelerator. Vagner Rosa et al. presented in [5] a hardware proposal of BAE. The throughput was improved by developing three different architectures of the renormalization step, presenting a processing rate from 0.68 to 1 bin per clock cycle. An RDO-support CABAC encoder was given by [6] and [7] to achieve the bit-rate saving of around 20 percent. In [6], an FPGA-based RISC CPU extension was proposed to accelerate the CABAC in a rate-distortion framework. This design achieved a coding speed of 1 bin per cycle and a clock frequency of 100 MHz. In [7], an efficient memory access was suggested to reduce the access frequency of the context RAM.

Most studies have mainly focused on ameliorating the throughput, but limited attention has been paid to reduce power consumption. Therefore, this paper aims to design BAE including a low-power technique. The main contributions of this paper are outlined as follows:

- 1) We implement a hardware of the BAE, which is the bottleneck of CABAC.
- 2) We further insert a low-power technique into the BAE architecture. In fact, a clock-gating technique is added into the design of a BAE sub-module, achieving reduced power consumption at a minor implementation effort.

The rest of this paper is organized as follows. Section II presents the CABAC encoding algorithm. Section III shows both encoding processes of the binary arithmetic coder and their corresponding proposed architecture. Section IV provides the FPGA synthesis results, and section V concludes the paper.

## II. CABAC ENCODING ALGORITHM IN H.264

As presented in Fig. 1, CABAC encoding consists of three main functions: binarization, context modeling, and binary arithmetic coding. The binarization part permits mapping the non-binary valued syntax elements into binary symbols, also known as *bins* or a *bin string*. Then each *bin* is arithmetically coded by a regular coding engine or a bypass coding engine. In the regular coding engine, a context model is used to encode each *bin*. In the bypass encoding engine, the context is not needed due to the equivalent probability of the appearance of these bins.

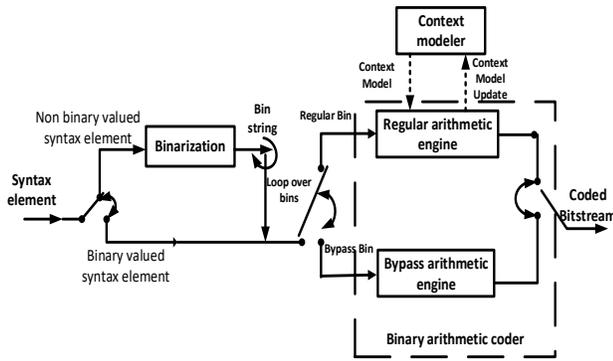


Fig. 1. Diagram of the CABAC encoder.

A. Binarization

In the binarization process, each syntax element is converted into a bin string. This step is done with different schemes: unary, truncated unary, fixed length and parameterized exp-Golomb. Each task is dedicated to some types of syntax elements, as given in Table 1. The input and output of the binarization process are the mapped syntax elements and the Context Index (CtxIdx) information. The next step is to use the CtxIdx information to fetch the context model from the context table.

TABLE. I. SYNTAX ELEMENTS AND ASSOCIATED TYPES OF BINARIZATION [1]

Syntax element	Binarization Method
<i>mb_type</i>	Table mapping
<i>mb_skip_flag</i>	Fixed length
<i>Sub_mb</i>	Table mapping
<i>Ref_idx_10</i>	Unary
<i>Ref_idx_11</i>	Unary
<i>mvd_10</i>	Truncated unary and exp-Golomb with =3 ,truncated value9
<i>mvd_11</i>	Truncated unary and exp-Golomb with =3 ,truncated value9
<i>Intra4x4_pre_mode</i>	Fixed length
<i>rem_intra_4x4_pre-mode</i>	Fixed length
<i>Chroma_pre_mode</i>	Fixed length
<i>Coded_block_pattern</i>	Fixed length and truncated unary, truncated value 2
<i>Mb_qp_delta</i>	Unary and table mapping
<i>Coded_block_flag</i>	Fixed length
<i>Significant_coefficient_flag</i>	Fixed length
<i>Last_significant_flag</i>	Fixed length
<i>Coeff_abs_level_minus1</i>	Truncated unary , exp-Golomb with =3 and truncated value 14
<i>Coeff_sig_flag</i>	Fixed length
<i>End_slice_length</i>	Fixed length

B. Context modeling

A context model is a probabilistic model with a statistical occurrence rate for each symbol, such that each type of syntax elements has a set of 399 context models as defined by the H.264 standard documentation [1]. Each context model comprises 6-bits representing the Probability State Indices (pStateIdx) and a 7th bit representing the value of the Most Probable Symbol (MPS).

C. Arithmetic coding

The aim of the arithmetic encoding process is to generate a bit stream from reading the bins and their context models, if the latter exist. Its principle is based on the division of an initial interval into two sub-intervals according to the context model (Fig. 2). One of two sub-intervals corresponds to the MPS, and the other refers to the Less Probable Symbol (LPS). After that, one of the two intervals is selected as a new one according to the bin value (MPS or LPS). Each interval is defined by two values: range (the length of the interval) and low (the bottom of the interval). These rules determine the updated value of the interval as follows:

If bin =LPS

$$\text{New range} = rLPS \text{ (range of LPS)}$$

$$\text{New low} = \text{low}$$

If bin= MPS

$$\text{New range} = \text{range} - rLPS = rMPS \text{ (range of MPS)}$$

$$\text{New low} = \text{low} + rLPS$$

Where, the value of *rLPS* is indexed by *pStateIdx*, read from context modeling.

For the bins that have the same probability, no context model is needed; and the bins are coded by a simpler bypass coding engine within a CABAC module.

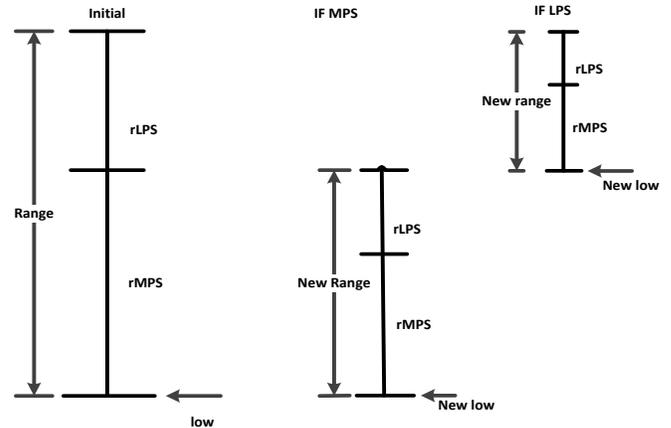


Fig. 2. Interval subdivision process of CABAC.

III. PROPOSED HARDWARE ARCHITECTURE OF BINARY ARITHMETIC CODER

At the binarization process, the syntax element of each MB can be treated in parallel. However, at the binary arithmetic coding process, all bin strings should be encoded sequentially. Thus, the binary arithmetic coder is the critical block that affects the throughput. This section firstly presents the processing flow of both regular and bypass BAE modes, and then its provides their corresponding hardware architectures. The clock gating technique is also presented in this section.

A. Regular BAE process and its proposed architecture

1) Regular BAE process

The regular BAE process is illustrated in Fig. 3. The chart consists of three steps: interval subdivision, probability-model updating and regular renormalization.

In the interval subdivision, the interval value is updated according to whether the current input bin (binval) is an MPS or not. The probability model  $pStateIdx$  is updated through two tables:  $TransIdxLPS$  and  $TransIdxMPS$ . The  $TransIdxMPS$  is selected when the bin value is equal to an MPS. Otherwise, the  $TransIdxLPS$  is used. The final update for low and range values is done by the regular renormalization process, which is needed to keep the interval range between 256 and 512. Fig. 4 shows the flowchart of the regular renormalization.

2) Regular arithmetic coder architecture

The hardware design of a regular BAE is depicted in Fig. 5. It consists of three main modules: probability-model updating, interval subdivision module, and regular renormalization.

The module of probability-model updating is constituted by three principal steps: context model read, context model update, and context read. When the context is read out, the context model will be updated according to bin value through the ROM of either  $TransIdxLPS$  or  $TransIdxMPS$ . Next, the new context model will be written back to the context table.

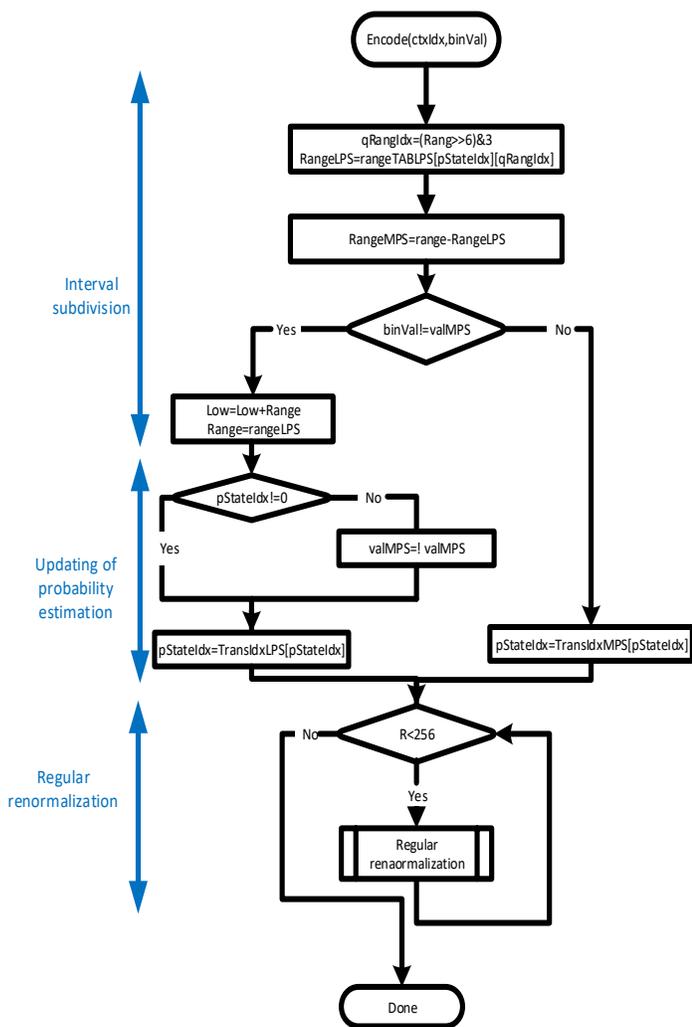


Fig. 3. Regular arithmetic encoding flowchart (from [1] with some modifications).

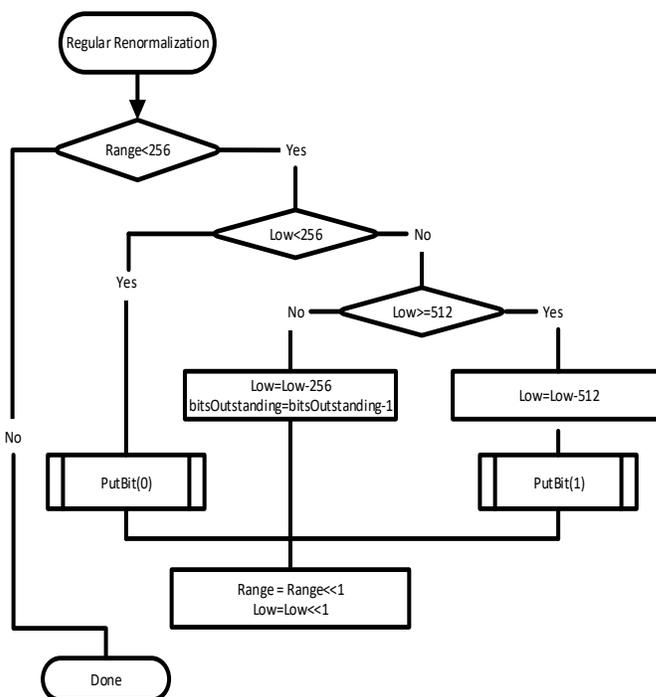


Fig. 4. Regular renormalization flowchart [1].

The module of interval subdivision will be performed when the context model  $pStateIdx$  is read from the RAM of the context table. Both  $pStateIdx[5:0]$  and  $range[7:6]$  are used to index the  $rLPS$  value from the  $rLPS$  table. After that, the interval values ( $range$  and  $low$ ) are calculated by using a ten-bit adder and ten-bit subtractor. According to the bin value, the two top and low multiplexers will select the appropriate value of  $low$  and  $range$ , respectively.

The module of regular renormalization will be carried after encoding each bin, when the range value is decreased to less than 256. This module is implemented by a finite state machine.

IV. BYPASS BAE PROCESS AND ITS PROPOSED ARCHITECTURE

A. Bypass BAE process

For the bypass mode, the bin is coded using a coding decision process. The context modeling is skipped as the bins show almost an equiprobable behavior. This encoding mode is a much simpler encoding process compared to the regular mode. Fig. 6 illustrates the bypass process, including the interval subdivision stage and the renormalization stage. There is no iteration loop in the renormalization process in the bypass mode unlike the renormalization in the regular coding mode.

B. Bypass BAE architecture

A hardware design of the bypass BAE is depicted in Fig. 7. Bypass coding generates valid coding states that conform to equations shown in the flowchart of Fig. 6. This mode is faster than the regular mode since there is no context modeling process. In addition, it is to note that there is no loop presented in bypass renormalization module. This latter is implemented by a simple finite state machine.

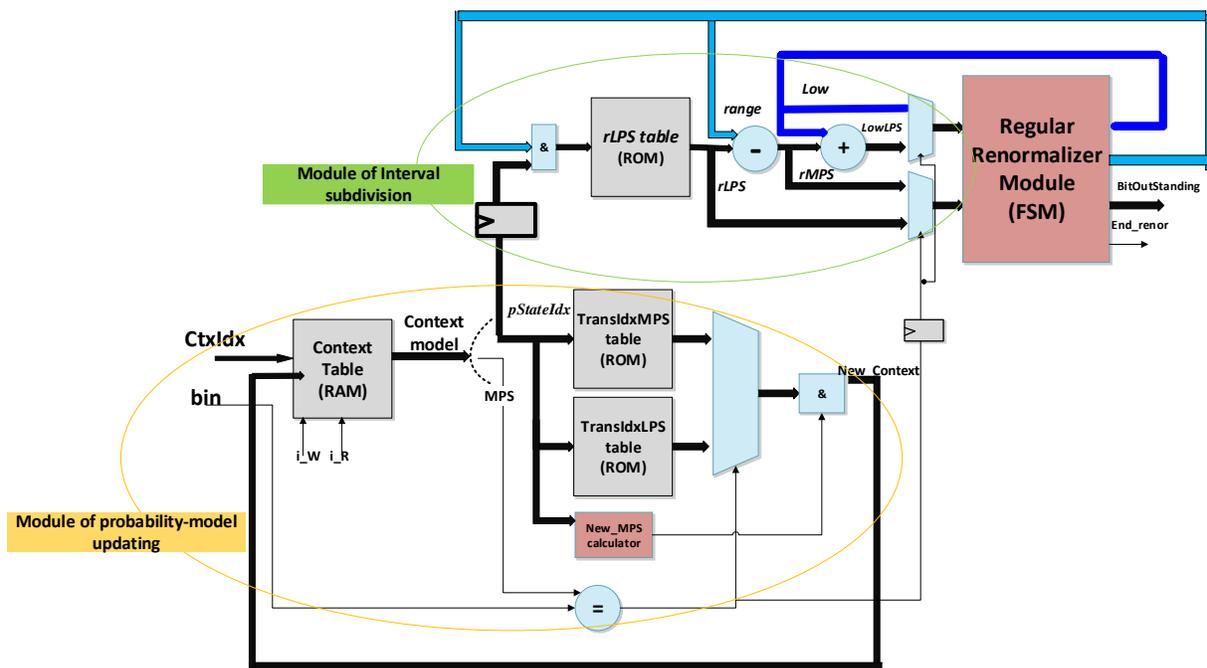


Fig. 5. Architecture representation of regular BAE.

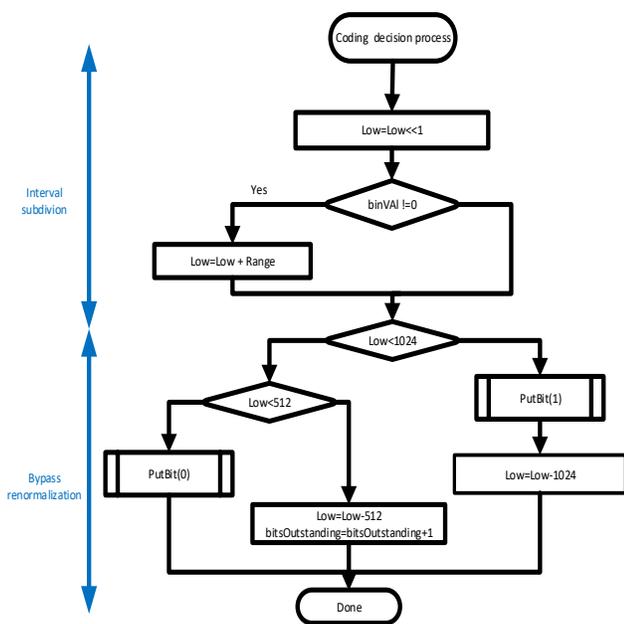


Fig. 6. Bypass BAE flowchart [1].

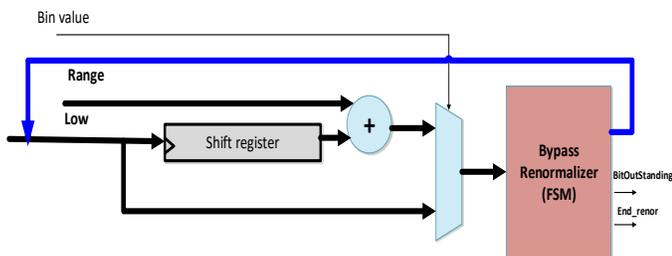


Fig. 7. Architecture representation of bypass BAE.

## V. CLOCK-GATING TECHNIQUE

Clock gating is among the techniques that are used for reducing dynamic power dissipation. This technique saves power by taking the enable conditions attached to registers and uses them to gate the clock.

At each input bin coming from the binarizer, one of the two coding modes (regular or bypass) is selected. The clock gating technique is inserted to prune the clock either for a regular arithmetic engine or for a bypass coder (i.e. by disabling the flip-flop registers in them).

The practical approach to insert the clock-gating technique in our proposed arithmetic coder is shown in Fig. 8. To avoid the glitch problem caused by clock switching, we use a latch-based clock-gating style.

## VI. IMPLEMENTATION RESULTS

Our design is synthesized and simulated by using the XILINX ISE and ModelSim tools, respectively. The synthesized circuit area of each component of the encoder is listed in Table 2. Synthesis results demonstrate that the BAE can work properly at a clock frequency of 268.5 MHz.

The design occupies 300 slices of which a regular BAE unit occupied 82%. It is to clear that the bypass BAE operates at a higher clock frequency compared to the regular mode.

Table 3 presents a comparison with previous work. Our design uses a higher frequency compared to the work [5], which was implemented in the same FPGA technology. Moreover, it is evident that the proposed architecture will achieve the lowest power consumption relative to power consumption of [7] when it is designed on ASIC-based technology. Indeed, as explained in [9], [10] and [11], the power consumption of ASIC designs was observed as being between 3 to 10 times greater than FPGA designs.

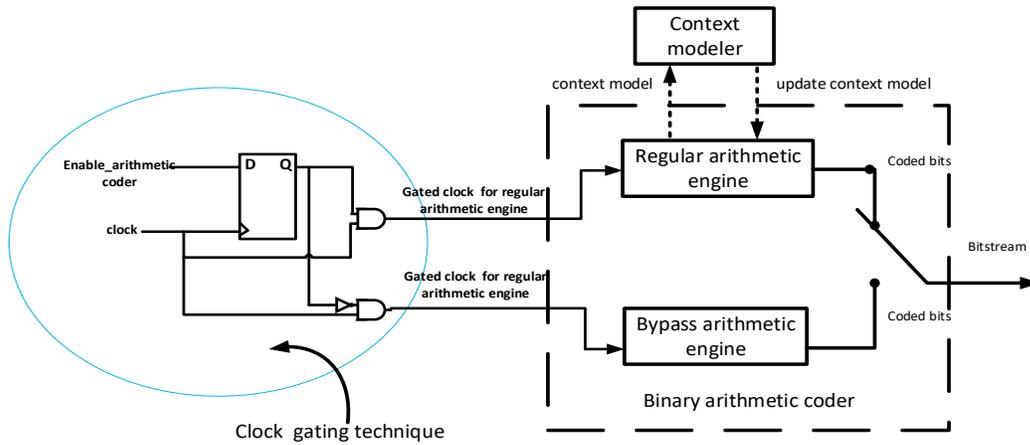


Fig. 8. Diagram of BAE with clock-gating technique.

TABLE. II. SYNTHESIS RESULTS OF EACH BAE UNITS ON VIRTEX 5

Unit Name	Area (slices)	Frequency (MHz)
Regular BAE	247	268.516
Bypass BAE	51	417.81
Total BAE (without clock gating)	298	268.516
Total BAE(with clock gating)	300	268.516

TABLE. III. COMPARISON OF PERFORMANCE RESULTS

	Process technology	Clock frequency (MHz)	Circuit Area (LUT slices)	Total power (mW)	Design parts
[5]	Virtex 5	189	436	Na	BAE
[6]	Startix II	130	603	Na	Total CABAC
[7]	Virtex4 FPGA	145	2559	Na	Total CABAC
	ASIC 0.13 $\mu$ m	200	Na	26.6	CABAC
[8]	ASIC 0.15 $\mu$ m	333	13.3K gates	Na	Total CABAC
Proposed	Virtex4 FPGA	219.479	298	43	BAE
	Virtex5 FPGA	268.516	300	17.77	

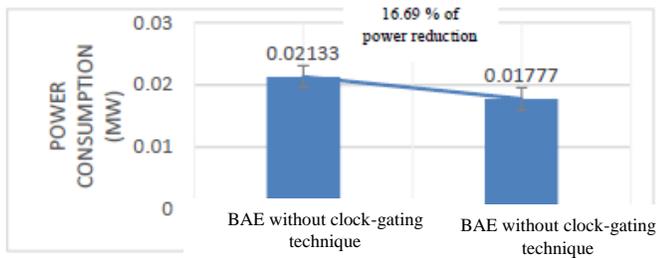


Fig. 9. Diagram of dynamic power consumption of our proposed BAE

Fig. 9 shows the power consumption for both designs (BAE without clock gating and BAE with clock gating). With the insertion of a clock-gating technique, there is about 17% of dynamic power consumption reduction.

## VII. CONCLUSION

In this paper, our design has focused on the BAE that presents the critical sub-block of the CABAC. Furthermore, a

clock-gating technique has been employed to reduce the power consumption. As a result, power consumption can be reduced by about 17%. Therefore, our design can be suitable for low power video coding applications. The synthesis results on Virtex 5 have indicated that the design is capable of operating at 268.516 MHz. Finally, it is important to mention that our BAE can fit both H.264/AVC and HEVC formats.

## REFERENCES

- [1] Joint Video Team (JVT) of ITU-T VCEG and ISO/IEC MPEG, Draft ITU-T Recommendation and final draft international standard of joint video specification ITU-T Rec. H.264/ISO/IEC 14496-10 AVC, May 2003.
- [2] Yu, Wei, and Yun He. "A high performance CABAC decoding architecture." IEEE Transactions on Consumer Electronics 51.4 (2005): 1352-1359.
- [3] Y. Li, S. Zhang, H. Jia, X. Xie, and W. Gao, "A high-throughput low-latency arithmetic encoder design for HDTV," in IEEE International Symposium on Circuits and Systems, 2013, pp. 998-1001.
- [4] R. R. Osorio and J. D. Bruguera, "A New Architecture for Fast Arithmetic Coding in H.264 Advanced Video Codec", 8th Euromicro Conference on Digital System Design (DSD'05)
- [5] V. Rosa, L. Max, S. Bampi, "High Performance Architectures for the Arithmetic Encoder of the H.264/AVC CABAC Entropy Coder", Electronics, Circuits, and Systems (ICECS), 2010 17th IEEE International Conference on
- [6] Y.L. Nunez-Yanez, V.A. Chouliaras, et al., "Hardware-assisted rate distortion optimization with embedded CABAC accelerator for the H.264 advanced video codec", IEEE Trans. CE, vol. 52, no. 2, pp. 590-597, May 2006.
- [7] X.H. Tian, T.M. Le SM-IEEE, X. Jiang, Y. Lian SM-IEEE, "A HW CABAC Encoder with Efficient Context Access Scheme for H.264/AVC", 2008 IEEE International Symposium on Circuits and Systems
- [8] C. Jian-long, L. Yu-kun and C. Tian-sheuan, member, IEEE, "a low cost context adaptive arithmetic coder for h.264/mpeg-4 avc video coding", IEEE international conference on acoustics, speech and signal processing, 2007. ICASSP 2007
- [9] A. Amara, F. Amiel, T. Ea, "FPGA vs. ASIC for low power applications", Microelectronics Journal, Vol. 37, p. 669-677, January 2006.
- [10] A. Chang and W. J. Dally. Explaining the gap between ASIC and custom power: a custom perspective. In DAC '05, pages 281{284, New York, NY, USA, 2005. ACM Press.
- [11] D. G. Chinnery and K. Keutzer. Closing the power gap between ASIC and custom: an ASIC perspective. In DAC '05, pages 275, 280.

# Dynamic Access Control Policy based on Blockchain and Machine Learning for the Internet of Things

Aissam OUTCHAKOUCHT  
PhD Student, Laboratory LISER  
IPI, Paris, France

Hamza ES-SAMAALI  
PhD Student, Laboratory LISER  
IPI, Paris, France

Jean Philippe LEROY  
Laboratory LISER  
IPI, Paris, France

**Abstract**—The Internet of Things (IoT) is now destroying the barriers between the real and digital worlds. However, one of the huge problems that can slow down the development of this global wave, or even stop it, concerns security and privacy requirements. The criticality of these latter comes especially from the fact that the smart objects may contain very intimate information or even may be responsible for protecting people’s lives. In this paper, the focus is on access control in the IoT context by proposing a dynamic and fully distributed security policy. Our proposal will be based, on one hand, on the concept of the blockchain to ensure the distributed aspect strongly recommended in the IoT; and on the other hand on machine learning algorithms, particularly on reinforcement learning category, in order to provide a dynamic, optimized and self-adjusted security policy.

**Keywords**—Internet of Things; security; access control; dynamic policy; security policy; blockchain; machine learning; reinforcement learning

## I. INTRODUCTION

Several works have dealt with the access control (AC) in the literature. Meanwhile, in constrained environments as the case in IoT, those concerns are not yet mature enough. This section is about introducing the IoT paradigm, basically from an AC point of view, and then will present how security policies are managed in the existing AC models.

### A. Internet of things paradigm

The Internet of things (IoT) is now a reality that surrounds us covering several parts of our lives, and will become more so in the future. Indeed, many researches consider IoT as one of the main technological revolutions of this century [1] and have moved from being a futuristic vision to an increasing market and research reality. It was in 2008 that the world passed the barrier of a single connected object per person and the statistics are now talking about numbers around 26 smart objects for every human being on earth by 2020 [2].

However, the Internet of Things, and despite all what has been said, is still maturing, in particular due to numerous challenges which slow down the full exploitation of the IoT, namely the computation constraints of the IoT devices, heterogeneity, identification, power supply, data storage/processing, etc. Meanwhile, one of the most crucial of these challenges concerns security and privacy, especially given the ubiquity of the smart objects in every corner of human life.

Unfortunately, what makes things worse; the traditional security solutions are not applicable in general in the context of IoT environments given the constraints of the IoT components which are characterized by low capabilities in terms of both energy and computing resources and thus, they cannot implement complex schemes supporting security. The OWASP Internet of Things Project has listed the most common IoT attacks and vulnerabilities [3]. According to this project, the risk arises because of the lack of adoption of well-known security techniques, such as encryption, authentication, access control and role-based access control. A reason for this lack of adoption is that existing security techniques, tools, and products may not be easily applied to IoT devices and systems.

To mitigate these risks, the deployed IoT services have to be “smart” and function in an open, dynamic and completely distributed environment. This requires that they gain a greater degree of autonomy and decision making.

### B. IoT and Access control

Authentication and access control technologies are known as the main elements to address the security issues in the Internet of Things. Actually, any effective access control system should satisfy the main security properties of the CIA triad: Confidentiality, integrity and availability. Note that one should not confuse AC with identification and authentication notions. Fig. 1 shows the boundaries of the access control process.

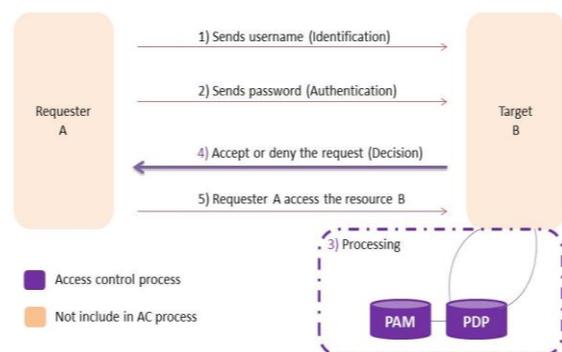


Fig. 1. Boundaries of access control.

Many access control models have been proposed in the literature to address security issues in IoT, but almost all of them are based on a centralized architecture, static security policy whose limitations in IoT context will be explained later.

As in the case of security mechanisms in general, applying current access control solutions on the device's side is not trivial. It requires intensive and computational capabilities which are not available in the most used IoT constrained devices. However, outsourcing the management of access control to non-constrained nodes presents serious security and privacy problems (e.g. break end-to-end security) and necessitates a high level of trust between the stakeholders. Furthermore, all interactions between them must be secured and mutually authenticated. To remedy what we have just cited, IoT needs an access control framework suitable to its distributed nature, where users may control their own privacy and, rather than being controlled by a centralized authority, and at the same time, the need arise for centralized entity handling authorization function to hardly constrained IoT devices.

### C. Security policy management

The Common Criteria defines an organizational security policy as: a set of security rules, procedures, or guidelines imposed (or presumed to be imposed) now and/or in the future by an actual or hypothetical organization in the operational environment [4].

Access-control policies have developed from trivial matrices to extremely complex representation expressed in sophisticated and advanced languages. It is then clear that this expansion and complexity require consequently robust automatic techniques to understand and manage them [5].

In traditional access control models access control policies are a set of rules stored somewhere in a server or, at best, distributed on several nodes in the network. In the case of the internet of the things it is necessary to have, in one hand a distributed policy that goes with the decentralized aspect of IoT and that is why (and for other reasons detailed later) we have chosen the blockchain technology as the basis of the proposed framework, and in the other hand a dynamic policy which takes into consideration the context in which the smart devices are, but also which can be improved over time, this improvement obviously does not, and cannot, be managed by a human being given the enormous and heterogeneous amount of data that the IoT generates. We therefore think in this paper to use the power of artificial intelligence algorithms, especially those of machine learning, to ensure this task.

## II. BACKGROUND

This section gives an overview of the basic concepts necessary to understand the proposed framework whether in terms of architecture or functioning.

### A. IoT and Machine learning algorithms

The internet of things is basically composed of various self-directed and low power devices. These nodes are able to collect information about their entourage with sensors, act on that environment (by using actuators) and communicate with each other and even with other entities like the Cloud.

The concept of machine learning (ML) was first treated as an artificial intelligence (AI) technique [6] then focused more and more in the complex algorithms that are difficult to manage by humans [7]. Nowadays, ML techniques are used in different domains and tasks including regression, classification,

speech recognition, fraud detection, and many others. Machine learning algorithms and techniques are inspired from several realms namely mathematics, neuroscience, statistics and computer science.

In general, machine learning algorithms are divided into two main steps: a training phase: the algorithm tries to learn based on the data; and a verification phase: the algorithm tests and tries to apply what is learnt.

The majority of the existing ML algorithms could be categorized in three classes: supervised, unsupervised and reinforcement learning [8]. The first class necessitates a labeled data set for the training phase in order to build a representation of the relations connecting the studied parameters. Unlike the first class, the unsupervised learning algorithms are not provided with input/output pairs. The emphasis here is mainly on classifying the data in different sets (clusters) by finding the connections between the given information. The third category, also known as the online learning, refers to process of handling the problems that an agent opposes when he must learn behavior via trial-and-error exchanges with an active environment [9].

Of course, some machine learning algorithms do not automatically fit into exactly one of these categories, there are some algorithms sharing features of both supervised and unsupervised learning approaches. The goal of these hybrid algorithms is mainly to benefit from the strengths of these two categories without inheriting their drawbacks [10].

Developing efficient algorithms that are suitable for many different application scenarios is a challenging task. Nevertheless, using reinforcement learning algorithms is the most suitable choice to solve the problem of static and non-contextual AC policies. Indeed, in our case it is sought that the algorithm must detect, progressively while accesses are made to resources and while the security policy is executing, the access control rules which are not optimal and even which present or lead to generate security problems. It is therefore an online learning.

### B. Blockchain concept

Originally introduced by Satoshi Nakamoto in 2008 [11] to underpin the Bitcoin cryptocurrency network, the blockchain is a computational paradigm that consists of a distributed ledger which contains all transactions ever executed within its network, enforced with cryptography and carried out collectively by a peer-to-peer nodes. Blockchains allow us to have a distributed peer-to-peer network where non-trusting members can interact with each other without a trusted intermediary, in a cryptographically verifiable manner.

Beyond the cryptocurrency field, blockchain is spreading over several other realms: Identity management [12], reputation system [13], storing system [14], IoT [15], access control [16], etc. Moreover, the continued integration of blockchains in the IoT domain will have a considerable impact on industry, home automation, healthcare, and so on.

Blockchain is a distributed database for transaction processing. All transactions in a blockchain are stored into a single ledger. The blockchain technology is built on top of four

fundamental building blocks, each building block has key properties, and each property is achieved through specific mechanisms:

1) Identifying the source and destination of a transaction: in a blockchain based ecosystem, users serve from digital identities called “addresses” to send and receive transactions. These addresses should be self-generated (independent from any given authority) and anonymous (reveal nothing about the real identity of its owner).

2) Transactions: A transaction records the transfer of a value (altcoin) from some source address to destination addresses. Transactions are generated by the sender and broadcasted the network of peers. Transactions are invalid unless they have been recorded in the public history of transactions, the blockchain. Note that these transactions are publically verifiable, furthermore, once a transaction is recorded in the blockchain it cannot be altered without that alteration being detected and rejected by the other nodes in the network.

3) Condition for auto-processing a transaction: The transfer of any value (e.g. altcoins, tokens) with the blockchain or the execution of any function through the blockchain should be locked by a logic conditions (e.g. low, contract) that have to be written as a code and automatically executed by nodes in the network. This condition should be self-executed.

4) Consensus: Every user or node in the network relies on algorithmically enforced rules to process transactions with no human interaction required to verify in an independent way the correct execution of the protocol, and obtains the same results. Each node has exactly the same ledger as all of the other users or nodes in the network. This ensures a complete consensus from all users or nodes in the corresponding currencies blockchain.

Fig. 2 shows the process of adding a transaction to the blockchain network. It gives an overview of the logic behind this technology in a five steps.

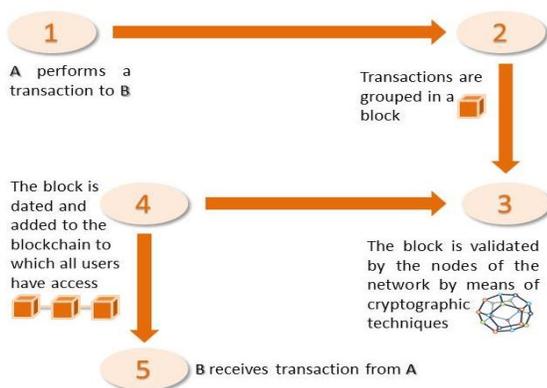


Fig. 2. Overall functioning of the blockchain concept

### C. AC and blockchain

An access control model is often used to rigorously specify and reason on the access control policy.

Providing an adequate access control model for IoT services is a vital but challenging topic. Actually, authentication and authorization concepts have been treated in numerous works. However in constrained environments, there is no considerable advancement. Moreover, IoT platforms need more and more dynamic, intelligent and fully distributed access control mechanism to be compatible with its nature.

The Blockchain applied to IoT provide a new world of promise and fascinating possibilities. Actually, the decentralization, automation, and trustless features of blockchain make it an ideal candidate to become a foundational element of IoT solutions.

One integration of the blockchain technology in the IoT is presented in [17] which consider that all the IoT devices of an organization work on the same blockchain network. The organization (or the device owner) deploys a smart contract that allows them to store the hash of the latest firmware update on the network. The devices can then query the contract, find out about the new firmware, and request it by its hash via a distributed peer-to-peer file system such as IPFS. Another approach to integrate the blockchain in the IoT is presented in the framework FairAccess [4] which will be detailed in section III.

In short, blockchain or distributed ledger technologies combined with IoT as underlying infrastructure can provide the next wave of innovation that streamlines the way business operates, the same way the web did, giving birth to a new collaborative economy [18].

### D. The need for a distributed AC architecture in IoT environments

The centralized approach consists in relieving smart device from the burden of handling a vast amount of access control-related information by outsourcing these functionalities to a back-end server or gateway which is responsible for security tasks. This approach presents many advantages: 1) possibility to reuse existing solutions and technologies; 2) authentication and access control policies are easier to manage in centralized IoT architectures. However, this approach presents several drawbacks: 1) prevent end to end security; 2) present single point of failure; 3) require trust foreign entities.

In distributed architecture, the access control process is carried out by the end component. This means that each device must be capable of handling authorization processes and have adequate resources to do so. In this proposal work, the concept of a distributed IoT is fundamental as a promising approach to release IoT. First of all, as devices increase their computational capacity, there are more opportunities to bring intelligence on the devices themselves. Moreover, this approach presents the following advantages: 1) end-devices act smartly, and are autonomous; 2) users have more control over the granularity of the data they produce as they are more enabled to define their own access control policies; 3) cost: it is less expensive than providing a cloud back end for each connected smart object; especially those that might need a connection for a decade; 4) trust could be supported in a better way with the decentralized approach than the centralized one because policies can be defined at the edge of the networks and there

will be no need to introduce any central entity; 5) This approach allows real time contextual information to become central to the authorization decision. However, the need to extend the constrained device with access control logic makes the implementation of this approach unfeasible in resource-constrained devices, and that is why going on with the totally distributed blockchain technology is strongly recommended.

### III. RELATED WORKS

Many access control models have been proposed in the literature to address security issues in IoT. Below is a summary of the most recent and relevant ones.

#### A. Role Based Access Control (RBAC)

RBAC (Role-Based Access Control) [19] refers to an access control model for governing user accesses to a system's resources, based on the notion roles. This model relies on four main blocks and each one of them provides RBAC with a number of features. These blocks are the core RBAC, the hierarchical RBAC, the static separation of duty relations and the dynamic separation of duty relations. The first block is composed of five components (users, roles, permissions, operations and objects). Roles and permissions are assigned, respectively, to users and roles. Moreover, there are two separate stages in RBAC: The design phase, where the administrator of the system or the security officer can describe a number of assignments between the system's components. The second phase (the run-time phase), that consists of enforcing the assignments in the system by the model as it is specified by the security policy, which was approved throughout the first phase.

#### B. Attribute Based Access Control (ABAC)

In ABAC model, accesses are allowed based on the notion of attributes. In fact, these later characterize every subject and object and identify them inside the system [20]. There is two parts that compose ABAC: The policy model and the architecture model which enforce this policy. ABAC model proclaims in his standard version that access are allowed according to the subject's attributes. Moreover, it is in the policy rules that conditions under which access is granted or rejected are defined. In ABAC model, the attributes are linked with the subject and the object features. Consequently, the user is given appropriate access control permissions suitable to his attributes at the time when he sends his access request to a given object. In the literature, several works using ABAC model have treated the AC from an IoT perspective: In 2014, Ye et al. [21] have proposed an authentication and access control model for the perception layer of the Internet of Things. The designed protocol provides low storage and communication overheads to deal with the constraints in resources of the IoT context, basically in perception layer. Furthermore, the fact that the model allows accessing the data according to user attribute guarantees fine-grained access control. Though, it necessitates on the other side complex management and slows down (even block) its large

deployment to constrained devices. Therefore, they only offer abstract outcomes of the proposed model.

#### C. Usage control (UCON)

Another famous AC model is the usage control (UCON) proposed in [22], it is considered as the next generation of access control models for the reason that it presents several novelties unavailable in traditional access control models such as RBAC and ABAC. It deals with the problems generated in the authorization phase, before the access execution, after the access execution, or even during the execution. In addition, it has the capability of supporting attribute's mutability; in other words, if a problem is produced in the security policy (during the execution) due to an alteration of some access attributes, the allowed access is canceled and the usage became invalid. Further information about UCON model is detailed in [23]. Many researches (like in [24]) have also applied UCON in collaborative system.

Before concluding this section, it is important to note that, in the state of the art level, there is a work that has stressed the particularity of UCON over usual AC models such as MAC, DAC and RBAC, and also that makes UCON more appropriate to the dynamic aspect of IoT, this work is exposed in [25].

#### D. Organization Based Access Control (OrBAC)

The OrBAC [26] model is one of the richest AC models in terms of components and applicability to many realistic situations; it was conceived to handle remaining issues in the extensions of RBAC. It presents an original dimension, namely the organizational concept; also it makes a clear distinction between the abstract level (roles, views, activities) and the concrete level (subject, object, action). In the decision making process, OrBAC takes into consideration various context information which can be temporal, spatial or declared by the subject (user). However, one of the big drawbacks of this model, especially when talking about IoT environments, is that it is based on a totally centralized architecture and does not provide or support the distribution, collaboration and interoperability requirements. That said, several works have done in order to extend OrBAC to overcome these limitation: PolyOrBAC [27] deals with this problem by using the OrBAC model to manage the internal policies of each organization, but to ensure the collaboration aspect between organizations, web services technology was. Nevertheless, such technologies that PolyOrBAC uses (e.g. SOA-based web services) are not systematically supported by IoT constrained nodes. To fix that, SmartOrBAC [28] and [29] objectives are to adapt OrBAC model to IoT situations. The major contribution of this proposition is the fact that it improves the notion of context (present in OrBAC) to respond to the IoT requirements. Unfortunately, SmartOrBAC does not precise any lightweight mechanisms to reduce the OrBAC complexity in order to be supported by IoT constraints devices.

The layers and components of OrBAC are shown in a simplified manner in the following Fig. 3.

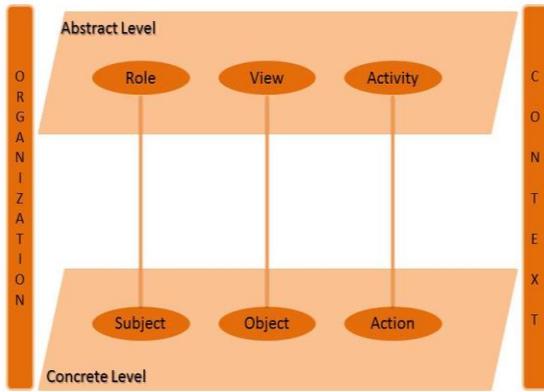


Fig. 3. Simplified presentation of OrBAC layers

#### IV. THE PROPOSED FRAMEWORK

In this section the focus is on the proposed framework that aims to solve the static and centralized problems of access control policy. This solution will be based on two essential concepts highlighted in the previous sections of this paper: Blockchain technology and machine learning algorithms.

##### A. Problem statement and research questions

Actually, most access control solutions today provide the ability for centralized authorities, whether governments, manufacturers or service providers to gain unauthorized access to and control devices by collecting and analyzing user’s data. That may cause ethical and privacy problems.

Section II-D has dealt with the problems coming from the centralized architecture. However, in IoT environments, a big obstacle blocks us from adopting a distributed architecture: The constrained devices generally used in the IoT do not have the capacity of calculation nor of storage to deal with a full distributed access control process where there is no central entity responsible for managing this latter. Therefore it is necessary to make a tradeoff between the two architectures or adopt a hybrid one as in SmartOrBAC. Except that fortunately the blockchain can respond with efficiency to this problem.

Another problem encountered by access control in the context of IoT is the difficulty of managing the security policy according to the contexts especially with the colossal number of smart devices supposed to be managed in IoT situations. This leads to adopt static policies where the manager or the security officer writes all the security or access control rules in a static manner. The major disadvantage of this approach is that this policy never detects if it contains rules that lead to security problems, which create conflicts or which are not optimal. This approach never takes into consideration feedback from the results of its operation. Moreover, given the number of the increasing number of smart objects, it is almost impossible to manage this policy manually in an efficient and totally personalized way.

This new framework responds to these problems, it gives people what properly belongs to them and also present an automatically-improved and dynamic security policy.

##### B. IoT-OrBAC

IoT-OrBAC access model, like SmartOrBAC [28], is specially conceived for the IoT context and it is designed through an abstraction layers’ perspective that makes use of a deep comprehension of the IoT paradigm as it is presented in the physical world. In the IoT that uses smart services as well as smart devices, contextual information is a key component in the decision making process, and only a real-time consideration of this information will reach smartness. In order to handle that, the authors improved the “context” concept (originally exposed in OrBAC) to fit the IoT needs. IoT-OrBAC separates the problem into several layers and then distributes processing charges between constrained devices and less constrained ones and at the same time addresses the collaborative aspect with a specific solution.

The layers IoT-OrBAC presented are shown in a simplified manner in the following Fig. 4.

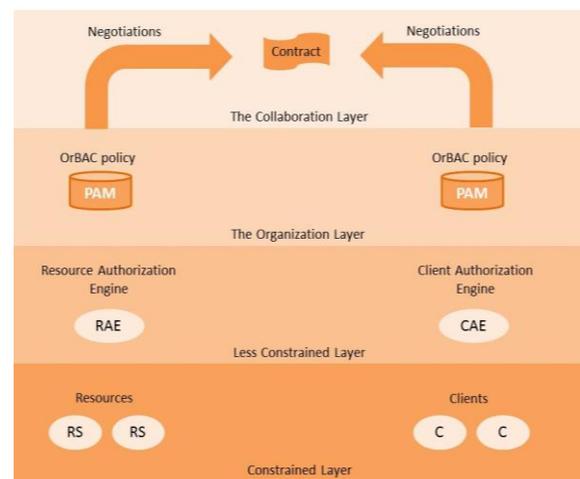


Fig. 4. Simplified presentation of IoT-OrBAC layers.

##### C. Fair Access

To explain how FairAccess works, let us take the following use case: Suppose that a subject (e.g., a requester device *A*, known with its address *rq*) wants to execute an action (e.g., read or alter) on a protected object (e.g., resource *B*, identified with its address *rs*). First, the subject must send this request to the authorization management point (AMP (AMP = wallet) which plays the role of a Policy Enforcement Point (PEP) that protects the requested object. The PEP formulates the received request to a *GetAccess* transaction. Then, the PEP broadcasts this transaction to the whole network of nodes with the aim of reaching miners, those later act as distributed Policy Decision Point (PDP), and accept or reject the transaction. The PDP evaluate the request and then it executes a SmartContract already deployed in the blockchain via a previous transaction called *GrantAccess*. The execution of SmartContract leads to decide whether the request should be permitted or not. Finally, if it is allowed, the SmartContract provide the requester with an access Token by sending it to his address through an *AllowAccess* transaction. After that, the Token will appear in the requester’s token database.

To summarize, and as it is shown in Fig. 5, authorization process in FairAccess framework consists of: 1) registration of a new resource with a corresponding address; 2) definition of an access control policies in the form of SmartContract deployed in the blockchain by a *GrantAccess* transaction; 3) access request; 4) access allowed; 5) access revoked/updated.

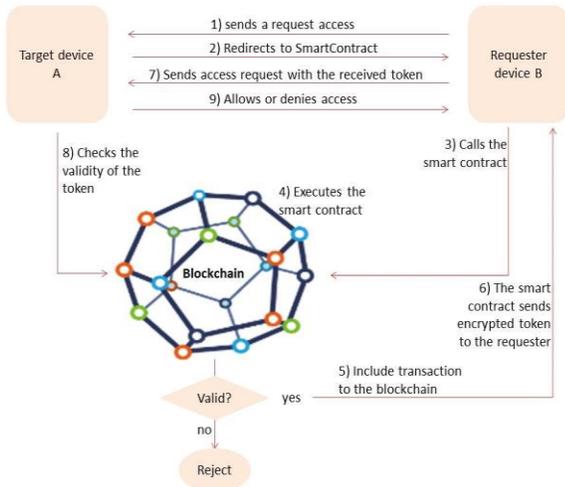


Fig. 5. Reload access control policies process in FairAccess.

#### D. Improve/upgrade security policy with ML algorithms

This framework uses the concept of SmartContract as a representation of an access control policy defined by a resource owner (RO), to manage access over one of his resources. It is a script stored on the blockchain. Since it resides on the chain, it has a unique address. This SmartContract is triggered by addressing a *RequestAccess* transaction type to it. It then executes independently and automatically in a prescribed manner on every node in the network, according to the data that was included in the triggering transaction. If the data fulfill access control policies, the *PolicyContract* will be correctly executed and then generates and assigns an Authorization Token to the sender of the *RequestAccess* transaction. For each end device, the RO defines one *PolicyContract* which is responsible for managing its access control functions.

Typically, the process of a classic reinforcement learning model begins by connecting an agent to its environment. Then, and in every interaction the agent receives some information (called feedback in this paper) about the present state of the environment; the agent then picks an action to make (output). The executed actions, obviously, updates the environment state and the value of this latter is transferred to the agent as a feedback. Note that the agent's behavior has to select actions in order to improve the situation of the environment especially in long term.

Formally, a typical RL model contains:

- A group of environment states,  $S$ ;
- A group of agent actions,  $A$ ; and
- A group of scalar reinforcement signals or weights if needed.

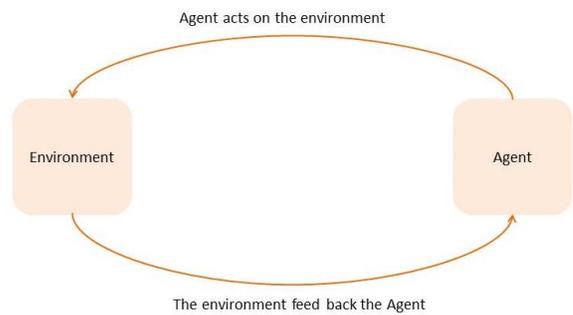


Fig. 6. Typical Reinforcement Learning (RL) scenario.

Moreover, it is important to emphasize that RL, generally, the fields of work of RL and the supervised learning (SL) are not the same (Fig. 6). Indeed, unlike SL, RL responds to problems where there is no arrangement of input/output information. Instead, the agent receives a reward (technically feedback information) after picking an action in a given state. It is indispensable that the agent passes several experiences to gather the maximum of rewards in order to know the best his environment and so he can make the right and optimal actions.

This proposed work relies on a fully distributed infrastructure based on blockchain technology as has been done in the work [4] of FairAccess. That said, it will use, as a cited before, the concept of SmartContract to distribute the security policy in the chain. Requesting access to an object will thus be managed by the detailed procedure of Fig. 5.

Once the IoT environment, now presented by the blockchain infrastructure, begins to function, it will send feedback information after each successful or unsuccessful transaction to the two entities involved in the communication, namely the subject (or requester) and the object (or resource). This information will be used as an evaluation of the transaction and its participants and will be taken into account to update the stakeholder data (e.g. update the trust, credibility or integrity levels of the participants) and also to update access control rules that allowed this transaction to be done.

#### E. Architecture

The procedure of this framework is detailed in the following organogram presented in Fig. 7.

Note that even if the feedback information are sent only to A and B, these information are spread to the blockchain after the update of SmartContract for example, so they become public.

Using Reinforcement Learning (RL) algorithms, the policy (the smart devices also) is trained to make particular decisions. It works this way: the SmartContract (which presents the access control policy of the smart object) is exposed to an environment where it trains itself continually using trial and error. Consequently, this SmartContract learns from past experience and tries to capture the best possible knowledge to make accurate business decisions.

The feedback information is represented by a vector with  $n$  number of components. These later may be binary represented (0 or 1) or with a level or weight of satisfaction (trust, integrity ...) between 0 and 1.

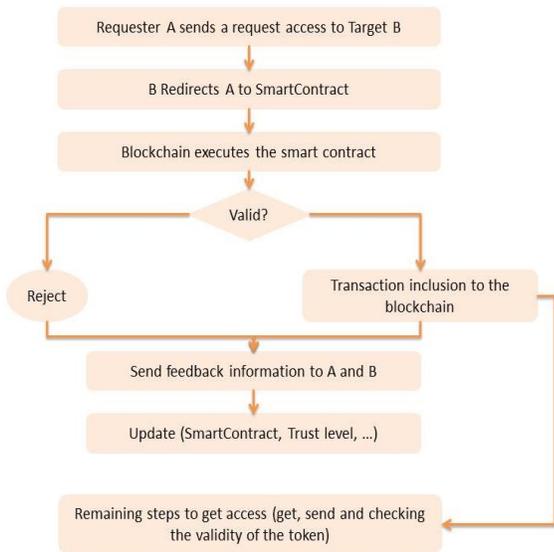


Fig. 7. Process organogram of our framework.

Let's take the very basic example: an organization H has some smart objects ( $O_1, \dots, O_n$ ), it adds them to its security policy, it creates then and publishes their SmartContracts in the network of the blockchain. Suppose that the subjects  $S_1, \dots, S_m$  have access to the objects  $O_1, \dots, O_n$  based on the rules contained in the SmartContracts.

Let us suppose that after each use of an object  $O_i$  by the subject  $S_1$  (which is a legitimate user), a breakdown is noticed in the object  $O_i$ , the feedback of  $O_i$  will be zero or a low note in the credibility scale given to  $S_1$ . The algorithm will determine, especially if it is repeated, that  $S_1$  will have to be removed from authorized users, and the security policy will update.

Another case, which does not concern the credibility/trust level of users is: Let us imagine that an object  $O_1$  frequently encounters problems when used in a given context  $C_i$ . The low notes coming from several transitions under the  $C_i$  context lead to detect the source of the problems and thus update the SmartContract by prohibiting the use of the object  $O_1$  under the context  $C_i$ .

Note that this framework is not limited to a specific access control model, that is why the components of the feedback vector are left open, they can include the context as seen before (e.g. OrBAC, SmartOrBAC), but also The attributes (ABAC), the level of credibility (I-OrBAC [30]).

Note also the choice of a category of algorithms (Reinforcement Learning) and not a single algorithm considering that the interest here is in the online learning aspect of this category which fits with the discussed requirements. That said, each smart-object owner or organization can choose the algorithms to implement according to their needs and their objectives.

F. Algorithm & inference system

In order to present the inference system of the proposed framework, hereafter the definitions of its components:

A: The requester how wants to access the resource

B: The resource or the object

$Req(A, B)$ : Requester A sends a request access to Target B

$SmartContract(B, A, S)$ : The target B redirects the subject A to SmartContract S

$GrantAccess(A, B, T)$ : Complete the remaining steps to get access (get, send and check the validity of the token) and then Allows A to access B. Also create a transaction T.

$AddBC(T)$ : Add the transaction T to the blockchain

$feed(A, B, T)$ : send feedback information of the transaction T to A and B

$update(B)$ : Update A knowledge (SmartContract, level of credibility, trust, integrity, ...)

$Reject(A, B, R)$ : Reject the request henceforth named R, deny the access of A to B

In the beginning an access request to resource B is sent by A. The output of this step is a SmartContract S. In case of failure, the result is a  $Reject(A, B, R)$  plus the feedback; otherwise the remaining steps to get access are executed: allowing A to access B and then generating a transaction T which can be added to the blockchain, without forgetting to send the feedback. After the feedback A's and B's knowledge are updated.

init	A	B	$Req(A, B)$
	$SmartContract(B, A, S)$		
Success	A	B	S
	$GrantAccess(A, B, T)$	$AddBC(T)$	$feed(A, B, T)$
Follow	A	B	$feed(A, B, T)$
	$update(B)$	$update(A)$	
failure	A	B	S
	$Reject(A, B, R)$	$feed(A, B, R)$	

Fig. 8. Inference system of the proposed framework.

V. CONCLUSIONS AND FUTURE WORKS

Today, IoT is surrounding us and its aptitudes of sensing, actuation, communication, and control become ever more sophisticated and ubiquitous; however these advantageous features are also examples of security and privacy (trust among users and things [31]) threats that are already nowadays slowing down the growth and expansion of the Internet of Things when not fulfilled properly.

This work has focused on the access control in the Internet of things environments. It proposed a framework that aims to solve two problems: 1) Problems that come with the centralized architecture, without being forced to transmit the management of the access control from a central entity to the nodes of the network. Indeed, the constrained devices generally used in the IoT do not have the capacity of calculation nor of storage to deal with a full distributed access control. 2) Problems of handling the access control policies

especially with the colossal number of smart devices supposed to be managed in IoT situations. This commonly leads to adopt static policies where the manager or the security officer writes all the security or access control rules in a static manner. The new framework proposed in this paper responds to these problems, it gives people total control of their IoT devices without being forced to trust in an outside entity and also present an automatically-improved and dynamic security policy.

The proposition presented in this article is based, on one hand, on the concept of the blockchain to ensure a totally distributed infrastructure to ensure an access control without trusting external central entities. This distributed aspect is strongly recommended in the Internet of things environments as well as privacy and unlinkability. On the other, the framework relies on an “online learning” mechanism of machine learning algorithms (Reinforcement Learning) in order to provide a dynamic, optimized and self-adjusted security policy.

In this paper, we presented an introduction of the internet of things paradigm, and how the access control and security policy management are among the IoT's priorities, both at the technological and research level. We then gave more details of some fundamental notions in this work namely the concepts of blockchain, machine learning and distributed architecture. Then section III was concentrated on previous/related works done in this domain. After that a presentation of the proposed framework was given by, first, mentioning the problem statement and research questions, then by explaining the contribution of this paper to improve/upgrade security policy using machine learning algorithms. Furthermore an explication of the architecture and algorithm that operate the framework was exposed. Finally, we conclude the paper with an explicit inference system for a better understanding of this work.

However, this contribution still has some limitations on which we intend to work in our future paper. Indeed, blockchain technology presents some intrinsic drawbacks especially when talking about privacy, required time for block validation, and so on. We also pretend to complete this framework with integrating the notion of collective intelligence which will respond to privacy concerns. As a final point, this model needs also a thorough case study as well as an implementation as a concrete proof of concept.

#### REFERENCES

- [1] J. Lopez, R. Rios, F. Bao and G. Wang, “Evolving Privacy: From Sensors to the Internet of Things”, 2017, p. 1.
- [2] S. Elbouanani, M. A. Elkiram, O. Achbarou, “Introduction To The Internet Of Things Security”, 2015. [Accessed: 06/2017].
- [3] A. Alkhalila, R. A. Ramadan, “IoT Data Provenance Implementation Challenges”, *3rd International Workshop on Tasks on High Performance Computing (THPC)*, 2017, p. 3.
- [4] A. Ouaddah, A. Abou Elkalam and A. Ait Ouahman, “FairAccess: a new Blockchain-based access control framework for the Internet of Things”, *Security and Communication Networks*, 2017, pp. 1-22.
- [5] D. J. Dougherty, K. Fisler and S. Krishnamurthi, “Specifying and Reasoning About Dynamic Access-Control Policies”, *IJCAR 2006*, pp. 632–646.
- [6] T. O. Ayodele, “Introduction to Machine Learning”
- [7] M. Abu Alsheikh, S. Lin and D. Niyato, “Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications”, *IEEE Communication Surveys & Tutorials*, vol. 16, no. 4, 2014,
- [8] Y. S. Abu-Mostafa, M. Magdon-Ismael, and H.-T. Lin, “Learning From Data”, *AMLBook*, 2012.
- [9] L. P. Kaelbling, M. L. Littman and A. W. Moore, “Reinforcement Learning: A Survey”, *Journal of Artificial Intelligence Research* 4, 1996, pp. 237-285.
- [10] O. Chapelle, B. Scholkopf, and A. Zien, *Semi-Supervised Learning*, vol. 2. Cambridge, MA, USA: MIT Press, 2006
- [11] S. Nakamoto, “Bitcoin : A Peer-to-Peer Electronic Cash System,” pp. 1–9, 2008
- [12] C. Fromknecht, D. Velicanu, and S. Yakoubov, “A Decentralized Public Key Infrastructure with Identity Retention” *IACR Cryptol. ePrint*, 2014.
- [13] A. Schaub, R. Bazin, O. Hasan, and L. Brunie, “A trustless privacy-preserving reputation system,” *IFIP Int. Inf.*, 2016.
- [14] S. Wilkinson, J. Lowry, and T. Boshevski, “Metadisk a blockchain-based decentralized file storage application,” 2014.
- [15] A. Ekblaw, A. Azaria, J. Haramka, and A. Lippman, “A Case Study for Blockchain in Healthcare: ‘MedRec’ prototype for electronic health records and medical research data,” 2016.
- [16] G. Zyskind and A. S. Pentland, “Decentralizing Privacy: Using Blockchain to Protect Personal Data,” 2015.
- [17] K. Christidis and M. Devetsikiotis, “Blockchains and Smart Contracts for the Internet of Things”, 2016
- [18] C. Scardovi, “Fin Tech Innovation and the Disruption of the Global Financial System,”.
- [19] R. S. Sandhu, “Role-based Access Control,” *Adv. Comput.*, vol. 46, pp. 237–286, 1998.
- [20] V. C. Hu, D. Ferraiolo, R. Kuhn, “Guide to Attribute Based Access Control (ABAC) Definition and Considerations”, *NIST Special Publication*, 2014.
- [21] N. Ye, Y. Zhu, R. Wang, R. Malekian, and L. Qiao-min, “An Efficient Authentication and Access Control Scheme for Perception Layer of Internet of Things,” *Appl. Math. Inf. Sci. An Int. J.*, vol. 1624, no. 4, pp. 1617–1624, 2014.
- [22] J. Park and R. Sandhu, “Towards usage control models: beyond traditional access control,” in *Proceedings of the seventh ACM symposium on Access control models and technologies - SACMAT '02*, 2002, p. 57.
- [23] A. Lazouski, F. Martinelli, and P. Mori, “Usage control in computer security: A survey,” *Comput. Sci. Rev.*, vol. 4, no. 2, pp. 81–99, 2010.
- [24] X. Zhang, M. Nakae, M. J. Covington, and R. Sandhu, “Toward a Usage-Based Security Framework for Collaborative Computing Systems,” *ACM Trans. Inf. Syst. Secur.*, vol. 11, no. 1, pp. 1–36, Feb. 2008.
- [25] G. Zhang, W. Gong, The research of access control based on UCON in the internet of things, *J. Softw.* (2011)
- [26] A. A. E. Kalam et al., “Organization based access control,” in *Proceedings POLICY 2003. IEEE 4th International Workshop on Policies for Distributed Systems and Networks*, 2003, pp. 120–131.
- [27] A. Abou El Kalam, Y. Deswarte, A. Baïna, and M. Kañiche, “PolyOrBAC: A security framework for Critical Infrastructures,” *Int. J. Crit. Infrastruct. Prot.*, vol. 2, no. 4, pp. 154–169, 2009.
- [28] A. Ouaddah, I. Bouij-Pasquier, A. Abou Elkalam, and A. Ait Ouahman, “Security analysis and proposal of new access control model in the Internet of Thing,” in *2015 International Conference on Electrical and Information Technologies (ICEIT)*, 2015, pp. 30–35.
- [29] I. Bouij-Pasquier, A. A. El Kalam, A. A. Ouahman, and M. De Montfort, “A Security Framework for Internet of Things,” *Springer International Publishing*, 2015, pp. 19–31.
- [30] A. Ameziane El Hassani et al., “Integrity-OrBAC: a new model to preserve Critical Infrastructures integrity”, *Int. J. Inf. Secur*, Springer-Verlag Berlin Heidelberg 2014.
- [31] S. Sicari, A. Rizzardi, L.A. Grieco, A. Coen-Porisini, “Security, privacy and trust in Internet of Things: The road ahead”, *Computer Networks*, 2015, pp. 146–164.

# Eye Controlled Mobile Robot with Shared Control for Physically Impaired People

Muhammad Wasim

Al-khawarizmi Institute of Computer Science,  
UET Lahore, Pakistan

Javeria Khan

Al-khawarizmi Institute of Computer Science,  
UET Lahore, Pakistan

Dawer Saeed

Al-khawarizmi Institute of Computer Science,  
UET Lahore, Pakistan

Dr. Usman Ghani Khan

Al-khawarizmi Institute of Computer Science,  
UET Lahore, Pakistan

**Abstract**—Physically impaired and disabled people are an integral part of human society. Devices providing assistance to such individuals can help them contribute to the society in a more productive way. The situation is even worse for patients with locked-in syndrome who cannot move their body at all. These problems were the motivation to develop an eye controlled robot to facilitate such patients. Readily available commercial headset is used to record electroencephalogram (EEG) signals for classification and processing. Classification based control signals were then transmitted to robot for navigation. The robot mimics a brain controlled wheelchair with eye movements. The robot is based on shared control which is safe and robust. The analysis of robot navigation for patients showed promising results.

**Keywords**—Locked-in syndrome; EEG; shared control; eye controlled robot

## I. INTRODUCTION

Brain Computer Interface (BCI) provides direct communication between computer and human brain by detecting electromagnetic brain signals and converting them into appropriate command signals[1]. The recorded brain activity can be used for assistive devices, gaming, and robotics. Signal recorded by brain is categorized as invasive or noninvasive. In Invasive BCIs electrical signals are recorded directly on or inside the cortex which require surgery to implant these electrodes. Different methods are used for recording these signals. Electrocorticogram (ECoG) signals are recorded which involves electrical signals for its subsequent analysis[2]. Another method uses an array of multiunit electrode placed in the cortex to record the natural activity for small population of neurons[3]. Both methods have greater signal-to-noise ratio, but require very less training of the user and are suitable for restoring the damaged neurons of the patients[4]. In non-invasive BCIs electrical activity of brain is recorded by placing the electrodes on the scalp so it does not require surgery[5]. Non-invasive BCIs use various types of signals from brain as inputs, such as electroencephalograms (EEG), blood-oxygen-level-dependent (BOLD), deoxyhemoglobin concentrations, and magnetoencephalograms (MEG) which detects weak magnetic fields due to the current flow in the cortex[6]. EEG signals are

most popular among all of them due to their low cost and convenient to use[7].

Recently robots are not only used in industry but also in daily life activities to provide assistance especially for the people suffering from disabilities like locked-in syndrome which is a neurodegenerative disorder [8]. In this syndrome person is fully aware of the environment but is not able to convey his/her commands so brain computer interface helps in providing a promising means to regain their mobility. Locked-in syndrome which leads to amyotrophic lateral sclerosis (ALS)[9] effects the motor neurons of the patients so the conventional EEG signal from motor imagery is not enough to control the robot which requires a lot of training and can be a tiring and stressful process. EEG signals from eye movements can provide aid for people suffering from such disorders. In previous work, electrooculogram (EOG) was used to detect eye movements which captures the corneal-retinal potential (CRP)[10]. The EOG varies from 0.05 to 3.5 mV and is proportional to eye movements. Electrodes are placed near the frontal part of brain near the eyes to measure electrooculographic potential (EOG), which is a function of eye position with respect to the head[11]. The strength of signal captured by this method is good but it causes discomfort for the user.

BCI robots based on EEG signals are most widely used in robotics and in assistive devices. The objective of BCI-robots is to convert the human intentions into appropriate commands so robots can perform various complex tasks which allows user to control robots naturally without any external signal. BCI robots are categorized into two main classes as mobile robots and brain-controlled manipulators[12]. Recently due to the development in brain computer interface mobile robots have gain the attention of most researchers due to their ability to transport disabled people [13] [14] but at the same time these mobile robots require higher safety as they are designed to provide transportation for disabled people. To implement these robots in real life much higher accuracy is required.

In this proposed method EEG based non-invasive BCI is used to detect the eye movement. Instead of using conventional high cost multiple channel EEG systems, low cost non-invasive EEG system is presented. On the basis of

the signals received by the physical movements of eye, the decisions are taken and hence the robot navigates accordingly.

The paper is organized in a way that related work discusses the previous methodologies used for brain signal extraction and shared control implementation. System model explains the overall proposed system and robot setup. Methodology section discusses the signal acquisition and its processing for the detection of eye movements and also the shared control implementation.

Robot movement under different scenarios explains the robot navigation under various situations. Experiments and results section describes the experiments designed to evaluate its performance and results are also shown. The conclusion section summarizes the results and discussions made in the paper. Finally the future works explains the future aspects of this work.

## II. RELATED WORK

BCI is most widely used in real time applications like gaming [15], [16], virtual reality [17] and in robotics [18], [19], [20]. Apart from these applications BCI can provide assistance for the individuals suffering from severe mobility disorders such as brain controlled wheelchairs for people suffering from disabilities [21], [22], telepresence robots to provide aid for disabled persons[23], and exoskeletons [24],[25].

In the past image processing techniques were used to detect the eye movements [26]. Eye gaze tracking technique which consist of camera and computer, by computing the distance traveled by pupil, the movement of eye is tracked[27] videooculography systems (VOG)[28] or infrared oculography (IROG) [29] based on position of eye were also used. These previous techniques increase the complexity of the system but with the recent advancement in the brain computer interface researchers are now working on brain signals to capture the desire brain signal for BCI applications[30].

Apart from the signal extraction shared control anatomy is very important for all the assistive devices. It is very essential to develop such a design in which system is aware of the environment and capable of deciding that when to give control to human, machine or both[7]. Research work is being done to establish a shared control that can be implemented in real time environment. A key feature in these shared control design is the use of several assistance modes[23]. Most researchers [31], [32], [33] used autonomous approach in which the user only specify the destination and the navigation system plans a short and safe path.

Zhang[31] presented autonomous navigation in which user selects the destination by using P300 based BCI. Obstacle map is constructed in 2D with the help of webcam and laser range finder sensor and the way-points are generated using generalized Voronoi diagram. The research work [32] presented the autonomous approach in which user selects the destination point from the screen which displays real time scenarios constructed by a laser scanner which is transmitted

to autonomous navigation system. It helps user to navigate in an unknown environment with the help of sensors. Once the destination is selected, user can relax which reduce the mental workload. Rebsamen and Cuntai Guan [33] also presented autonomous approach where paths are already defined and user can select only from these paths. This method require continuous update in the predefined paths in case of any change in environment. It is only suitable for a room equipped with webcams and not suitable for real time environment. Autonomous approach has a serious drawback that the user has no control on the robot and it does not work well in the situation where the environment stochastic so under these circumstances it is better to control the robot by the user and use semi-autonomous approach instead of autonomous approach.

Leeb and Robert [23] used semi-autonomous approach in which user control high level commands (turn left, right) while low level commands (avoiding obstacle) are implemented by the robot. Only two classes are used and the default behavior of the robot is to move forward with a constant speed. In case of repellers or attractors (representing obstacle and target) the motion of device changes in order to avoid them. In this proposed shared control semi-autonomous approach is implemented. Instead of using two classes, four classes (forward, backward, left, right) are used to give maximum control to the user. The default behavior of the robot is to stop in order to avoid any dangerous situations.

In this paper EEG is used to capture brain signal corresponding to the eye movements. Four classes are used for the movement of robot. Forward command is executed when the user lift his eyebrows. Left and right movement is implemented by the pupils movements to the left and right receptively. Eye blink is used to stop the robot. In order to differentiate from normal blinking of eye user have to blink three times. In this proposed shared control method semi-autonomous approach is implemented to give the maximum control of robot to the user.

## III. SYSTEM MODEL

The proposed BCI system consists of three main components: Signal acquisition unit which records electrical signals corresponding to the eye movements; the control unit which decides the desired motion commands on the basis of EEG signal received from EPOC headset; and the shared control unit ensuring the safety of user by avoiding obstacles. The block diagram of the system is shown in the figure 1.

### A. Robot Setup

The test robot is made from readily available and open source equipment to demonstrate the proof of concept. Short range ultrasonic sensors are used to detect the obstacles in front of robot. Each sensor covers a 15° of range in front of it so to cover more range in front of the robot, three sensors have been used. While only one sensor has been used on rest of the three sides. To control the motor an H-Bridge is implemented in between the motors and microcontroller. The communication of the robot and control unit is done using Bluetooth.

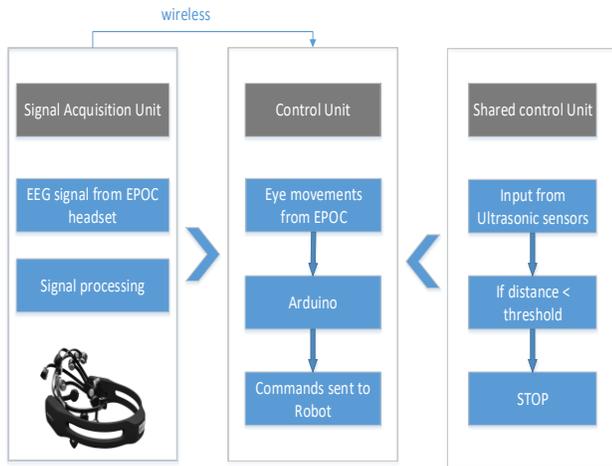


Fig. 1. Proposed Shared Control System Model

All the above mentioned setup is implemented using cheap and readily available components. An image of the developed robot is shown in figure 2. The robot has custom made wheels which are developed in order to make it very stable and aligned to move in a straight path. The wheels are also light weight and already have an O ring which has reduced its surface friction.

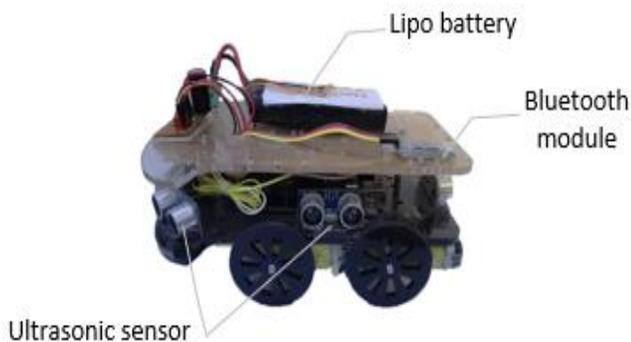


Fig. 2. Eye Controlled Mobile Robot

#### IV. METHODOLOGY

The proposed system provides a method to detect the muscle activity of eyes through EEG signals which filtered the acquired signal for identifying eye movements. Emotiv API processes the acquired signals and performs classification. Once classified the control signals are transmitted to the robot. A high level system perspective is shown in the figure 3.



Fig. 3. System overview

#### A. Signal Acquisition and Processing

A widely used and commercially available wireless EEG headset Emotiv EPOC<sup>1</sup> with 14 channels is used for BCI application. Software Development Kit (SDK) from Emotiv is also used for signal acquisition and classification. The application consists of three emotiv suites which are named as affectiv, cognitiv, and expressiv suite. These suites are developed to recognize the user's emotional state, thoughts and facial expressions respectively.

This experiment focuses on expressive suit which filters out the facial expression and hence eye blinks. The specific signals related to eye movement are classified from a wide range of facial expressions. This classification mainly uses the electroencephalogram (EEG) for eye blink and for the detection of muscle movements for looking right and left. Since all the users have same facial expressions so it does not need training. Sensitivity of the signal can be changed through the classifier. Certain facial expressions for some users are weak to detect so the sensitivity was tuned accordingly. Since every expression is a different expression so the sensitivity of each and every action was also fine-tuned accordingly.

#### B. Shared Control and Robot Movement

One of the most important and tedious task of controlling through BCI is its accuracy and high latency in processing the signals which can be tedious for the user some times. Also user's safety is very important. To ensure the user safety and to reduce the mental work load from user shared control is implemented. Shared control takes over the user in case of any obstacle is critically close. To implement the shared control, signal received from brain is classified into 4 classes (forward, left, right, and stop). The shared control implements the lower level control like obstacle avoiding, critical distance handling and collision prevention while the user controls the high level commands. The default behavior of the robot is to stop. To move the robot from its position either the obstacle has to be moving and the robot will counter for it accordingly or the robot must receive a signal from brain. As the robot receives signal, it keeps on moving in the respective direction until the user sends a signal again or an obstacle is detected. The flow chart of the algorithm implemented on the robot is also shown in the figure 4.

<sup>1</sup> <https://www.emotiv.com/epoc/>.

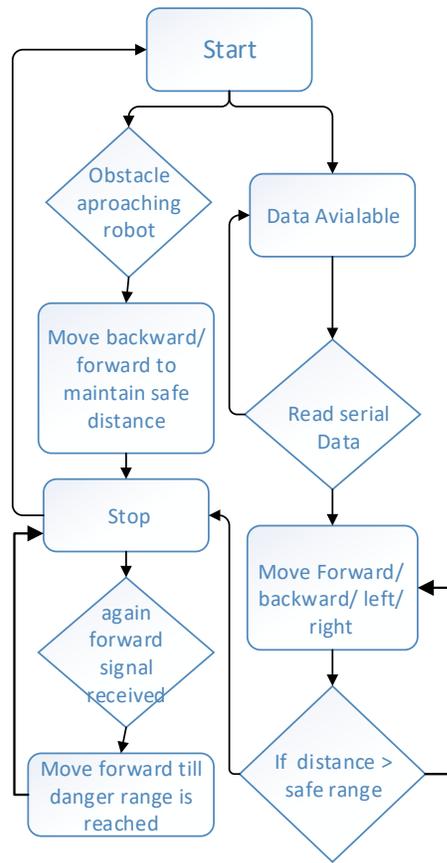


Fig. 4. Flow chart of the shared control algorithm.

V. ROBOT MOVEMENT UNDER DIFFERENT SCENARIOS

The shared control along with safety and other very promising features also comes with difficult control and working situations under different circumstances. The basic obstacle avoidance of the robot is shown in figure 5.

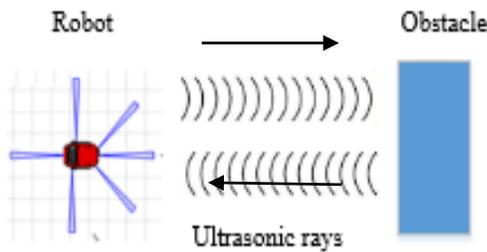


Fig. 5. Obstacle detection through ultrasonic sensor

The robot in presented work also behaves on different scenarios which are explained as follows:

A. External Signal

The robot moves forward as soon as the forward command is received from BCI. Whenever an object comes within the safe range the robot stops so that the safety of the user is maintained.

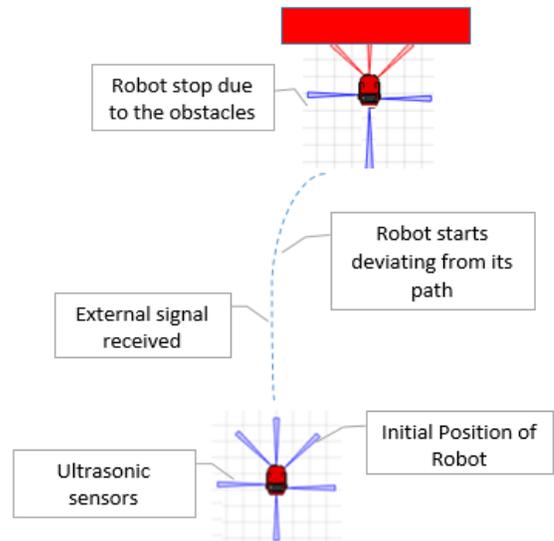


Fig. 6. Robot movement in case of external signal received

An important feature of the robot is to differentiate between the obstacle and target if the user wants to move towards target the robot will move further after reaching safe range till danger range is reached which is very small distance as compared to the safe range.

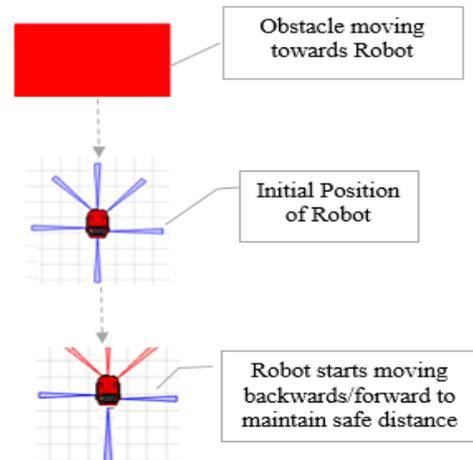


Fig. 7. When obstacle is moving towards robot

After danger range robot will not move any further irrespective of forward signal received. It is also demonstrated in figure 6. When left signal is received from BCI the robot turns left at some radius and continue to move in circle unless next signal is received and when an object comes in safe range the robot stops. As soon as the safe range is reached robot stops irrespective of the signal received from the user. Similar is the case when the right command or signal is received.

B. Moving Obstacle

In the case when no external signal from the user is received, robot does not move at all, if un-disturbed it behaves

like an obstacle avoiding robot. If the obstacle is coming towards BCI robot front sensor detects the obstacle and robot starts moving backwards and maintains a safe range from the obstacle similarly if obstacle is following the robot it will move forward and maintains a proper distance from the obstacle which is shown in figure 7.

VI. EXPERIMENTS AND RESULTS

In order to test the effectiveness of proposed system two experiments were designed. For this, two different paths were designed which differ on the basis of their path complexity. Each experiment had 3 trials. User is directed to go to the specified goal in the described path in the specified time period. Before the actual experiment randomized list of task was designed which included 20 trials of eye movements in order to test the correctness of BCI. The performance is evaluated on the basis of time delay between the eye movements and the execution of command. Workload as well as stress level on the user is also studied by the user feedback at the end of each experiment.

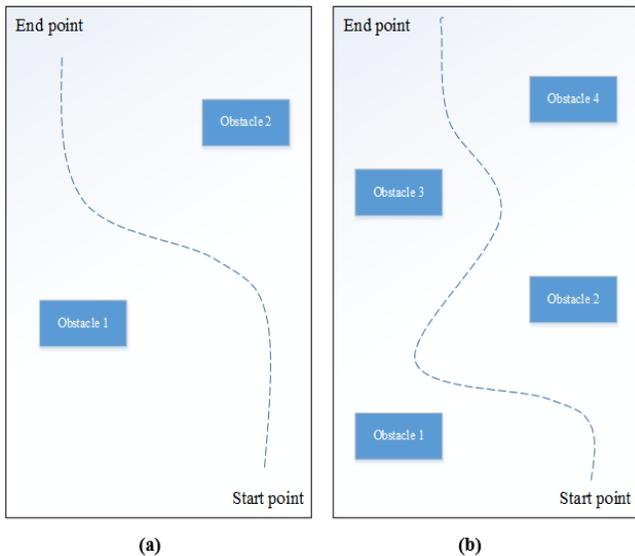


Fig. 8. (a) Path designed for experiment 1 (b) Path designed for experiment 2

The extent of control of user over robot is another important parameter for the evaluation of proposed system. Speed of the robot has no effect on the BCI but it affects the user in terms of mental workload. At high speed it was difficult for the user to control the robot effectively and it was very tiring and stressful for the user. So the speed of robot was kept low which decreased the workload as well as enabled user to completely control the robot.

In the first experiment only few obstacles were added. The user has to reach the goal by avoiding obstacles. The results showed that the user was fully able to control the robot without any mental workload. While in second experiment a complex path was designed which includes sharp turns and user has to issue more commands in order to reach the goal. Both of the path scenarios are also shown in the figure 8 and the performance evaluation which is on the basis of path 1 versus path 2 is also shown in figure 9.

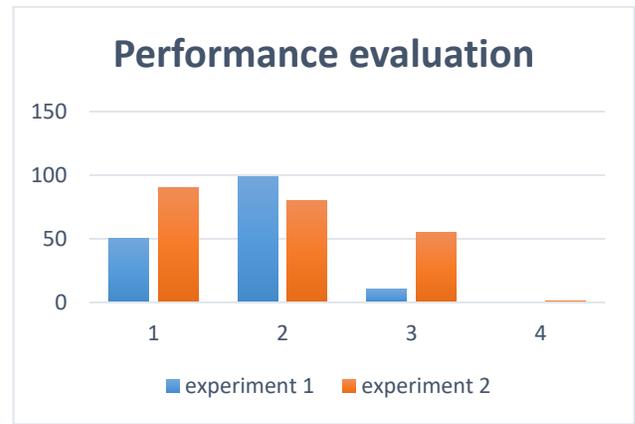


Fig. 9. Performance evaluation

TABLE. I. COMPARISON OF BOTH EXPERIMENTS

	Time to reach goal (1)	Control of robot (2)	No. of commands issued by user(3)	No. of collisions (4)
Experiment 1	50%	99%	30%	0
Experiment 2	80%	80%	60%	0

The comparative analysis of both experiments is shown in table 1. The results showed that user was able to reach the specified destination with 80% accuracy and the time to reach the goal was very high although there were no collisions with the obstacles. The feedback of the user showed that this was a bit tiring and requires more attention of the user. This experiment also showed that this was very safe there were no collision due to the obstacles which suggest that this system can be implemented on wheelchair without any safety issues.

VII. CONCLUSION

As prototype of brain controlled wheelchair, brain controlled robot was developed. The experiments were conducted in order to test whether the robot moves in accordance with the provided command and the system accuracy was tested by letting a person to control the robot in different path scenarios of variable complexity. The results demonstrated that the robot was successfully controlled by the person using EEG signals from eye movement without a single collision. The shared control is very crucial in this regard and helps the user to avoid any obstacle coming its way. The mental work load from the user is reduced by making the robot autonomous on the basis of last received command which was confirmed by the feedback from the user. However the workload also depends upon the navigation path complexity. The proposed strategy can also be used on industrial robots which are controlled by physically impaired people.

VIII. FUTURE WORK

Authors are working on controlling the robot with different methods which include gestures, muscle activity and other brain signals. Further improvements in this work can also be

done by implementing a manual control for a third person along with the brain and shared control.

REFERENCES

- [1] J. del R. Millán et al., "Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges," *Front. Neurosci.*, vol. 4, p. 161, 2010.
- [2] B. J. He, "Electrocorticogram (ECoG)," in *Encyclopedia of Computational Neuroscience*, D. Jaeger and R. Jung, Eds. New York, NY: Springer New York, 2015, pp. 1070–1074.
- [3] J. A. Wilson, E. A. Felton, P. C. Garell, G. Schalk, and J. C. Williams, "ECoG factors underlying multimodal control of a brain-computer interface," *IEEE Trans. neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 246–250, 2006.
- [4] A. Nijholt et al., "Brain-computer interfacing for intelligent systems," *IEEE Intell. Syst.*, vol. 23, no. 3, 2008.
- [5] J. del R. Milan and J. M. Carmena, "Invasive or noninvasive: Understanding brain-machine interface technology [conversations in bme]," *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 1, pp. 16–22, 2010.
- [6] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based brain-computer interface for cursor control," *Electroencephalogr. Clin. Neurophysiol.*, vol. 78, no. 3, pp. 252–259, 1991.
- [7] L. Bi, X.-A. Fan, and Y. Liu, "EEG-based brain-controlled mobile robots: a survey," *IEEE Trans. human-machine Syst.*, vol. 43, no. 2, pp. 161–176, 2013.
- [8] M. S. Cardwell, "Locked-in syndrome.," *Tex. Med.*, vol. 109, no. 2, pp. e1–e1, 2013.
- [9] R. Golby, B. Poirier, M. Fabros, J. J. Cragg, M. Yousefi, and N. Cashman, "Five-Year Incidence of Amyotrophic Lateral Sclerosis in British Columbia (2010-2015)," *Can. J. Neurol. Sci. Can. des Sci. Neurol.*, pp. 1–5, 2016.
- [10] J. A. Müller, D. Wendt, B. Kollmeier, and T. Brand, "Comparing eye tracking with electrooculography for measuring individual sentence comprehension duration," *PLoS One*, vol. 11, no. 10, p. e0164627, 2016.
- [11] R. Barea, L. Boquete, M. Mazo, and E. López, "Guidance of a wheelchair using electrooculography," in *Proceeding of the 3rd IMACS International Multiconference on Circuits, Systems, Communications and Computers (CSCC'99)*, 1999.
- [12] S. M. Grigorescu, T. Lüth, C. Fragkopoulou, M. Cyriacks, and A. Gräser, "A BCI-controlled robotic assistant for quadriplegic people in domestic and professional life," *Robotica*, vol. 30, no. 3, pp. 419–431, 2012.
- [13] A. Chella et al., "A BCI teleoperated museum robotic guide," in *Complex, Intelligent and Software Intensive Systems, 2009. CISIS'09. International Conference on, 2009*, pp. 783–788.
- [14] L. Tonin, R. Leeb, M. Tavella, S. Perdakis, and J. del R. Millán, "The role of shared-control in BCI-based telepresence," in *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, 2010*, pp. 1462–1466.
- [15] P. Martinez, H. Bakardjian, and A. Cichocki, "Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm," *Comput. Intell. Neurosci.*, vol. 2007, p. 13, 2007.
- [16] E. C. Lalor et al., "Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment," *EURASIP J. Appl. Signal Processing*, vol. 2005, pp. 3156–3164, 2005.
- [17] L. Mingyu, W. Jue, Y. Nan, and Y. Qin, "Development of EEG biofeedback system based on virtual reality environment," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the, 2006*, pp. 5362–5364.
- [18] F. Galán et al., "A brain-actuated wheelchair: asynchronous and non-invasive brain-computer interfaces for continuous control of robots," *Clin. Neurophysiol.*, vol. 119, no. 9, pp. 2159–2169, 2008.
- [19] J. del R. Millán, F. Renkens, J. Mourino, and W. Gerstner, "Non-invasive brain-actuated control of a mobile robot," in *Proceedings of the 18th international joint conference on Artificial intelligence, 2003*, no. EPFL-CONF-82919.
- [20] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic," *Comput. Intell. Neurosci.*, vol. 2007, 2007.
- [21] I. A. Mirza et al., "Mind-controlled wheelchair using an EEG headset and arduino microcontroller," in *Technologies for Sustainable Development (ICTSD), 2015 International Conference on, 2015*, pp. 1–5.
- [22] R. Zhang, Y. Li, Y. Yan, H. Zhang, and S. Wu, "An intelligent wheelchair based on automated navigation and BCI techniques," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, 2014*, pp. 1302–1305.
- [23] R. Leeb, L. Tonin, M. Rohm, L. Desideri, T. Carlson, and J. del R. Millán, "Towards independence: a BCI telepresence robot for people with severe motor disabilities," *Proc. IEEE*, vol. 103, no. 5, pp. 969–982, 2015.
- [24] A. Frisoli et al., "A new gaze-BCI-driven control of an upper limb exoskeleton for rehabilitation in real-world tasks," *IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev.)*, vol. 42, no. 6, pp. 1169–1179, 2012.
- [25] J. Webb, Z. G. Xiao, K. P. Aschenbrenner, G. Herrstadt, and C. Menon, "Towards a portable assistive arm exoskeleton for stroke patient rehabilitation controlled through a brain computer interface," in *Biomedical Robotics and Biomechanics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on, 2012*, pp. 1299–1304.
- [26] S. W. Wijesoma, D. F. H. Wolfe, and R. J. Richards, "Eye-to-hand coordination for vision-guided robot control applications," *Int. J. Rob. Res.*, vol. 12, no. 1, pp. 65–78, 1993.
- [27] M. Lin, B. Li, and Q.-H. Liu, "Identification of eye movements from non-frontal face images for eye-controlled systems," *Int. J. Autom. Comput.*, vol. 11, no. 1, pp. 543–554, 2014.
- [28] J. P. Hansen, A. W. Andersen, and P. Roed, "Eye-gaze control of multimedia systems," *Adv. Hum. Factors/Ergonomics*, vol. 20, pp. 37–42, 1995.
- [29] C. H. Morimoto and M. R. M. Mimica, "Eye gaze tracking techniques for interactive applications," *Comput. Vis. image Underst.*, vol. 98, no. 1, pp. 4–24, 2005.
- [30] A. Vallabhaneni, T. Wang, and B. He, "Brain-computer interface," in *Neural engineering*, Springer, 2005, pp. 85–121.
- [31] R. Zhang et al., "Control of a Wheelchair in an Indoor Environment Based on a Brain-Computer Interface and Automated Navigation," *IEEE Trans. neural Syst. Rehabil. Eng.*, vol. 24, no. 2, pp. 128–139, 2016.
- [32] I. Iturrate, J. M. Antelis, A. Kubler, and J. Minguez, "A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation," *IEEE Trans. Robot.*, vol. 25, no. 3, pp. 614–627, 2009.
- [33] B. Rebsamen et al., "A brain controlled wheelchair to navigate in familiar environments," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 4, pp. 590–598, 2010.

# Fast-ICA for Mechanical Fault Detection and Identification in Electromechanical Systems for Wind Turbine Applications

Mohamed Farhat

Department of Electrical  
Engineering, University of Tunis  
El Manar, LA.R.A. Automatique,  
National Engineering School of  
Tunis, BP 37, 1002 Tunis, Tunisia

Yasser Gritli

Department of Electrical  
Engineering, University of Tunis El  
Manar, LA.R.A. Automatique,  
National Engineering School of  
Tunis, BP 37, 1002 Tunis, Tunisia

Mohamed Benrejeb

Department of Electrical  
Engineering, University of Tunis El  
Manar, LA.R.A. Automatique,  
National Engineering School of  
Tunis, BP 37, 1002 Tunis, Tunisia

**Abstract**—Recently, the approaches based on source separation are increasingly adopted for the fault diagnosis in several industrial applications. In particular, Independent Component Analysis (ICA) method is attractive, thanks to its simplicity of implementation. In the context of electrical rotating machinery with a variable speed, namely the wind turbine type, the interaction between the electrical and mechanical parts along with the fault is complex. Therefore, the essential system variables are affected and it thereby requires to be analyzed in order to detect the presence of certain faults. In this paper, the target system is the classical association of a doubly-fed induction motor to a two stage gearbox for wind energy application system. The investigated mechanical fault is a uniform wear of two gear wheels for the same stage. The idea behind the proposed technique is to consider the fault detection and identification as a source separation problem. Based on the analysis into independent components, Fast-ICA algorithm is adopted to separate and identify the sources of the gear faults. Afterwards, a spectral analysis is applied on the signals resulting from the separation in order to identify the fault components related to the damaged wheels. The efficiency of the proposed technique for the separation and identification of the fault components is evaluated by numerical simulations.

**Keywords**—Source separation; fault diagnosis; independent component analysis; fast-ICA; spectral analysis

## I. INTRODUCTION

Wind power increasingly gain ground, thanks to its characteristics as an inexhaustible and clean source of energy, which has made it a privileged field of scientific research and technological development in the world. A recent report shows the large-scale expansion of the installation of wind farms in the world [1]. Yet, an electric machine, whether running as a motor or as a generator, is rather sized in torque.

In small powers, the speed is relatively high, however in the case of large powers, (several hundred KW to a few MW), the low speeds lead to very high torques and prohibitive generator masses. For this reason, a gearbox is typically interposed between the turbine and the generator. Consequently, the fast shaft of the gearbox is coupled to the shaft of the electric generator [2], [3], [4]. A recent study of faults in the wind energy conversion systems revealed that

about 10% of the identified defects are related to the gearbox [5], [6]. Although this proportion is apparently low, this type of fault often leads to prohibitive production stops. That's from where comes the need to continuously monitor the proper functioning of this essential component in the energy conversion chain. That is why, several diagnostic techniques for the fault detection in these speed multipliers have been developed. These techniques include: Analysis of acoustic emissions [7], [8], oil analysis [9], [11] and specifically vibratory analysis. In particular, the investigation of vibratory signals has been proposed in different works, using different approaches: statistical analysis [10]-[11], temporal and/or frequency domain [7], [12], [13].

In reality, the vibratory signals collected during operation contain relevant informations which reflect several sources of faults relating to the speed multiplier itself and to those associated with the machine coupled with it. This is clearly justified in the references [14], [15], where the characterization of bar breaking faults, as well as the unbalance was based on the time-frequency analysis of the vibratory signals.

However, the measured observations are often mixtures of the vibrations of the defects mentioned before. This makes the diagnosis of defects a very difficult task. To solve this problem, several techniques have been used to identify the sources of defects from the spectral mixtures resulting from vibratory signals [16], [17], [18].

In the literature, Independent Component Analysis (ICA) has been widely applied for the separation of sources in different domains, including medical imagery, telecommunications, and more recently for the diagnosis of faults in electromechanical systems [19], [20], [21], [22].

More recently, new ICA-based techniques have been proposed for fault diagnosis in the electromechanical systems. In fact, the most used algorithms of the (ICA) can be classified as follows

- The InfoMax algorithm [23] solves the ICA problem by maximizing the differential entropy of the output of an invertible non-linear transformation of the whitened observations;

- JADE [24]-[25] consists in jointly diagonalizing the set of the eigen-matrices constructed from the eigenvectors associated to the  $P$  greatest eigenvalues of the covariance matrix of the whitened observations;
- Fast-ICA [26] tries, after the whitening step, to maximize a contrast function based on negentropy.

In the present work, the fast temporal algorithm, known as Fast-ICA, has been adopted for the identification of gear faults because of its appealing characteristics: high convergence speed and low computational cost. Moreover, this technique is interesting since it is relatively insensitive to the increase in the number of sources. This paper is organized as follows: Fast-ICA is formulated for fault diagnosis in the second section. Then, the gear vibration data is described in the third section. Afterwards, the fourth section is dedicated for the spectral analysis. Finally, the paper ends with a conclusion.

## II. FORMULATION OF THE FAST-ICA FOR FAULT DIAGNOSIS

The Fast-ICA algorithm is an advanced version of the ICA, characterized mainly by a very fast convergence, whose separation into independent components takes place in a whitened space [27],[28]. In fact, instantaneous linear mixtures (signals from sensors) are preprocessed. This consists in their projection into a whitened space. Then, they are separated by the Fast-ICA algorithm itself. The details of these two preprocessing steps and the Fast-ICA processing are described in the following. Furthermore, several nonlinearity functions are presented because of their impact on the performance of Fast-ICA algorithm.

### A. Preprocessing step

Let  $n$  sources of faults  $s_j$  denoted by  $[s_1, \dots, s_n]^T$ , and mixed before being retrieved by the sensors. Thus,  $m$  mixtures  $x_i$  of length  $N$ , are represented as rows of a  $m \times N$  matrix denoted  $X = [x_1, \dots, x_m]^T$ .

$$X = \begin{pmatrix} X_{1,1} & \dots & \dots & \dots & X_{1,N} \\ \vdots & \dots & X_{i,j} & \dots & \vdots \\ X_{m,1} & \dots & \dots & \dots & X_{m,N} \end{pmatrix} \quad (1)$$

Moreover, it can be represented by a linear model as

$$X = AS + bh \quad (2)$$

where  $A$  is an  $m \times n$  mixing matrix,  $h$  is the additive noise with the corresponding Gaussian weight vector given by  $b = [b_1, \dots, b_n]^T$ .

In order to apply Principal Component Analysis (PCA) to the mixtures, they should be considered differently. Indeed, the mixtures  $X$  should be seen as a set of  $N$   $m$ -dimensional points. Now, each column of  $X$  is interpreted as the coordinates of a point in the space  $\mathbb{R}^m$ .

First of all, PCA computes the mean of the  $N$  points, denoted  $\mu = (\mu_1, \dots, \mu_m)^T$ , as follows

$$\mu_i = \frac{1}{N} \sum_{j=1}^N X_{i,j}, \text{ for all } i=1, \dots, m \quad (3)$$

Then, PCA centers each point relatively to  $\mu$  as follows:

$$(X'_{1,j}, \dots, X'_{m,j})^T = (X_{1,j}, \dots, X_{m,j})^T - (\mu_1, \dots, \mu_m)^T, \quad \text{for all } j=1, \dots, N \quad (4)$$

Therefore, the resulting matrix denoted  $X'$  has as rows the centered mixtures.

$$X' = \begin{pmatrix} X'_{1,1} & \dots & X'_{1,j} & \dots & X'_{1,N} \\ \vdots & \dots & \vdots & \dots & \vdots \\ X'_{m,1} & \dots & X'_{m,j} & \dots & X'_{m,N} \end{pmatrix} \quad (5)$$

Afterwards, the covariance matrix  $\mathcal{C}$  of  $X'$  is computed as follows

$$\mathcal{C} = \frac{1}{N-1} X' X'^T \quad (6)$$

Then, the  $\mathcal{C}$  matrix is diagonalized as follows

$$\mathcal{C} = EDE^T \quad (7)$$

Therefore, two matrices are obtained

- a diagonal matrix denoted  $D$  composed of decreasingly sorted eigenvalues of the covariance matrix of  $X'$ .
- a matrix denoted  $E$  whose columns are the eigenvectors of the covariance matrix of  $X'$ . These eigenvectors are pairwise orthogonal.

Once PCA achieved, the whitening matrix denoted  $U$  is calculated by the following expression

$$U = D^{-\frac{1}{2}} E^T \quad (8)$$

Finally, this step results in the matrix composed of whitened mixtures, denoted  $V$ , is obtained by

$$V = UX \quad (9)$$

### B. Processing Step: Implementation of the "Fast-ICA" fixed point algorithm

The ICA method defines a separation model in order to estimate the sources  $\hat{S}$  given the whitened mixtures  $V$

$$\hat{S} = W^T V \quad (10)$$

Therefore, the goal of the ICA subsequently is to estimate  $W^T$ , called the whitened separation matrix. In particular, the Fast-ICA estimates the independent components by maximizing the non-gaussianity, defined as the opposite of the deviation of this signal distribution relatively to a gaussian signal distribution of the same power. It is thus possible to separate the sources of a linear mixture by maximizing the non-gaussianity of the obtained output signal by a linear combination of the observations.

There are multiple approaches to measure the non-gaussianity. After several trials with different approaches, mainly: normalized kurtosis, negentropy, [26], [29], the authors in the literature opted for negentropy. Next, the steps of the Fast-ICA algorithm are described, represented by the flowchart of Figure 1.

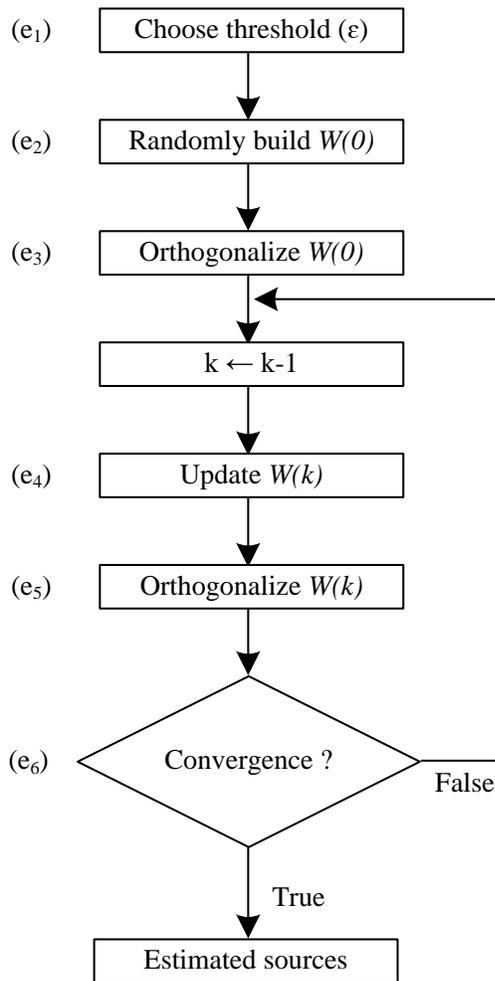


Fig. 1. Flowchart of FastICA.

Indeed, the application of the Fast-ICA algorithm starts with

step (e<sub>1</sub>) which assigns a positive and infinitely small value to a parameter, called convergence threshold and denoted  $\epsilon$ .

The next initialization step denoted (e<sub>2</sub>) consists in constructing  $W(k)$ , as well as its zero-order orthogonalization in step (e<sub>3</sub>). Thereafter, the algorithm iteratively performs the following two steps

- step (e<sub>4</sub>) of updating the matrix  $W$  to the order  $k$  is performed by the following equation of the fixed point of the negentropy

$$W(k) = E\{Vg(W^T(k-1)V)\} - E\{Vg'(W^T(k-1)V)\} \quad (11)$$

where the function  $g$  is representing the non-linearity of the Fast-ICA algorithm, which will be detailed later.

- the orthogonalization step (e<sub>5</sub>) based on the symmetric method, which does not favor any vector  $w$ , consists of starting directly from any matrix  $W$ , orthogonalizing it by the Gram-Schmidt approach, as follows

$$W(k) \leftarrow (W(k)W^T(k))^T W(k) \quad (12)$$

- finally, at the end of each iteration, the algorithm checks in step (e<sub>6</sub>) whether it has reached a maximum of the negentropy, which is based on the thresholding process given by

$$1 - \|W^T(k)W(k-1)\| < \epsilon \quad (13)$$

### C. Choice of the nonlinearity:

The function  $g$  of equation (11) is the non-linearity of Fast-ICA, which can be, as shown in the literature

- tangent-hyperbolic noted  $g_1$  which is effective for any type of situation

$$g_1 = \tanh(y) \quad (14)$$

- kurtosis noted  $g_2$  which is used only in the case of subgaussian variables

$$g_2 = y^3 \quad (15)$$

- exponential or Gauss noted  $g_3$  which is more suitable in the case of supergaussian variables

$$g_3 = ye^{-\frac{y^2}{2}} \quad (16)$$

- skewness noted  $g_4$  is the third level moment which measures the asymmetry of the data

$$g_4 = \frac{y^2}{2} \quad (17)$$

The choice of the function  $g$  has a direct impact on the updating of  $W$  as indicated in equation (12), and consequently on the overall performance of the algorithm.

## III. PRESENTATION OF GEAR VIBRATION DATA

### A. System Description

In order to evaluate the efficiency of the method described above, the asynchronous double-feed machine-speed multiplier combination of Figure 2 has been considered.

More precisely, the defects relating to the gear- A two-stage speed multiplier are interesting. Indeed, the gear in question is composed of four toothed wheels (R1, R2, R3 and R4). The system under consideration is assumed to operate at nominal speed of 1012 rpm on the side of the generator (Wheel R4) and 46 rpm on the turbine side (Wheel R1).

In fact, the vibrations resulting from the gearbox operation are due to the forces of mutual contact between the teeth of the wheels in contact. For two wheels, of the same stage, making contact, a meshing frequency is given by

$$f_{mesh} = f_{r,i}Z_i = f_{r,i+1}Z_{i+1} \quad (18)$$

where the  $f_{r,i}$  and  $f_{r,i+1}$  are the rotational frequencies of the wheels for the same considered stage. The numbers of teeth relative to each wheel are denoted  $Z_i$  and  $Z_{i+1}$ .

Under healthy gear, the vibration spectrum typically shows the harmonic chain in (19) with small amplitudes.

$$f_{sideband} = kf_{mesh} \pm f_{r,i} \quad (19)$$

On the other hand, in presence of a uniform wear fault on all the teeth of the same wheel, the amplitude of the harmonics in (19) shows a noticeable increase, making it possible to

identify the wheel affected by the fault. In the considered system, the wheels R1 and R2 have the same meshing frequency  $f_{mesh1,2}$  and have respectively different lateral frequencies ( $f_{l1}, f_{r1}$ ) and ( $f_{l2}, f_{r2}$ ).

Likewise, the wheels R3 and R4 have the same meshing frequency  $f_{eng3,4}$  and have respectively different lateral frequencies ( $f_{g3}, f_{d3}$ ) and ( $f_{g4}, f_{d4}$ ), as detailed in Table I.

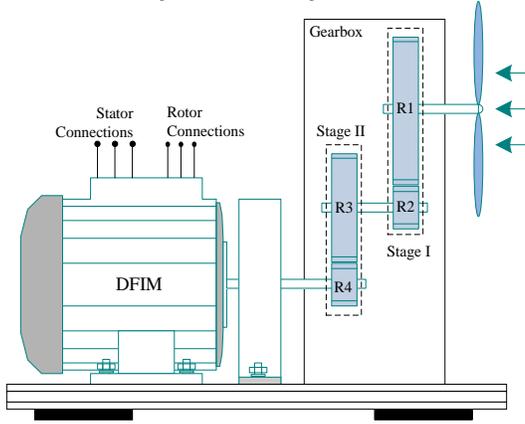


Fig. 2. Simplified representation of the generator association with doubled multiplier of speed composed of two stages.

TABLE I. FREQUENCIES OF FAULT IDENTIFICATION

	Rotating Speed (tr/min)	Number of teeth $Z_i$	$f_{lk}, (k=1..4)$	$f_{rk}, (k=1..4)$
<b>R1</b>	46	69	52.12	53.66
<b>R2</b>	186.7	17	49.78	56
<b>R3</b>	186.7	92	283.62	289.84
<b>R4</b>	1012	17	269.87	303.59

### B. Description of mixtures

The mixtures are linear combinations of the sources as indicated in equation (2), where matrix  $A$  and vector  $b$  must be specified

- the adopted mixture matrix  $A$  is chosen as

$$A = \begin{pmatrix} 0.9 & 0.53 & 0.7 & 0.6 \\ 0.45 & 0.9 & 0.68 & 0.55 \\ 0.55 & 0.71 & 0.9 & 0.6 \\ 0.6 & 0.52 & 0.7 & 0.9 \end{pmatrix} \quad (20)$$

- the vector  $b$ , of dimension  $4 \times 1$ , for the weighting of noise in the mixtures

$$b = \begin{pmatrix} 0.3 \\ 0.5 \\ 0.7 \\ 0.1 \end{pmatrix} \quad (21)$$

On the other hand, a mixture can be represented either in the temporal space or in the spectral space. Nevertheless, the choice of the appropriate representation is required. It allows to know whether a mixture is in healthy mode or in faulty mode.

The temporal representation of the mixtures makes it possible to distinguish the healthy mode from the faulty mode. Indeed, the amplitudes in the faulty mode shown in Figure 4 are generally greater than the amplitudes in the healthy mode illustrated in Figure 3. However, the problem is that the temporal representation does not make possible to display exactly which wheels are affected by the fault. For this purpose, it is preferred to use the spectral representation instead of the temporal representation.

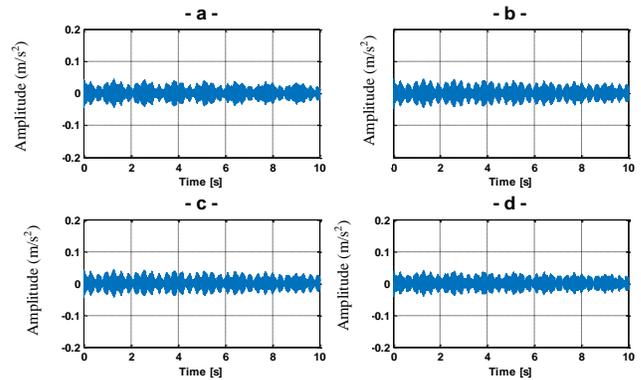


Fig. 3. Mixtures used for the separation under healthy condition.

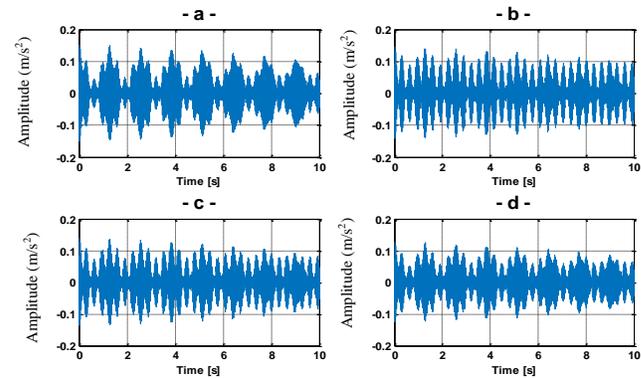


Fig. 4. Mixtures used for the separation under a uniform wear of R1 and R2.

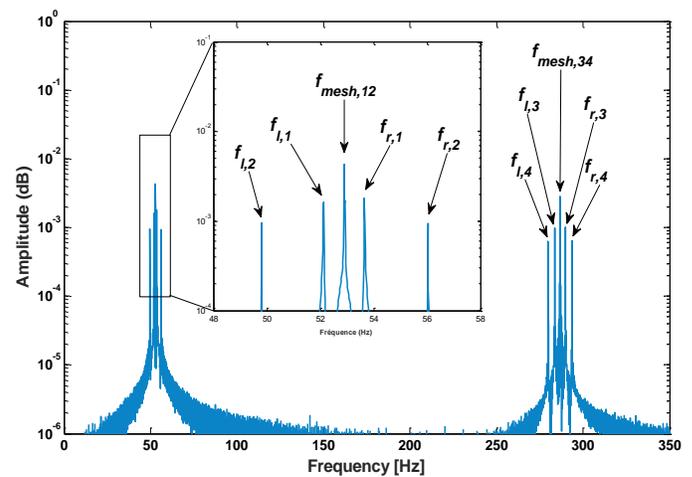


Fig. 5. Spectrum of mixture 1, in healthy mode.

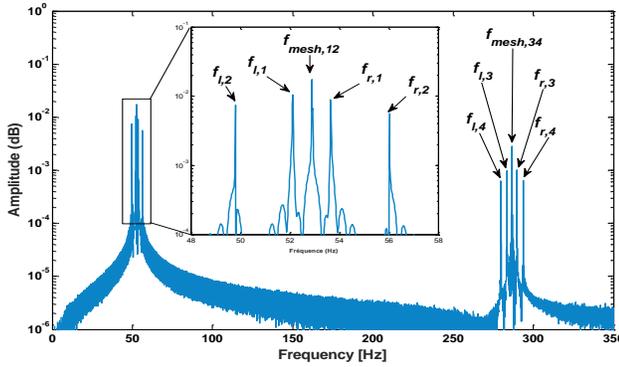


Fig. 6. Spectrum of mixture 1, under a uniform wear of R1 and R2.

For instance, for the frequency-band centered at  $f_{mesh34}$ , the spectrum of mixture 1 in the faulty mode, presented in Figure 6, is similar to the spectrum of this mixture in the healthy mode, in Figure 5. This result shows that the wheels R3 and R4 are healthy.

On the other hand, for the frequency-band centered in  $f_{mesh12}$ , the spectrum of mixture 1 in the fault mode, in Figure 6, is different from the spectrum of the same mixture in the healthy mode, in Figure 5. This proves that the wheels R1 and R2 are affected by the uniform wear fault.

### C. Study of the whitening preprocessing

The mixtures are firstly whitened using Principal Component Analysis (PCA) technique.

Let  $x$  be the sample composed of points in  $\mathbb{R}^4$  extracted from the mixtures  $X$  in the faulty mode, Figure 7. PCA computes the two following moments of  $x$

- the arithmetic mean

$$\mu = (0 \ 0 \ 0 \ 0)^T \quad (22)$$

- the covariance matrix

$$C = 10^{-4} \begin{pmatrix} 17 & 18 & 15 & 13 \\ 18 & 21 & 17 & 14 \\ 15 & 17 & 14 & 12 \\ 13 & 14 & 12 & 10 \end{pmatrix} \quad (23)$$

The points of  $x$  are centered relatively to  $\mu$ . It comes the two matrices

- $D$ : matrix whose diagonal values are the eigenvalues of  $C$

$$D = \begin{pmatrix} 2,610^{-6} & 0 & 0 & 0 \\ 0 & 6,310^{-6} & 0 & 0 \\ 0 & 0 & 1,710^{-4} & 0 \\ 0 & 0 & 0 & 6,110^{-3} \end{pmatrix} \quad (24)$$

- $E$ : matrix whose columns are the eigenvectors of  $C$

$$E = (\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4) = \begin{pmatrix} -0.1619 & -0.3393 & 0.7673 & 0.5196 \\ 0.3354 & -0.5046 & -0.5449 & 0.5796 \\ -0.7269 & 0.3911 & -0.3030 & 0.4763 \\ 0.5770 & 0.6909 & 0.1503 & 0.4089 \end{pmatrix} \quad (25)$$

In Figure 7 and 8, only the three eigenvectors  $\vec{v}_1$ ,  $\vec{v}_2$  and  $\vec{v}_3$  are displayed.

And, then, the obtained whitening matrix  $U$  expressed in (8) based on  $D$  and  $E$

$$U = \begin{pmatrix} -100.56 & 208.38 & -451.68 & 358.48 \\ -135.20 & -201.11 & 155.86 & 275.35 \\ 58.75 & -41.72 & -23.20 & 11.50 \\ 6.65 & 7.42 & 6.10 & 5.23 \end{pmatrix} \quad (26)$$

Therefore, the whitened mixtures  $V$  is the projection of  $X$  into  $U$ , as shown in equation (9). These resulting mixtures  $V$  are more appropriate for the source separation than the original mixtures  $X$ . Indeed, the sample  $v$  composed of the points belonging to  $V$ , shown in Figure 8, has the following appealing characteristics

- the zero arithmetic mean:

$$\mu = (0 \ 0 \ 0 \ 0)^T \quad (27)$$

- the identity covariance matrix

$$C' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (28)$$

- the orthonormal basis composed by the eigenvectors corresponding to the columns of  $E$ .

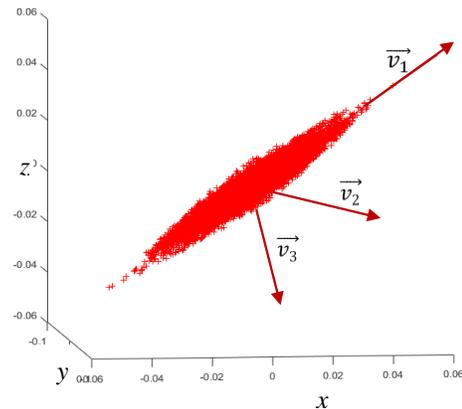


Fig. 7. Dispersion of points from mixtures before Whitening.

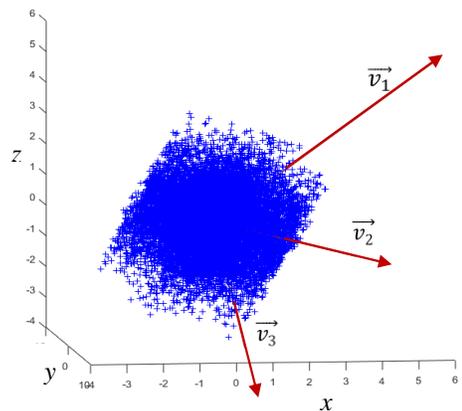


Fig. 8. Dispersion of points from mixtures after Whitening

D. Study of the source separation processing

In this section, the performance of the Fast-ICA algorithm is evaluated for the source separation task.

1) Performance measures

In the context of the separation of vibratory signal sources, performance measurement is an essential task for assessing separation quality. Therefore, the following measures are adopted [30]

- the Signal-to-Distortion Ratio (SDR)

$$SDR = 10 \log_{10} \left( \frac{\|S_j^{target}\|^2}{\|e_j^{inter} + e_j^{artif}\|^2} \right) \quad (29)$$

- source-to-Interference Ratio (SIR)

$$SIR = 10 \log_{10} \left( \frac{\|S_j^{target}\|^2}{\|e_j^{inter}\|^2} \right) \quad (30)$$

- the Souce-to-Artifact Ratio (SAR)

$$SAR = 10 \log_{10} \left( \frac{\|S_j^{target} + e_j^{interf}\|^2}{\|e_j^{artif}\|^2} \right) \quad (31)$$

where

- $S_j^{target} = f(S_j)$  is a version of the original source modified using an allowed distortion  $f \in F$ , such that  $F$  encompasses several time-invariant gains distortions,
- $e_j^{interf}$  and  $e_j^{artif}$  are, respectively, the error terms relative to interferences and artifacts.

2) Results and discussion

Our goal in this section is to identify the non-linearity results in the best performance of source separation using Fast-ICA algorithm. Furthermore, the experiments are conducted with healthy and faulty gears.

In the case of healthy mode, the results of table II are obtained. It is obvious that the tanh outperforms the other non-linearities. Indeed, it gives rise to the highest average values of SIR: 69.3 and SDR: 41.47.

Kurtosis gives the second best performance in terms of SIR: 65.87 and SDR: 38.02. The gauss non-linearity gives significantly lower average values of SIR: 27.23 and SDR: 21.95.

The worse performance is obtained by Skew non-linearity. It gives very low average value of SIR: 3.96 and SDR: 3.97. All the non-linearities result in close high SAR values between 41 and 44. Therefore, Fast-ICA leads to low overlap artifact in the estimated sources. In the case of faulty mixtures, the results of table III are obtained. Particularly, Fast-ICA performs the best separation using tanh non-linearity. It results in SIR average value 81.27 and SDR average value 48.15.

What is interesting in this faulty mode is that the sources of the damaged gear wheels R1 and R2 have been well separated based on tanh. Indeed, tanh gives rise the highest values of SIR: 88.67 and SDR: 59 for the estimated source of damaged wheel R1. Similarly, tanh gives the highest values of SIR: 75.75 and SDR: 38.85 for the estimated source of damaged wheel R2.

Moreover, the obtained average SAR values are between 45 and 48 for all the non-linearities. These values are higher than average SAR values obtained in the case of healthy mode. Therefore, Fast-ICA results in less overlap artifact in the faulty mode.

TABLE II. PERFORMANCE MEASURE OF SOURCE SEPARATION IN HEALTHY MODE

		Tanh	Kurtosis	Gauss	Skew
<b>Nb of iterations</b>		5	5	7	15
<b>SIR (dB)</b>	R1	70.4	70,1	60.8	5.7
	R2	64.3	61,4	3.86	1.98
	R3	65.9	62	40.4	2.4
	R4	76.6	70	3.86	5.77
	Average	69.3	65.87	27.23	3.96
<b>SDR (dB)</b>	R1	37.69	35.8	44.3	5.7
	R2	45	42.23	3.8	1.98
	R3	38.8	40.39	35.9	2.44
	R4	44.4	33.69	3.8	5.77
	Average	41.47	38.02	21.95	3.97
<b>SAR (dB)</b>	R1	37.69	38.8	44.4	47.63
	R2	45	45.26	56.6	40.19
	R3	38.8	44.41	37.7	36.34
	R4	44.4	37.69	37.9	46.86
	Average	41.47	41.54	44.15	42.75

TABLE III. PERFORMANCE MEASURE OF SOURCE SEPARATION IN FAULT MODE

		Tanh	Kurtosis	Gauss	Skew
<b>Nb of iterations</b>		5	5	7	15
<b>SIR (dB)</b>	R1	88.67	71.94	84.7	2.39
	R2	75.75	72.01	71.5	-4.1
	R3	73	71.07	70.3	1.3
	R4	87.67	86.5	66.4	3.5
	Average	81.27	75.38	73.22	0.77
<b>SDR (dB)</b>	R1	59	37.68	55	2.39
	R2	38.85	36.7	37.7	-4.13
	R3	37	38.8	38.8	1.37
	R4	57.75	59	37.68	3.5
	Average	48.15	43.04	42.29	0.78
<b>SAR (dB)</b>	R1	59	37.68	59	41.54
	R2	38.85	57.75	57.7	37.89
	R3	37.68	38.85	38.8	61.29
	R4	57.75	59	37.6	41.6
	Average	48.32	48.32	48.27	45.58

IV. SPECTRAL ANALYSIS

In this section, the results obtained by the Fast-ICA algorithm are studied. First, Fast-ICA converges quickly in up to 15 iterations, which confirms that this algorithm is a fast variant of the ICA. On the other hand, in our experiments, Fast-ICA is applied on two types of mixing: healthy mode and fault mode. Therefore, two questions that arise: Is the Fast-ICA able to separate the gear signals associated to the four

wheels R1, R2, R3 and R4? And, in the case of a fault mode, can it distinguish between damaged wheels and healthy wheels ?

By observing the spectrums of the estimated sources, Fast-ICA succeeded in separating the gear signals corresponding to each wheel. The spectrum presented in Figure 9 is composed of a fundamental frequency  $f_{mesh,12}$  and two lateral frequencies  $f_{l,1}$  and  $f_{r,1}$ . Thus, the source 1 corresponds to the wheel R1. The second spectrum illustrated in Figure 10 is composed of a fundamental frequency  $f_{mesh,12}$  and two lateral frequencies  $f_{l,2}$  and  $f_{r,2}$ . Thus, the source 2 identifies the wheel R2. The third spectrum shown in Figure 11 is composed of a fundamental frequency  $f_{mesh,34}$  and two lateral frequencies  $f_{l,3}$  and  $f_{r,3}$ , leading to a clear identification of wheel R3. Finally, the spectrum given in Figure 12 is composed of a fundamental frequency  $f_{mesh,34}$  and two lateral frequencies  $f_{l,4}$  and  $f_{r,4}$ . Thus, the source 4, corresponding to the wheel R4, is clearly identified.

By comparing the results obtained in fault mode with the results obtained in healthy mode, Fast-ICA distinguishes between the faulty sources and the healthy sources of the gears

- The wheel R1 has two slightly different spectrums. Indeed, the spectrum of the wheel R1 mentioned in Figure 13 is slightly different from the spectrum of R1 in Figure 9 which is in a healthy mode.
- The wheel R2 has two slightly different spectrums. Indeed, the spectrum of the wheel R2 mentioned in Figure 14 is slightly different from the spectrum of R2 in Figure 10 which is in a healthy mode.
- On the other hand, the other two wheels R3 and R4 are healthy. Indeed, each of these two wheels keeps almost the same spectrum in the healthy mode and in the fault mode, as shown in Figures (15, 11) and Figures (16, 12) respectively.

## V. CONCLUSION

In this paper, a diagnostic technique is presented for separating and identifying uniform wear in two-stage gearbox, classically associated to a double-fed induction machine in modern wind energy conversion systems. Based on Fast-ICA, the main contribution of the proposed technique is its ability to isolate the fault frequency components, representative of a uniform wear for each pinion or gear of the gearbox.

The obtained results show also clearly the ability of the Fast-ICA for separating the characteristic frequency components of the gears from noisy mixtures. Moreover, the spectral analysis allows us to distinguish for each estimated source associated to a gear whether it is healthy or faulty. As a perspective, further faults than gear fault would be taken into account in a future work.

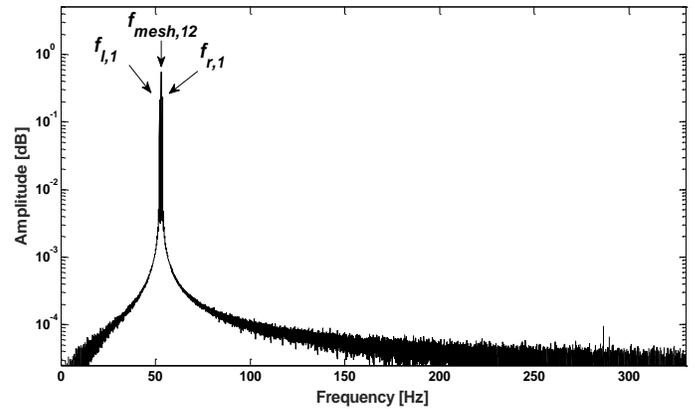


Fig. 9. Spectrum of signal resulting from the estimation of component relative to the wheel R1, in healthy mode.

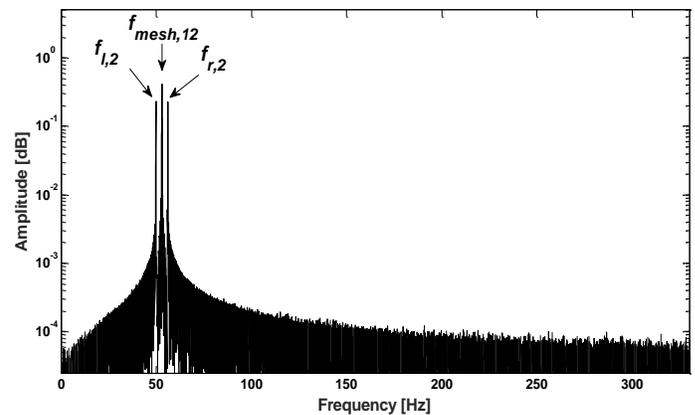


Fig. 10. Spectrum of signal resulting from the estimation of component relative to the wheel R2 in healthy mode.

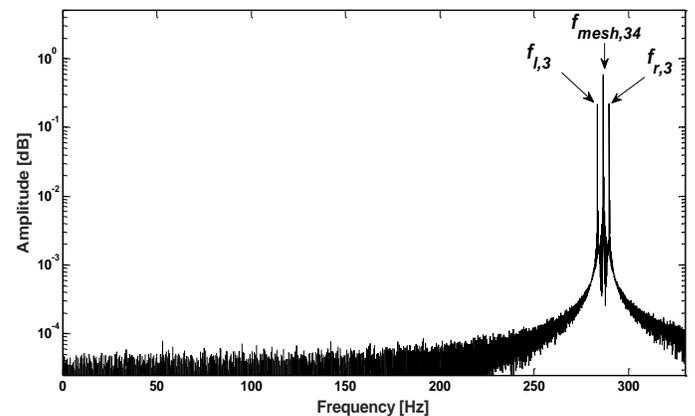


Fig. 11. Spectrum of signal resulting from the estimation of component relative to the wheel R3, in healthy mode.

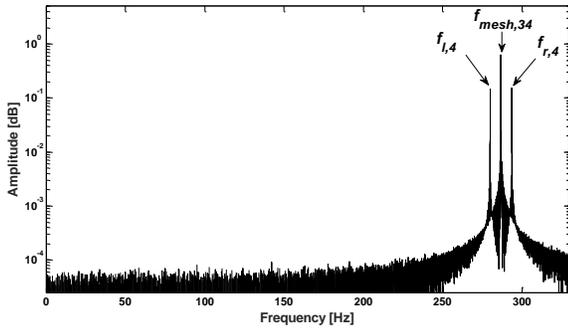


Fig. 12. Spectrum of signal resulting from the estimation of component relative to the wheel R4, in healthy mode.

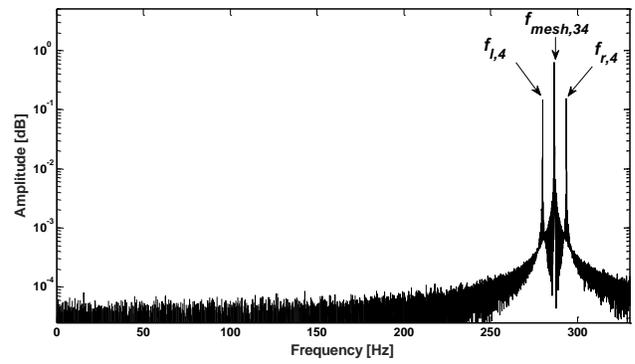


Fig.

Fig. 16. Spectrum of signal resulting from the estimation of component relative to the wheel R4, in fault mode with uniform wear of R1 and R2.

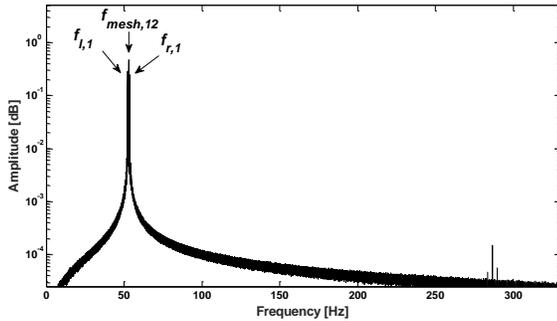


Fig. 13. Spectrum of signal resulting from the estimation of component relative to the wheel R1, in fault mode with uniform wear of R1 and R2.

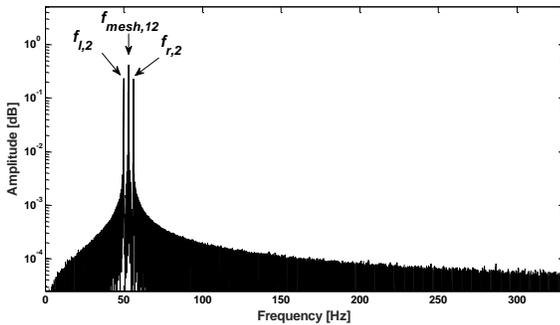


Fig. 14. Spectrum of signal resulting from the estimation of component

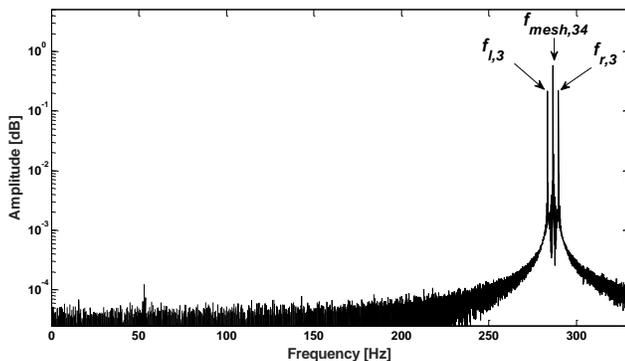


Fig. 15. relative to the wheel R2, in fault mode with uniform wear of R1 and R2.

#### REFERENCES

- [1] H. Polinder, F.A. Van Der Pijl, G-J. de Vilder and P.J. Tavner, "Comparison of Direct-Drive and Geared Generator Concepts for Wind Turbines", IEEE Transactions on Energy Conversion, vol. 21, no. 3, Sept. 2006, pp. 725-733.
- [2] H. Li, Z. Chen, "Design optimization and evaluation of different wind generator systems", The 2008 International Conference on Electrical Machines and Systems, 17-20 Oct. 2008, pp. 2396 – 2401.
- [3] Y. Amirat, M.E.H. Benbouzid, E. Al-Ahmar, B. Bensaker, S. Turri, "A brief status on condition monitoring and fault diagnosis in wind energy conversion systems", Renewable and Sustainable Energy Reviews, vol. 13, no. 9, Dec. 2009, pp. 2629–2636.
- [4] M. Todorov, I. Dobrev, and F. Massouh, "Analysis of Torsional Oscillation of the Drive Train in Horizontal-Axis Wind Turbine", Electromotion 2009 – EPE Chapter 'Electric Drives' Joint Symposium, Lille, 1-3 July 2009.
- [5] N. Baydar, A. Ball, "Detection of gear failures via vibration and acoustic signals using wavelet transform", Mechanical Systems and Signal Processing, vol. 17, no. 4, pp. 787-804, 2003.
- [6] S. Slim, Van L. Paul, P. Asanka, Gan Tat-Hean, B. Bryan", Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring", Renewable Energy, vol. 51, 2013, pp. 175–81.
- [7] M. Heydarzadeh, M. Nourani, J. Hansen and S.H. Kia, "Non-invasive gearbox fault diagnosis using scattering transform of acoustic emission", IEEE-ICASSP, New Orleans, pp. 371-375, March 2017.
- [8] J. Wei and J. McCarty, "Acoustic emission evaluation of composite wind turbine blades during fatigue testing", Wind Eng., vol. 17, no. 6, pp. 266-274, 1993.
- [9] D. Coronado, C. Kupferschmidt, "Assessment and validation of oil sensor systems for on-line oil condition monitoring of wind turbine gearboxes", 2nd International Conference on System-Integrated Intelligence: Challenges for Product and Production Engineering, vol. 15, pp 748 –755, 2014.
- [10] D. Dyer, R.M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis", Trans ASME, J Mech Design, vol. 100, no. 2, pp. 229–235, 1978.
- [11] N.J. Cozens, S.J. Watson, "State of the art condition monitoring techniques suitable for wind turbines and wind farm applications", Report for CONMOW, Project 2003.
- [12] Y. Gritli, A. Di Tommaso, R. Miceli, C. Rossi and F. Filippetti, "Diagnosis of mechanical unbalance for double cage induction motor load in time-varying conditions based on motor vibration signature analysis", International Conference on Renewable Energy Research and Applications, Madrid, pp. 1157-1162, 2013.
- [13] R. Miceli, Y. Gritli, A. Di Tommaso, F. Filippetti and C. Rossi, "Vibration signature analysis for monitoring rotor broken bar in double squirrel cage induction motors based on wavelet analysis", The

- International Journal for Computation and Mathematics in Electrical and Electronic Engineering, vol. 33, pp. 1625-1641, 2014.
- [14] V. Capdevielle, C. Serviere and J.L. Lacoume, "Blind separation of wide-band sources: Application to rotating machine signals", 8th European Signal Processing Conference, Trieste, September, 1996.
- [15] T.D. Popescu, "Blind separation of vibration signals and source change detection – Application to machine monitoring", Applied Mathematical Modelling, vol. 34, pp. 3408-3421, 2010.
- [16] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications", Neural networks, vol. 13, pp. 411–430, 2000.
- [17] F. Miao and R. Zhao, "Application of Independent Component Analysis in machine fault diagnosis", Advanced Materials Research, vol. 905, pp. 524-527, 2014.
- [18] Z. Xing yuan, T. Ji Liang, and D. De cun, "Separation of the Train Vibration Signal Based on Improved Fast-ICA", IEEE Computer Society, Fourth International Conference on Intelligent Computation Technology and Automation, Washington, Vol. 1, pp. 912-914, 2011.
- [19] H. Nakamura and al., "The application of independent component analysis to the multi-channel surface electromyographic signals for separation of motor unit action potential trains: part I – measuring techniques", J. Electromyogr Kinesiol, vol. 14, pp. 423-432, 2004.
- [20] D. Obradovic, N. Madhu, A. Szabo, C.S. Wong, "Independent component analysis for semi-blind signal separation in MIMO mobile frequency selective communication channels", IEEE International Joint Conference on Neural Networks, Munich, pp. 53–58, 2004.
- [21] A. Ypma, A. Leshem, and R.P. Duijn, "Blind separation of rotating machine sources: bilinear forms and convolutive mixtures", Neurocomputing, vol. 49, pp. 349–368, 2002.
- [22] G. Yu, "Fault feature extraction using independent component analysis with reference and its application on fault diagnosis of rotating machinery", Neural Computing and Applications, vol. 26, pp. 187-198, 2015.
- [23] A-J. Bell, T-J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution", Neural Comput., vol. 7, pp. 1129–1159, June 1995.
- [24] J.F. Cardoso, A. Souloumiac, "Blind beamforming for non gaussian signals", in IEE Proceedings-F., vol. 140, no. 6, pp. 362–370, Dec. 1993.
- [25] J. Zhang, A. C. Sanderson, "JADE: Self Adaptive Differential Evolution with fast and reliable convergence performance", IEEE Congress on Evolutionary Computation, Troy, vol. 13, no 5, pp. 945-958, 2009.
- [26] A. Hyvarinen and E. Oja, "A fast fixed-point algorithm for independent component analysis", Neural Computation, vol. 9, pp. 1483–1492, 1997.
- [27] A. Hyvarinen, J. Karhunen and E. Oja, "Independent Component Analysis", Hoboken, NJ: Wiley-Interscience, 2001.
- [28] M. Farhat, Y. Gritli, M. Benrejeb, "Application de l'analyse en composantes indépendantes pour l'identification de défauts mécaniques d'une chaîne de conversion éolienne", ICME International Conference on Mechatronics Engineering, Sousse, 2017.
- [29] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", IEEE Trans. on Neural Networks, vol. 10, no. 3 pp. 626-634, mai 1999.
- [30] E. Vincent, R. Gribonval and C. Févotte, "Performance measurement in blind audio source separation", IEEE Transactions on Audio, Speech and Language Processing, vol. 14, no. 4, pp. 1462–1469, July 2006.

# An Enhanced Approach for Detection and Classification of Computed Tomography Lung Cancer

Wafaa Alakwaa  
Faculty of Computers & Info.  
Cairo University, Egypt

Mohammad Nassef  
Faculty of Computers & Info.  
Cairo University, Egypt

Amr Badr  
Faculty of Computers & Info.  
Cairo University, Egypt

**Abstract**—The paper presents approaches for nodule detection and extraction in axial lung computed tomography. The goal is to detect correctly pulmonary nodule to recognize and screen lung cancer patients. The pulmonary nodule detection is very challenging problem. The proposed model developed a hybrid efficient model based on affine-invariant representation and shape of segmented nodule. Due to large number of extracted features for all slices on patient, feature selection is an important step to select the most important feature for classification. We apply forward stepwise least squares regression that maximizes the R-squared value, this criterion provides a fast preprocessing feature selection assessment for systems with huge volumes of features based on a linear models framework. Moreover, gradient boosting have been suggested to select the relevant features based on boosting approach. Classification of patients has been done by support vector machine. Kaggle DSB dataset is used to test the accuracy of our model. The results show major improvement in accuracy and the features are reduced.

**Keywords**—Lung cancer; computed tomography; affine invariant moments; pulmonary nodules; R2; feature selection; support vector machine

## I. INTRODUCTION

Lung cancer has the second highest incidence of cancers worldwide for both the male and female population, and remains the cancer with the highest mortality. This is because it remains asymptomatic for a long time, and is therefore diagnosed mostly at such a late stage that treatment outcome is poor. Despite this, most countries currently do not have a lung cancer screening programme for early detection of lung cancer. This is not only due to the high costs involved if applied to a large proportion of the population, but also the lack of a sufficiently sensitive diagnostic test, including imaging. Current research in screening for lung cancer is therefore limited to patients identified at high risk of developing lung cancer, such as smokers or patients with COPD (or both), but it is anticipated that this research could form an important foundation for a future national screening programme [1].

CT scan is an extended version of X-ray in which computer is attached to the X-ray machine. Pictures that are taken from angles and distances are processed in the computer and presented in the 3-dimensional, cross-sectional (tomographic) and in slices form. In this way, bones, tissues, blood vessels, and organs are shown up clearly. The imaging of CT scan is

useful for diagnosis, treatment and progress of medication. Recently, helical or multi-slice scanning is introduced that almost eliminated gaps in the collection of slides [2]. The radiologists miss detecting lung nodules in early stage due to dramatic expanding in number of image slices in high resolution images. A lung Cancer screening computer-aided detection/diagnosis (CAD) system can reduce cost and speed up screening. CAD systems help radiologists in building decisions and enhance process of detection and observation of diseases in screening. CAD can enhance nodule detection step by detecting missed nodules, reduce reading time so that the screening process is made possible and helps differentiate between benign and malignant lesions.

In this paper, we introduce an efficient model to detect and diagnose lung cancer patients. Based on watershed segmentation, nodules are detected and shape features are applied to describe the nodules using affine moments. Gradient boosting is used which can identify a robust feature selection through ensemble learning by combining weak classifiers to yield strong, robust and accurate classifier. The variations in the target classes are identified by the best selected features through R-Squared regression criterion.

The paper's arrangement is describes as follows: Related work is summarized briefly in Section II. The model architecture is presented in Section III. The nodule segmentation is introduced in Section IV based on watershed algorithm. The feature extraction process based on affine moments and shape features are presented in Section V. Section VI presents feature selection models based on ensemble-based feature selection models which include Gradient Boosting and regression-based feature selection using R-squared model. The classification process using SVM is mentioned in Section VII. Our discussion and results are described in details in Section VIII. Section IX summarizes the conclusion of paper.

## II. RELATED WORK

Recent research tries to encourage developing an image-based model that is able to improve, as a second opinion, in conjunction with the radiologist, the detection accuracy of a radiologist, and reduce mistakes related to false positives. A CAD system generally consists of several steps when processing medical images. Images are preprocessed to remove

noise and enhance quality. Then Region of Interest is segmented from other structures. Features are extracted from these ROIs, such as geometrical, textural, and statistical features. Accordingly, a classification step is done, to decide if the image contains malignant nodule. There has been exhaustive efforts on computer aided diagnosis for lung images.

In [3], [4], Hamada et al. evaluated their system on the Japanese Society of Radiological Technology (JSRT) standard dataset of chest radiographs. The two preprocessing techniques were histogram equalization and Laplacian filter. Contrast was enhanced and the rapid intensity change was examined. Wavelet transform was used for feature extraction. To select the most important features the proposed model calculated the variance and the energy. The dimensions of the overall features is then reduced. For classification K-nearest neighbor classifier was employed. The proposed model was tested on 154 nodule regions with 100 malignant and 54 benign nodules. The Accuracy was 99.15% for normal versus abnormal and 98.70% for benign and versus malignant.

In [5] many techniques were applied for lung region detection. Bit plane slicing algorithm is used to generate different binary slices which then were enhanced by erosion algorithm and dilation and median filters. After detection of lung region, segmentation was applied to identify the lung nodules. Fuzzy Possibilistic C Mean (FPCM), which is a clustering algorithm that combines the characteristics of a fuzzy and possibility c-means, was applied for segmentation. Area and the mean intensity value of the candidate region are the features that were used to classify the nodule on. Support Vector Machine was used for binary classification. The proposed model was tested on experimentation data consists of 1000 lung images obtained from the reputed hospital.

In [6], Ada and Rajneet K. proposed a hybrid approach on feature extraction and Principal Component Analysis (PCA). Histogram Equalization is used for preprocessing of the images. Features were Extracted using Binarization and Masking Approach. A Grey Level Co-occurrence Method was created to make different combinations of pixel brightness. The features used in this approach were entropy, contrast, energy, and maximum probability. The exact output and results were not clearly specified.

In [7], FFT, Auto enhancement and Gabor filtering were used for image enhancement step. Topology surface and watershed algorithm were applied to the marker location and segmentation progress. The features that were extracted from ROI were area, perimeter, eccentricity and average intensity. In [8] Kamil Dimililer et al. used many image processing techniques: grayscale conversion, thresholding, erosion, median filtering and image subtraction.

In [9], many filters were applied in the preprocessing step, such as low pass filters, contrast stretching histogram equalization, negativity and power law transformation. For segmentation modified thresholding, labeling algorithm and edge detection were taken off. Features such as geometric properties, textural properties and mathematical properties were calculated. Gray Level Co-occurrence Matrix (GLCM) is a used to examine relationship of image pixels.

In [10] a computer aided diagnosing system was proposed to detect lung cancer based on texture features take out from

the slice of DICOM Lung CT images. For preprocessing step K Nearest Neighbors and Weiner filters were used. Sobel Methods was suggested for segmentation. The set of texture features that were used for diagnosis are area of the interest, Calcification, Shape, Size of nodule and Contrast Enhancement. Artificial neural network was used for classification. This CAD system neglects all the false positive cancer regions and detects the cancer regions. The used dataset was obtained from NIH/NCI Lung Image Database Consortium (LIDC). There were about 1000 lung images. This approach showed sensitivity of 90% with 0.05 false positives per image.

In [11], Hashemi et al. proposed a system based on fuzzy inference. Starting with image enhancement and noise removal, Linear-Filtering was used. A region growing based technique was used for segmentation. A Fuzzy Inference System was implemented to determine the type of the mass diagnosed. The system was 95% accurate. Features such as area and color were used. This method was tested on 1000-tumor contained 10000 CT slices from 1000 lung tumor patients. The accuracy of the proposed systems was 95%.

### III. PROPOSED METHODOLOGY

The suggested framework of Lung cancer detection and classification is composed of four stages: Nodule segmentation, feature extraction, feature selection and patient classification. As shown in Fig. 1, the overall architecture is drawn. Watershed segmentation is used to detect the nodule in lung cancer slices in CT scan as feature detection. Then, shape features and invariant affine moments are applied to describe the extracted nodules. For feature selection, we developed ensemble models and regression model to select the best important and relevant features to avoid the over-fitting problems. Finally, patient is classified by SVM. The main tasks of our model are presented in details through the next sections.

### IV. NODULE DETECTION AND SEGMENTATION

One of the main tasks in medical diagnosis is the segmentation, especially in lung cancer using CT scans. Segmentation is a commonly preprocessing step for more enhancement in anomalies and lung structures, such as nodules.

The watershed algorithm is a common segmentation technique based on morphology mathematics. It depends on an intensity based topographical representation. The higher altitudes (hills) are represented by brighter pixels and the valleys are represented by dark pixels to determine the path of a falling raindrop would follow. The different regions are separated by watershed lines in watershed algorithm. Fig. 2 shows the resulted nodules extracted using watershed.

### V. FEATURE EXTRACTION

#### A. Moments Invariant Features

Moments are applied in many applications. Many of these techniques are essentially based on the general moment theory widely known and applied in research in several areas of statistics and mechanics. In particular, geometric moments have vast practical applications in many area of computer vision and invariant pattern recognition, ranging from lower-level recognition such as pose estimation to higher-level recognition such as

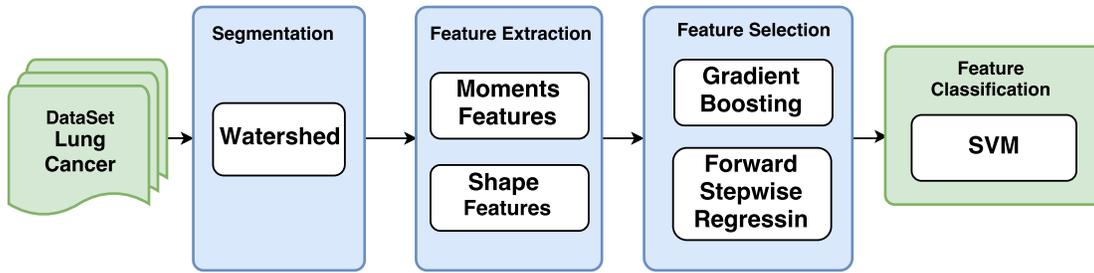


Figure 1: Pipeline of the proposed model

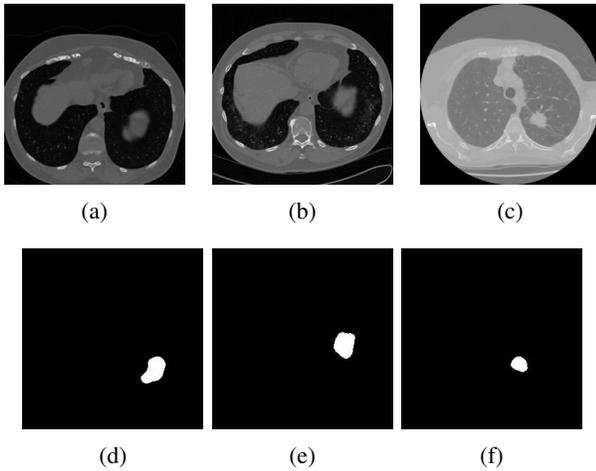


Figure 2: An axial slices of CT scans for Three Different Patients 2a, 2b, 2c and Segmented Nodules Based on Watershed Algorithm for Each Patient 2d, 2e, and 2f

activity recognition and analysis. When applied to images, they were identified to be most descriptive of the image contents (i.e., intensity distribution) with respect to its axes. Once such moments are properly defined, both global and detailed geometric information of image contents can be reasonably expected to be detected robustly. In such a scenario, moments would be able to characterize various image objects such that the properties with analogies in statistics or mechanics are extracted, and thus the shape of all objects of interest can be described well. Formally speaking, in continuous domain, an image is viewed as a 2-D Cartesian density distribution function  $f(x, y)$ . The general form of the geometric moments of order  $(p + q)$  for the function  $f(x, y)$ , evaluated over the entire plane  $\Omega$  is defined by the following discrete form:

$$M_{pq} = \sum_y \sum_x \varphi_{pq}(x, y) I(x, y), \quad p, q = 0, 1, 2, \dots, \infty \quad (1)$$

Where,  $\varphi_{pq}$  is a basis function or weighting kernel by which a weighted description for the image function  $f(x, y)$  across the entire plane  $\Omega$  is generated. It is perhaps worthwhile to point out here that the choice of above basis functions  $\varphi_{pq}$  greatly depends on the application of use, and on the invariant properties desired. Furthermore, it is expected that choosing a

specific basis function results in some constraints, such as to restrict the range of the image coordinates,  $x$  and  $y$ , enable the image and its descriptors to be translated to other coordinates (e.g., polar coordinates), etc. In [12], Hu stated that the 2-D Cartesian moment of order  $(p + q)$  for an  $m \times n$  discretized image,  $I(x, y)$  can be defined by taking the basis function in (1) as a monomial of power  $p + q$  (product of powers of the variables  $x$  and  $y$ , i.e.,  $\varphi_{pq}(x, y) = x^p y^q$  as follows:

$$M_{pq} = \sum_{y=0}^{n-1} \sum_{x=0}^{m-1} x^p y^q I(x, y), \quad p, q = 0, 1, 2, \dots, \infty \quad (2)$$

The full moment set of order  $k$  that includes all moments,  $M_{pq}$ , such that  $p + q \geq k$  comprises of exactly  $\frac{1}{2}(k + 1)(k + 2)$  elements. Ever since the pioneering work of Hu [12] on moment functions that has explored quite thoroughly the use of moments for image analysis and object representation, a broad range of new applications utilizing moment invariants in image analysis and pattern recognition fields has started to evolve. It is clear that the Cartesian moments given by (2) are not invariant to geometric transformations. To achieve invariance under translation, these moments are calculated with respect to the center of mass as follows:

$$\mu_{pq} = \sum_{y=0}^{n-1} \sum_{x=0}^{m-1} (x - \bar{x})^p (y - \bar{y})^q I(x, y), \quad p, q = 0, 1, 2, \dots, \infty \quad (3)$$

Where,  $\bar{x}$  and  $\bar{y}$  are the coordinates of the centroid and given by:

$$\bar{x} = \frac{M_{10}}{M_{00}}, \quad \bar{y} = \frac{M_{01}}{M_{00}} \quad (4)$$

After a bit tedious but straightforward manipulation, (2) and (3) lead to the following relation between the Cartesian and centralized moments:

$$\mu_{pq} = \sum_i^p \sum_j^q \binom{p}{i} \binom{q}{j} (-\bar{x})^{p-i} (-\bar{y})^{q-j} M_{ij} \quad (5)$$

However, it should be emphasized that the expression in (3) suggests that the centralized moments are only invariant to translation. To enable invariance under scale changes, the

2-D centralized moments  $\mu_{pq}$  need to be normalized to obtain scale-normalized centralized moments  $\eta_{pq}$  as follows:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}}} \quad (6)$$

Where, the exponent  $\gamma$  is given in terms of p and q as follows:

$$\gamma = \frac{p+q}{2} + 1, \quad p+q \geq 2$$

Strictly speaking, the moments present the shape properties for appearance of a nodule. Affine moments are invariant under six transform and derived based on central moments [13] as follows:

$$\begin{aligned} I_1 &= \frac{1}{\eta_{00}^4} [\eta_{20}\eta_{02} - \eta_{11}^2], \\ I_2 &= \frac{1}{\eta_{00}^{10}} [\eta_{03}^2\eta_{30}^2 - 6\eta_{30}\eta_{21}\eta_{12}\eta_{03} + 4\eta_{30}\eta_{12}^3 \\ &\quad + 4\eta_{03}\eta_{21}^3 - 3\eta_{21}^2\eta_{12}^2], \\ I_3 &= \frac{1}{\eta_{00}^7} [\eta_{20}(\eta_{21}\eta_{03}\eta_{21} - \eta_{12}^2) - \eta_{11}(\eta_{30}\eta_{03} - \eta_{21}\eta_{12}) \\ &\quad + \eta_{02}(\eta_{03}\eta_{12} - \eta_{21}^2)], \\ I_4 &= \frac{1}{\eta_{00}^{11}} [\eta_{20}^3\eta_{03}^2 - 6\eta_{20}^2\eta_{11}\eta_{12}\eta_{03} - 6\eta_{20}^2\eta_{02}\eta_{21}\eta_{03} \\ &\quad + 9\eta_{20}^2\eta_{02}\eta_{12}^2 + 12\eta_{20}\eta_{11}^2\eta_{21}\eta_{03} + 6\eta_{20}\eta_{11}\eta_{02}\eta_{30}\eta_{03} \\ &\quad + 18\eta_{20}\eta_{11}\eta_{02}\eta_{30}\eta_{12} - 8\eta_{11}^3\eta_{30}\eta_{03} - 6\eta_{20}\eta_{02}^2\eta_{30}\eta_{12} \\ &\quad + 9\eta_{20}\eta_{02}^2\eta_{21}^2 + 12\eta_{11}^2\eta_{02}\eta_{30}\eta_{12} + \eta_{02}^3\eta_{30}^3], \\ I_5 &= \frac{1}{\eta_{00}^6} [\eta_{40}\eta_{04} - 4\eta_{31}\eta_{13} + 3\eta_{22}^2], \\ I_6 &= \frac{1}{\eta_{00}^9} [\eta_{40}\eta_{04}\eta_{22} - 4\eta_{31}\eta_{13}\eta_{22} - \eta_{40}\eta_{13}^2 - \eta_{04}\eta_{13}^2 - \eta_{22}^2] \end{aligned} \quad (7)$$

## B. Shape Features

After segmentation, the nodule candidate is selected and two different types of features are extracted, namely, 2-D geometric, 3-D geometric. A median slice  $I_{NC,m}$  is extracted from 2-D features because the area of the segmented object is the largest. The shape of nodule candidates are worthy features to recognize the objects in Lung. The shape of nodules are described as 2-D and 3-D geometric features. Area, Perimeter, and Eccentricity are the most common used in our paper to describe the segmented regions in lung cancer slices.

## VI. FEATURE SELECTION

Feature selection is a worthy stage in medical diagnosis to choose the best features that enhance the model accuracy. Furthermore, the models can be simpler and faster in understanding and building with the least number of features. In our paper, we applied two approaches: regression-based feature selection and tree-based feature selection.

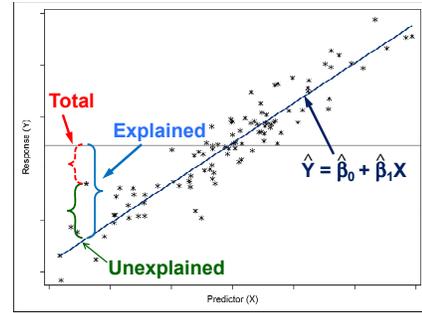


Figure 3: Explained vs. Unexplained variability.

### A. Regression-Based Feature Selection

Fitting regression line with statistical measure of how the data is close called R-squared statistical measure or coefficient of determination. It employs a forward step-wise least squares regression that optimize the model r-squared value. It is used in huge data as a preparatory step to assess the features very fast and can identify quickly the useful features. The variations in target that explained by single feature with deleting the calculations of other features called squared correlation coefficient. The range of values between 0 and 1 (1 means the input feature can explain totally the variation in target) and 0 denotes that the target and input feature have not a relationship. In lung cancer recognition, the calculated squared correlation coefficient in a simple linear regression for all input features that are interval is mentioned as follows:

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (8)$$

where  
X: input feature,  
Y: response variable or target,  
 $\beta_0$ : intercept parameter,  
 $\beta_1$ : slope parameter  
 $\varepsilon$ : error deviation of Y about  $\beta_0 + \beta_1 X$ .

The feature that explains the target is a worthy feature, thus it is selected in simple linear regression. In a baseline model, the target class and the input features have not a relationship. thus, any feature value does not improve predictions of the target class over simply using the mean of the target class for everyone.

R-Squared the ratio of variations explained via regression line in the observed data. The R-Squared is equal to  $R^2 = \frac{SSM}{SST} = 1 - \frac{SSE}{SST}$ , where SSM indicates sum square of model. It is the total variations explained by regression model and equal to  $SSM = \sum (\hat{Y}_i - \bar{Y})^2$ , SSE indicates to sum square of error. It is the total variations unexplained by regression model which means the error, and equal to  $SSE = \sum (Y_i - \hat{Y}_i)^2$ . Finally, SST indicates to total sum of square and equal to  $SST = \sum (Y_i - \bar{Y})^2$ . It is the correct total variations in the target class. Fig. 3 describes visually the relationships between the data, baseline model, total variability, explained variability and unexplained variability. A comparison between the squared correlation coefficient and the default Minimum R-Square of 0.005 is calculated and the feature is rejected if its value is

less than the cut-off criterion. The feature is selected if its R-square is greater than than the cut-off criterion. The sequential process of feature selection starts by choosing the feature that explains the large amount of changes in the target class. The stepwise process terminates when no remaining input feature can meet the Stop R-Square criterion.

### B. Ensemble Based Feature Selection

Tree algorithms (decision trees, random forest, and gradient boosting) are powerful predictive models. They are the most widely used supervised learning due to their stability, high accuracy and ease of interpretation, considered to be one of the best and mostly used supervised learning methods. Decision tree is a type of supervised learning algorithm that is mostly used in classification problems. As shown in Fig. 4, a decision tree separates the data into segments, and a target value is assigned to each identical segment. A greedy, top-down recursive separating method is used. It employs exhaustive search at each phase by attempting all compositions of features and partition values to gain the maximum decrease in impurity. Subsequently, feature selection can be identified in tree building process. The process of selecting a specific feature based on its relative importance to split in impurity reduction can consider as a kind of feature selection. In our

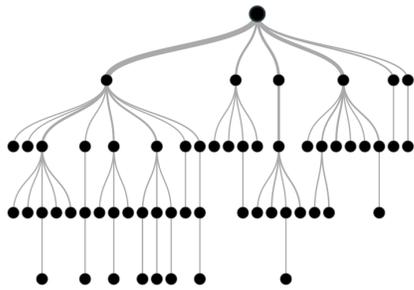


Figure 4: Decision Tree Diagram

model, the measures for feature importance or selection is based on the following metrics: count, surrogate count, residual sum square (RSS), and relative importance. The count-based feature importance simply counts the number of times in the tree that a particular feature is used in a split. Similarly, the surrogate count calculates the number of times that a variable is used in a surrogate splitting rule.

Feature importance measure is calculated for a single decision tree as:

$$VI(x_i, T) = \sum_{t \in T} \Delta I(x_i, t) \quad (9)$$

Where,  $\Delta I(x_i, t) = I(t) - p_L I(t_L) - p_R I(t_R)$  is the reduction in impurity on feature  $x_i$  in tree T at a node t during split.  $p_L$  is the percentage of left observations by  $x_i$  and  $p_R$  for right. Gini index is calculated for classification of node t as:

$$Gini(t) = \sum_{i \neq j} p_i^t p_j^t \quad (10)$$

Where,  $p_i^t$  is the percentage of observations in t with class target equal i ( $y=i$ ) and  $i, j$  run through target class. The Entropy=  $-\sum_i p_i^t \log(p_i^t)$  is similar to Gini index which evaluates impurity at a node t and its value is zero when node has observations from one class. When node has observations from mixture of classes, then entropy value is maximum.

Gradient boosting is a boosting approach that divides the dataset several times using random sampling to create outputs that form a weighted average of the re-sampled data set. Tree boosting generates a series of decision trees which together form a single predictive model. A tree in the series is fit to the residual of the prediction from the earlier trees in the series. The residual is defined in terms of the derivative of a loss function. For the stochastic tree ensemble (Gradient Boosting) of M trees, the generalized importance measure is calculated over the trees:

$$M(x_i) = \frac{1}{M} \sum_{j=1}^M VI(x_i, T_j) \quad (11)$$

In Gradient Boosting, separate models  $f_k(x)$  are built to classify every k classes.

$$F_k(x) = \sum_{j=1}^M T_{kj}(x) \quad (12)$$

The general ((11)) is calculated as:

$$M(x_i, k) = \frac{1}{M} \sum_{j=1}^M VI(x_i, T_{kj}) \quad (13)$$

The total importance of  $x_i$  can be calculated with all classes as:

$$M(x_i) = \frac{1}{K} \sum_{k=1}^K M(x_i, k) \quad (14)$$

The proposed methodology for feature selection using gradient boosting is described in Algorithm 1 as:

## VII. SUPPORT VECTOR MACHINE

Lung Cancer detection and recognition is formulated by binary classification problem. Each patient is classified as normal or abnormal. The goal is labeling a patient to detect the cancer and do the required steps. Many supervised learning methods are learned as computer aided system. In this section, we describe Support Vector Machines (SVMs) as an activity classifier we used in most of the experimental work presented in this field. SVMs are seen as relatively new supervised ML methodology developed by Cortes & Vapnik [14], which were first applied as an alternative to multi-layer neural networks.

```

1 begin
2   Compute variable importance for a single decision
   tree as:  $VI(x_i, T) = \sum_{t \in T} \Delta I(x_i, t)$ 
3   Compute the reduction in impurity on feature  $x_i$  in
   tree T as:  $\Delta I(x_i, t) = I(t) - p_L I(t_L) - p_R I(t_R)$ 
4   For M trees, the generalized importance measure is:
    $M(x_i) = \frac{1}{M} \sum_{j=1}^M VI(x_i, T_j)$ 
5   For every class, the generalized importance measure
   is:  $M(x_i, k) = \frac{1}{M} \sum_{j=1}^M VI(x_i, T_{kj})$ 
6   The total importance of  $x_i$  can be calculated with
   all classes as:  $M(x_i) = \frac{1}{K} \sum_{k=1}^K M(x_i, k)$ 
7 end

```

**Algorithm 1:** PROPOSED GRADIENT BOOSTING FEATURE SELECTION ALGORITHM

To obtain the optimum decision boundary, SVM attempts to maximize the minimal distance from the decision boundary to the labeled data. Once this decision boundary is decided, a given unseen activity can be checked on which side of the decision boundary it lies (Fig. 5).

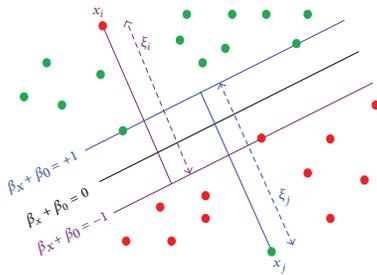


Figure 5: Support Vector Machine

Formally, let  $\mathbf{S} = \{\{x_i, y_i\}_{i=1}^n \mid x_i \in \mathbb{R}^d, y_i \in \{-1, +1\}\}$  be the training samples (i.e., feature vectors of patients), and  $y_i \in \{-1, +1\}$  be the class label of  $x_i$ , thus two parallel separating hyperplanes can be formed such that:

$$y_i = \begin{cases} +1, & w^T x_i + b \geq 1 \\ -1, & w^T x_i + b \leq -1 \end{cases} \quad (15)$$

Where, T denotes the transpose operator, w is a perpendicular vector to the two hyperplanes and b is the bias, as shown in Fig. 5. Thus, the separating decision boundary (i.e. the optimal hyperplane) that maximizes the margin between the two classes is created by solving the following constrained optimization problem:

$$\begin{aligned} & \text{Minimize :} && \frac{1}{2} \|w\|^2 \\ & \text{subject to} && y_i(w^T x_i + b) \geq 1 \quad \forall_i \end{aligned} \quad (16)$$

By Lagrange duality, after some lengthy but straightforward calculations, the dual problem of the primal problem in (16) is given as:

$$\begin{aligned} & \text{Maximize :} && \mathbf{W}(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \\ & \text{subject to} && \alpha_i \geq 0, \quad \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned} \quad (17)$$

Where,  $\alpha_i \geq 0$  are the lagrangian multipliers. Since (17) describes a Quadratic Programming (QP) problem, and a global maximum always exists for  $\alpha_i, \omega$  can be deduced as:

$$\omega = \sum_{i=1}^n \alpha_i y_i x_i \quad (18)$$

## VIII. SIMULATION RESULTS

The experiments are applied on kaggle DSB dataset. In this dataset, a thousand low-dose CT images from high-risk patients in DICOM format is given. The DSB database consists of 1397 CT scans and 248580 slices. Each scan contains a series with multiple axial slices of the chest cavity. Each scan has a variable number of 2D slices (Fig. 6), which can vary based on the machine taking the scan and patient. The DICOM files have a header that contains the necessary information about the patient ID, as well as scan parameters such as the slice thickness. It is publicly available in the Kaggle<sup>1</sup>. DICOM is the defacto file standard in medical imaging. This pixel size/coarseness of the scan differs from scan to scan (e.g. the distance between slices may differ), which can hurt performance of our model.

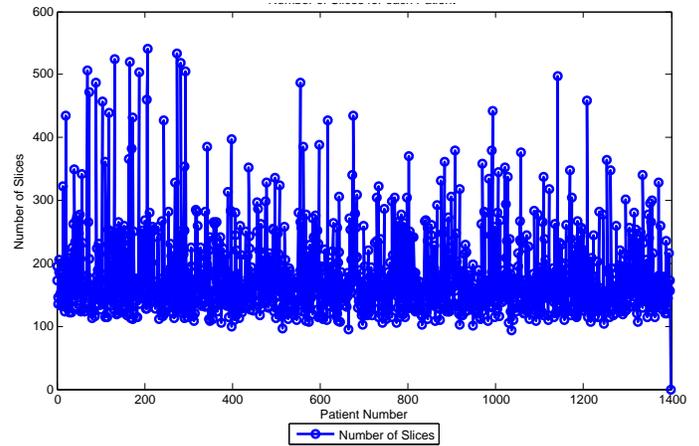


Figure 6: Number of slices per patient in data science bowl dataset.

The experiments are implemented on computer and its properties are described as follows: CPU i7, 2.6 GHz, 16 RAM, Matlab 2016b, R-Studio, and Python. Initially speaking, The nodules in Kaggle DSB dataset are detected and segmented using the watershed algorithm. The diameters of the nodules range from 3 to 30 mm. Each slice has  $512 \times 512$  pixels and 4096 gray level values in Hounsfield Unit (HU), which is a measure of radiodensity.

After segmentation process, binarization process is done. In the screening setting, the annual low-dose CT study is one of the most difficult decisions whether CT or another investigation is needed. The nodule is very complex to be guided using current clinical guidelines due to its size and appearance. Moreover, the most important features of lung cancer are the size, number of nodules, location of the nodule,

<sup>1</sup><https://www.kaggle.com/c/data-science-bowl-2017/data>

and type. The shape features are extracted which describe the shape of extracted nodules. Also, six moments are extracted for each slice of CT lung cancer that has a nodule. The concatenated features are selected based on forward step-wise least squared regression and gradient boosting to select the most important/relevant features and avoid the irrelevant, redundant features, and over-fitting problems.

Gradient Boosting builds a sequential decision tree to form a predictive model. Each iteration, the residuals of classification are updated using loss function derivatives from previous decision trees. The number of iterations in the boosting series is 50 iterations with 60% train proportion. In feature selection algorithm using R-square step-wise regression, the minimum R-square is 0.005 which is the cut-off threshold of a feature to be selected for R-square model selection and other features are irrelevant or redundant.

To evaluate the effectiveness of discrimination subset of features, we apply SVM with a linear kernel using 30% split for testing. Kaggle DSB dataset is divided into 60% for training, 10% for validation and 30% for testing. The accuracy of our proposed model is shown in Table I. As shown from Table I, the best accuracy is 88.07% with gradient boosting, watershed segmentation, and combination of moments and shape features.

Table I: Accuracy Results of our Algorithms on DSB Dataset

Method	Accuracy
Watershed+Moments+SVM	82.34%
Watershed+Shape+SVM	81.62%
Watershed+(Moments+Shape)+SVM	85.68%
Watershed+(Shape+Moments)+R2+SVM	87.11%
Watershed+(Shape+Moments)+Gradient Boosting+SVM	88.07%

The recognition results are shown by confusion matrix achieved on the DSB dataset with gradient boosting feature selection as shown in Table II. As shown from the Table II, Accuracy of model is 88.06%, Mis-classification rate is 11.93%, False positive rate is 11.29%, and False Negative is 13.76%. Almost all patients are classified correctly. Additionally, there is an enhancement on accuracy due to feature selection and efficient feature extraction.

Table II: Confusion Matrix of Watershed, Moments, Shape Feature, and Gradient Boosting Feature Selection using 30% Testing with SVM

Actual	Predicted	
	Abnormal	Normal
Abnormal	94	15
Normal	35	275

## IX. CONCLUSION

Lung cancer recognition and detection based on gradient boosting and regression feature selection and watershed segmentation is presented. Due to the high dimensional data in

medical images with hundreds of CT slices, feature selection is an important step to remove the irrelevant or redundant features. Also, the accuracy of models may be degraded with large number of features if there are not enough training observations to learn all parameters in model activities. In this paper, the features of CT scan for patients are extracted from simple and advanced discriminating method called shape and moments features, then feature selection methods are applied. The gradient boosting and regression models achieved the best accuracy compared to original features and state of the art. Also, ensemble-based or regression-based feature selection methods using random forest, gradient boosting, and R2 reduced the size of features which contribute to avoid over-fitting problems. In the future, we plan to investigate the problem of high dimensional data with different features, different datasets and different approaches like deep learning. Three dimensional convolution neural network (3D CNN) can improve the accuracy of model but it has more powerful machines to run with GPU.

## REFERENCES

- [1] J. Schnabel and B. Glocker, "Deep learning for early detection of lung cancer in patients at risk," 2015.
- [2] W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classification approach," *Entropy*, vol. 15, no. 2, pp. 507–523, 2013.
- [3] S. S. Hamada R. H. Al-Absi, Brahim Belhaouari Samir, "A computer aided diagnosis system for lung cancer based on statistical and machine learning techniques," *JOURNAL OF COMPUTERS*, vol. 9, no. 2, pp. 425–431, 2014.
- [4] B. B. Samir, K. B. Shaban, and S. Sulaiman, "Computer aided diagnosis system based on machine learning techniques for lung cancer," *International Conference on Computer & Information Science*, vol. 9, no. 2, pp. 295–300, 2012.
- [5] M. Gomathi and P. Thangaraj, "A computer aided diagnosis system for lung cancer detection using support vector machine," *American Journal of Applied Sciences*, vol. 7, no. 12, pp. 1532–1538, 2010.
- [6] Ada and R. Kaur, "Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT scan Images," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 3, pp. 187–190, 2013.
- [7] S. L. Kumar, M. Swathy, S. Sathish, J. Sivaraman, and M. Rajasekar, "Identification of lung cancer cell using watershed segmentation on ct images," *Indian Journal of Science and Technology*, vol. 9, no. 1, 2016.
- [8] K. Dimililer, B. Ugur, and Y. K. Ever, "Tumor detection on ct lung images using image enhancement," *The Online Journal of Science and Technology www.tojsat.*, vol. 7, no. 1, pp. 133–138, January 2017.
- [9] N. S. Lingayat and M. R. Tarambale, "A Computer Based Feature Extraction of Lung Nodule in Chest X-Ray Image," *International Journal of Bioscience, Biochemistry and Bioinformatics*, vol. 3, no. 6, pp. 624–629, 2013.
- [10] D. Sharma and G. Jindal, "Computer Aided Diagnosis System for Detection of Lung Cancer in CT Scan Images," *International Journal of Computer and Electrical Engineering*, vol. 3, no. 5, pp. 714–718, 2011.
- [11] A. Hashemi, A. H. Pilevar, and R. Rafah, "Mass Detection in Lung CT Images Using Region Growing Segmentation and Decision Making Based on Fuzzy Inference System and Artificial Neural Network," *International Journal of Image, Graphics and Signal Processing*, vol. 5, no. 6, pp. 16–24, 2013.
- [12] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, February 1962.
- [13] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognition*, vol. 26, no. 1, pp. 167 – 174, 1993.
- [14] C. Cortes and V. Vapnik, "Support-vector networks," *Journal of Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

# Short Survey on Static Hand Gesture Recognition

Huu-Hung Huynh

University of Science and Technology  
The University of Danang, Vietnam

Duc-Hoang Vo

University of Science and Technology  
The University of Danang, Vietnam

**Abstract**—This paper presents a survey of methods which have been recently proposed for recognizing static hand gestures. These approaches are first summarized and then are assessed based on a common dataset. Because mentioned methods employ different types of input, the survey focuses on stages of feature extraction and classification. Other former steps, such as pre-processing and hand segmentation, are slightly modified. In experiments, this work does not only consider the recognition accuracy but also suggests suitable scenarios for each method according to its advantages and limitations.

**Keywords**—Hand gesture; rank-order correlation matrix, Gabor filter; block; centroid distance; Fourier transform

## I. INTRODUCTION

Hand gesture recognition is one of the important problems in vision-related fields such as human-machine interaction, communication, and robotic. There are two gesture types including static and dynamic ones. The application of each gesture type depends on the system objective and gesture definition. Static hand gestures are usually identified based on the hand appearance, e.g. contour and shape, while gestures of the other type are mostly recognized according to the change of hand pose and motion trajectory. In this paper, we focus on some recent methods that recognize static hand gestures. The survey considers four approaches proposed in [1]–[4]. These methods are selected because they used different inputs, described gestures by various features, and employed typical classifiers. Each of such approaches can thus be extended to be appropriate with a wide variation of practical gesture collections.

The work [1] aimed to identify static gestures representing the Vietnamese alphabet, in which a character may be represented by either a single gesture or a combination of two hands. The researchers decomposed this task into a problem of static gesture recognition and a combination based on the alphabetic rules. In this survey, we consider the former task. The input image of the method [1] is a depth map captured by a Kinect 1 at 30 fps and resolution of  $640 \times 480$ . In [2], the authors introduced a system identifying Arabic alphabet and number sign language that helps the communication between hard-of-hearing people. A color camera was employed for data acquisition. The gestures are distinguished according to the hand silhouette and hidden Markov model (HMM) technique. The study [3] also captured input images via a RGB color camera. Differently from [2], the feature extraction was performed in frequency domain, and the stage of classification employed Bayesian techniques. The last considered work [4] combined both spatial and frequency domains for describing the hand characteristics before identifying each input gesture by the

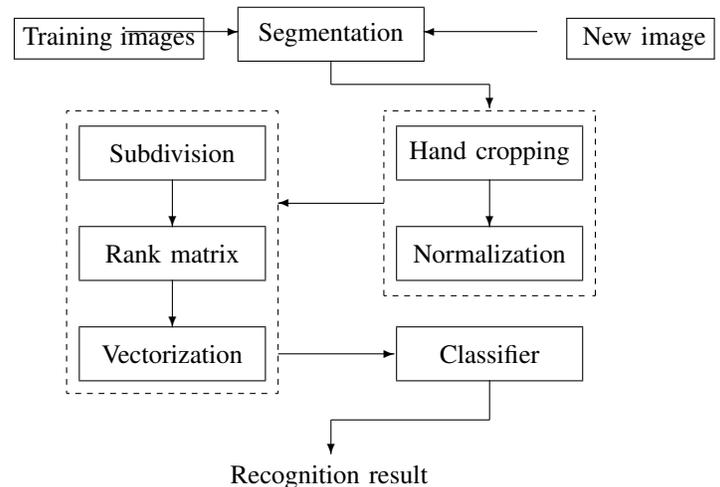


Fig. 1. System overview of the study [1]. The input of the system is a depth map captured by a Kinect 1.

simple nearest neighbor method. Details of these methods are presented in the next section.

## II. LITERATURE REVIEW

Since mentioned studies in this survey employed state-of-the-art classification techniques, the features describing the hand are emphasized. However, the overall processing of each method is also presented.

### A. Rank-order Correlation Matrix

The flowchart of the approach [1] is shown in Fig. 1. Since the input is a depth map which directly captures the scene, the hand segmentation is necessary. The authors introduced two methods for isolating the hand from the background and other body parts according to the fact that hand is the object which is nearest to the camera. The difference between these two methods is the definition of the depth range of interest. The first one sets this range at a fixed distance from the camera, thus a user has to remember it and stand in a suitable region when performing the gestures. The second method is more flexible since it defines the range based on the nearest part. This is also the general segmentation for the object of interest in a depth map in some other studies. The segmentation results a depth map of only the hand. A step of normalization is then applied to synchronize the image size of segmented hands. Instead of using a regular resizing method, the researchers pad

black regions to some borders in order to keep the dimensional ratio of the hand while the image has a square shape. These padded depth maps are finally resized into a fixed size (e.g.  $100 \times 100$  pixels).

In order to extract characteristics describing the considering hand, the normalized depth image is divided into square blocks. The hand of interest is thus represented as a square matrix, in which each element is estimated from a block at the corresponding position. Two matrices are proposed, i.e. there are two values that are extracted from each block. They are statistical descriptions including the mean

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

and standard deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (2)$$

of  $n$  pixel's depths  $x_i$  inside the considering block. In summary, a normalized depth hand gives two matrices, in which the first one contains  $d^2$  values of  $\mu$  and the other one has  $d^2$  elements of  $\sigma$  where  $d$  is the number of blocks corresponding to each image dimension. The values of these two matrices are not directly employed in the task of classification. Instead, each one is converted into a rank matrix, i.e. each element is replaced by its order after sorting all values. The order is in the range  $[0, d - 1]$ .

Such rank matrices are then transformed into vectors according to a specific rule, in which two adjacent vector's elements correspond to two adjacent blocks in the hand image. Each vector is finally converted into correlation vector that describes the relation between two continuous elements. The combination of two correlation vectors is used as the feature of the considering hand. This feature is called Rank-order correlation matrix (ROCM) though the final representation is a vector. The stage of classification is performed by the support vector machine with the one-vs-one strategy [5].

### B. Sorted Block-Histogram

Differently from the work [1], a normal color camera was employed in the study [2]. The overall of this work is presented in Fig. 2.

Since the input is a color image, the authors first detect body parts by applying a skin detection [6]. The image is converted into YCbCr color space in order to separate color channels from the intensity one. To reduce the computational cost, only the chrominance channel (Cr) is then considered. The binary mask of skin-filter result  $S_{i,j}$  is formed by applying a checking on pixel's values of the  $Cr_{i,j}$  image as

$$S_{i,j} = \begin{cases} 0, & 10 < Cr_{i,j} < 45 \\ 1, & \text{otherwise} \end{cases} \quad (3)$$

The noise removal is then performed to reduce possible noises in the obtained binary image. In details, the mask is decomposed into blocks of  $5 \times 5$  pixels. Each block is fully filled by white or black pixels according to the ratio of white pixels inside the block. In normal situations, the filtered body parts usually consist of the face and hand. Therefore, the possible face is detected and discarded using a face detector.

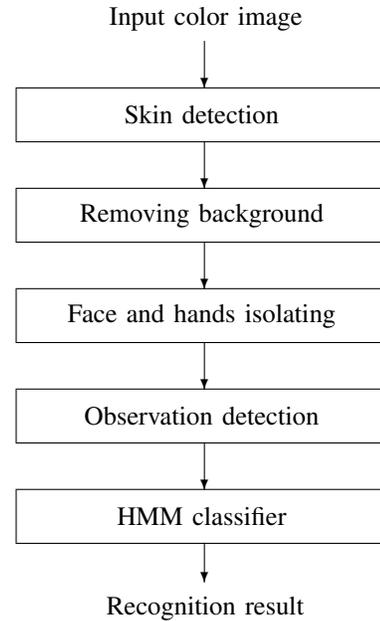


Fig. 2. Main stages of the work [2]. The principal tasks are hand segmentation, feature extraction, and classification.

In the next stage, the segmented hand is processed to provide a feature vector. Similarly to [1], the hand silhouette is also divided into blocks. However, instead of employing statistical descriptions, each block is represented by a simple value which corresponds to the number of white pixels inside it. Every block position is assigned a label, and the feature is formed as a sequence of such labels, in which the label's order is determined based on the sorted pixel-count values.

In order to perform the classification, HMM with discrete observations is employed. The number of observations is the same with the number of blocks (and labels). Many HMMs are built corresponding to the number of gestures. Given the feature vector representing a unknown hand gesture, the returned class is determined as the HMM providing the highest likelihood.

### C. Gabor Filtering

The pipeline of the study [3] is shown in Fig. 3. Similarly to the study described in Section II-B, a skin filter is also applied to detect the hand region. Instead of YCbCr, the L.A.B color space is used for representing the input image. The mask corresponding to skin regions is formed by thresholding automatically the channel B. The noise removal is then performed on the resulted binary image using morphological operations such as erosion and dilation, in which the structure element has the size of  $5 \times 5$  pixels.

In the stage of feature extracting, a collection Gabor filters is employed to emphasize hand characteristics in different orientations. A Gabor filter can be considered as a Gaussian kernel function that is modulated by a sinusoidal wave. It provides two components representing orthogonal directions that include the real and imaginary parts which are respectively

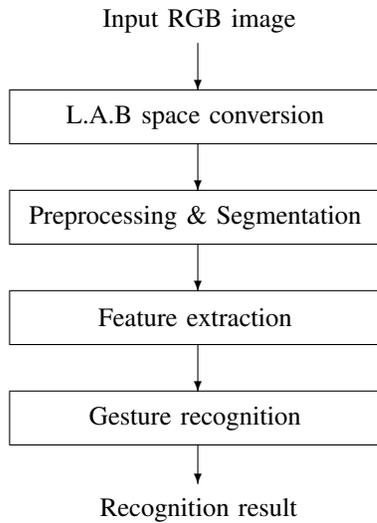


Fig. 3. Pipeline of the work [3]. The stage of feature extraction is performed in frequency domain.

calculated as

$$g_r(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (4)$$

$$g_i(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (5)$$

in which

$$\begin{cases} x' = x \cos(\theta) + y \sin(\theta) \\ y' = y \cos(\theta) - x \sin(\theta) \end{cases} \quad (6)$$

where  $\lambda$  is the sinusoidal factor of the wavelength,  $\theta$  is the orientation characterizing the Gabor filter,  $\psi$  is the phase offset,  $\sigma$  is the standard deviation, and  $\gamma$  is the spatial aspect ratio. These parameters are assigned by the collections of values presented in Table I.

TABLE I. PARAMETER VALUES DEFINING GABOR FILTERS

Parameter	Notation	Values
Orientation	$\theta$	$\{0, \frac{\pi}{8}, \frac{2\pi}{8}, \frac{3\pi}{8}, \frac{4\pi}{8}, \frac{5\pi}{8}, \frac{6\pi}{8}, \frac{7\pi}{8}\}$
Wavelength	$\lambda$	$\{4, 2^{1/4}, 8, 2^{1/8}, 16\}$
Phase	$\psi$	$\{0, \frac{\pi}{2}\}$
Aspect ratio	$\gamma$	$\gamma = \lambda$

After applying the defined Gabor filters on the segmented hand, a classifier is used to determine the corresponding class of the input gesture. The Bayesian techniques was employed in this work.

#### D. Centroid Distance Signature

In the final considered approach [4], the spatial and frequency domains are continuously employed for feature extraction. Similarly to the study [1], the input is a depth map captured by a Kinect 1 at 30 fps with resolution of  $640 \times 480$  pixels. The overall of the system is presented in Fig. 4. In order to segment the hand from input depth image, the researchers first employ a functionality in the Kinect SDK to isolate the human body from the background. Possible hands are then determined and separated by applying a threshold on depth

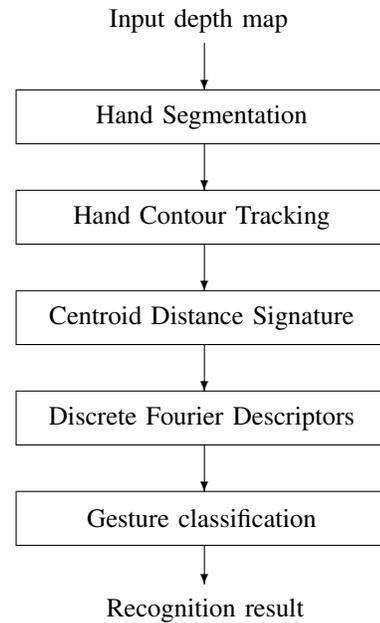


Fig. 4. The gesture recognition approach proposed in [4].

layers representing the whole body. The  $k$ -mean clustering technique is then performed to check the number of hands. The number of clusters is thus assigned to 2. In the case that there is only one hand, two clustered regions are combined. This case is detected by comparing the distance between two centroids with a predefined threshold. The next stages are to recognize each hand gesture.

Instead of considering the entire pixels representing the hand, only the contour is focused. The 8-connectivity algorithm [7] is applied to perform this task. The centroid of these determined pixels is also estimated. In the next step, the feature vector is formed as a sequence of distances between contour pixels and the centroid. This vector is finally converted into frequency domain using the Fast Fourier Transform (FFT) [8]. However, the FFT requires that the input must have a power-of-two elements. Besides, the lengths of sequences of contour points corresponding to different hands also need to be normalized for the classification stage. Therefore the researchers converted feature vectors to 128-element ones based on equal angle sampling. The recognition is finally performed based on the nearest neighbor with Euclidean distance.

These described studies are summarized in Table II.

### III. EXPERIMENTS

The survey is performed on four different methods for static hand gesture recognition. The selection of dataset is thus important since it must be appropriate for all of these approaches. We employ the dataset provided in the study [1]. This is the collection of segmented depth hands corresponding to 23 static gestures. The dataset consists of 4637 images that were captured by the depth sensor of a Microsoft Kinect 1 and segmented based on depth thresholds. A visualization of these gestures is presented in Fig. 6. This dataset is suitable for methods processing on depth images as well as hand silhouettes. Therefore the stage of hand segmentation in considering

TABLE II. SUMMARY OF FOUR STUDIES CONSIDERED IN THIS PAPER

Method	Input	Hand segmentation	Feature extraction	Classification
[1]	Depth map	Depth thresholding	Block dividing & Statistical values	Support vector machine
[2]	Color image	Skin filtering in YCbCr	Block dividing & Sorted indices	Hidden Markov model
[3]	Color image	Skin filtering in L.A.B	Gabor filtering	Bayesian technique
[4]	Depth map	Body detection & Depth thresholding	Centroid distance signature & FFT	Nearest neighbor

TABLE III. GESTURE RECOGNITION ACCURACIES CORRESPONDING TO FOUR CONSIDERING METHODS

Approach	Vo et al. [1]	Abdo et al. [2]	Ashfaq et al. [3]	Gani et al. [4]
Recognition accuracy	94.22%	87.71%	86.98%	98.32%

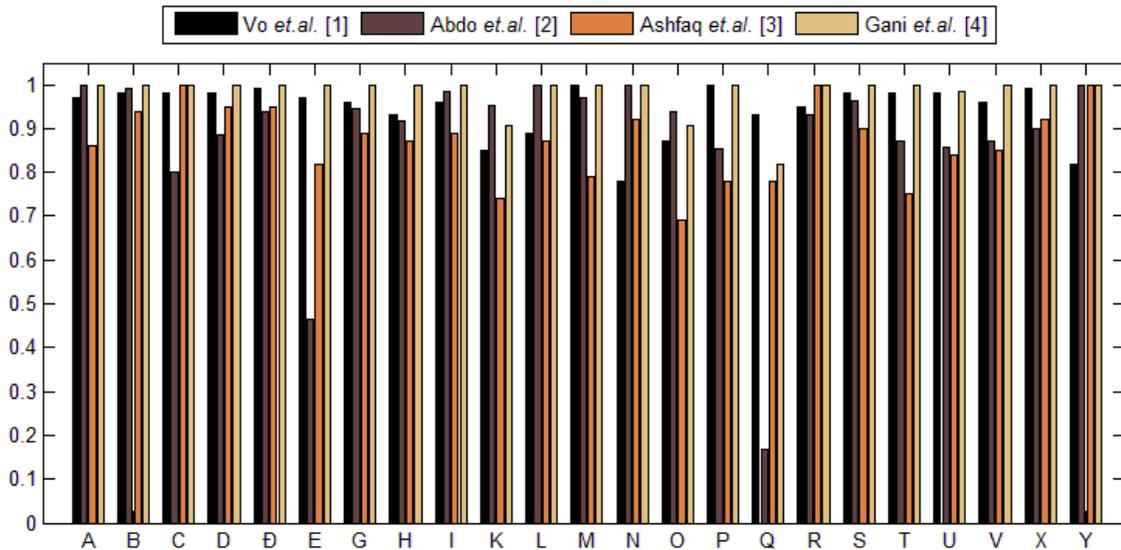


Fig. 5. Accuracies when recognizing each gesture of the benchmark dataset.



Fig. 6. Benchmark dataset including 23 static hand gestures which represent 23 letters in Vietnamese alphabet.

studies is not necessary to perform in these experiments. In other words, this paper accesses only the combination of feature extraction and classification. The overall accuracies of considered methods in recognizing gestures corresponding to 23 classes are shown in Table III. The details are also presented in Fig. 5.

According to Table III, the method proposed in [4] has best ability in classifying 23 gestures in the benchmark dataset. Frequency domain is thus an appropriate choice for improving current works on hand gesture recognition. This method also provides absolute accuracies when identifying many gestures (see Fig. 5). Therefore the processing flow in [4] should be

applied at the beginning of a study on gesture recognition. The approach gives the second highest accuracy is the study [1]. The partial recognition accuracies corresponding to 23 classes are mostly similar. Therefore the statistical information extracted from regions and/or blocks of depth maps should also be focused when dealing with the problem of gesture recognition. However, the work [1] requires a depth camera for data acquisition while the study [4] does not. The accuracies of two other methods are slightly different. However, the experimental values corresponding to several gestures in [2] are quite low (e.g. E and Q). It means that the simple feature proposed in [2] may be significantly affected by the gesture selection. In other words, the ability of gesture recognition may be very different when applying the method on various collections of gestures. A validation for each particular dataset is thus necessary. The remaining study [3] employs a collection of Gabor filters for feature extraction. This method thus requires a large number of operations. In our experiments, the computational cost of [3] is very high compared with the others. Therefore this approach is only suitable for (1) systems with high strength of computation and/or (2) low-resolution input images.

#### IV. CONCLUSION

This paper presents a survey on some recent methods which perform the static hand gesture recognition. These studies employ different input data types, extracted features, and classification techniques. Details of each processing stage

corresponding to each method are described. The experiments are performed on a common dataset in order to provide a comparison on recognition ability of these approaches. According to the obtained results, advantages and limitations of them are also given.

In future works, a combination of these approaches and some recent ones will be focused to improve the ability of recognizing more complicated static gestures. In detail, we attempt to estimate gesture scores based on a weighted sum of likelihoods computed from each individual method. Such classification fashion is expected to take advantage of each approach. In addition, each considered method can be extended, by employing more features and/or combining with different weak classifiers, to be appropriate with a specific practical gesture collection.

#### ACKNOWLEDGMENT

The authors would like to thank Trong-Nguyen Nguyen, DIRO, University of Montreal, for his technical support.

#### REFERENCES

- [1] D. H. Vo, T. N. Nguyen, H. H. Huynh, and J. Meunier, "Recognizing vietnamese sign language based on rank matrix and alphabetic rules," in *2015 International Conference on Advanced Technologies for Communications (ATC)*, Oct 2015, pp. 279–284.
- [2] M. Z. Abdo, A. M. Hamdy, S. A. E.-R. Salem, and E. M. Saad, "Arabic alphabet and numbers sign language recognition," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 11, pp. 209–214, 2015.
- [3] T. Ashfaq and K. Khurshid, "Classification of hand gestures using gabor filter with bayesian and naïve bayes classifier," *International Journal Of Advanced Computer Science And Applications*, vol. 7, no. 3, pp. 276–279, 2016.
- [4] E. Gani and A. Kika, "Albanian sign language (albsl) number recognition from both hand's gestures acquired by kinect sensors," *International Journal Of Advanced Computer Science And Applications*, vol. 7, no. 7, pp. 216–220, 2016.
- [5] J. Milgram, M. Cheriet, and R. Sabourin, "'One Against One' or 'One Against All': Which One is Better for Handwriting Recognition with SVMs?" in *Tenth International Workshop on Frontiers in Handwriting Recognition*, G. Lorette, Ed., Université de Rennes 1. La Baule (France): Suvisoft, Oct. 2006.
- [6] Y. t. Pai, S. j. Ruan, M. c. Shie, and Y. c. Liu, "A simple and accurate color face detection algorithm in complex background," in *2006 IEEE International Conference on Multimedia and Expo*, July 2006, pp. 1545–1548.
- [7] T. Pavlidis, *Algorithms for graphics and image processing*. Springer Science & Business Media, 2012.
- [8] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.

# Mobility based Net Ordering for Simultaneous Escape Routing

Kashif Sattar

University Institute of Information Technology  
Arid Agriculture University  
Rawalpindi, Pakistan

Anjum Naveed

School of Electrical Engineering and Computer Science (SEECS)  
National University of Sciences and Technology (NUST)  
Islamabad, Pakistan

Aleksandar Ignjatovic

School of Computer Science and Engineering (CSE)  
University of New South Wales (UNSW)  
Sydney, Australia

Muhammad Zeeshan

School of Electrical Engineering and Computer Science (SEECS)  
National University of Sciences and Technology (NUST)  
Islamabad, Pakistan

**Abstract**—With the advancement in electronics technology, number of pins under the ball grid array (BGA) are increasing on reduced size components. In small size components, a challenging task is to solve the escape routing problem where BGA pins escape towards the component boundary. It is often desirable to perform ordered simultaneous escape routing (SER) to facilitate area routing and produce elegant Printed Circuit Board (PCB) design. Some heuristic techniques help in finding the PCB routing solution for SER but for larger problems these are time consuming and produce sub-optimal results. This work propose solution which divides the problem into two parts. First, a novel net ordering algorithm for SER using network theoretic approach and then linear optimization model for single component ordered escape routing has been proposed. The model routes maximum possible nets between two components of the PCB by considering the design rules based on the given net ordering. Comparative analysis shows that the proposed net ordering algorithm and optimization model performs better than the existing routing algorithms for SER in terms of number of nets routed. Also the running time using proposed algorithm reduces to  $O(2^{N_E/2}) + O(2^{N_E/2})$  for ordered escape routing of both components. This time is much lesser than  $O(2^{N_E})$  due to exponential reduction.

**Keywords**—Net ordering; optimization model; ordered escape routing; PCB routing; simultaneous escape routing

## I. INTRODUCTION

With the advancement in technology, the demand for small size electronic components with larger number of pins increases [1]. These components can be Integrated Circuits (ICs) or BGA type with hundreds of connectivity pins [2]–[4] in the form of grid. Usage of such components necessitates sophisticated design and advanced technologies for producing small sized electronic circuit boards. Multiple BGAs need to be connected on the PCB, which require connectivity of pins by routing of nets. This kind of routing has two parts, one is escape routing which is from pin to the boundary of the component and the other is area routing which is between the components as shown in Fig. 1. Presence of BGA components on smaller sized PCBs significantly increases the

routing complexity.

Routing in PCBs is the problem of connecting the pins of different components on the PCB to make it operational. The circuit wire lying on the PCB, between the pins pair (that needs to be connected) is called a net. Routing of nets becomes more complex, if there are more number of nets and needs ordered escape routing. Ideally, the routing algorithm shall connect all nets as a planar graph (nets don't cross each other) and the nets shall be entirely contained within the area of the PCB. Given the high density of components on smaller sizes PCBs, and hundreds of nets to route. A single layer is not enough to route all nets in a planar fashion particularly if ordered escaping is required. Therefore, multiple routing layers are used to ensure planarity as shown in Fig. 1.

PCB routing for given set of connectivity pins is similar problem as constructing of planar graph for a given set of nodes, which is known to be NP-Hard [5]. BGA components have small size balls/pins under the surface, in the form of a grid. This high density grid reduces the net escaping space and make the routing more complex. Therefore, PCB routing is divided into two parts: 1) escape routing where routing is to be done from pin to the component boundary; and 2) area routing where the net between one component to other component is established as shown in Fig. 1. To facilitate the area routing, escape routing is required in some particular order. Therefore ordered escape routing is a major challenge in PCB routing, which becomes even more challenging for simultaneous escape routing (SER) [6]–[12]. In SER pin escaping for multiple components is done simultaneously in a specific order to reduce the complexity of area routing as shown in Fig. 2. To reduce the complexity of SER, the problem has been divided into two parts i.e. ordered escape routing of two components separately. In this case, net order has necessarily required.

Focus of this research is to find the net order for simultaneous escape routing. In this paper, an algorithm to find the net order has been proposed. This reduces the problem into half and reduces the complexity of routing by many times. In this way the area routing becomes very simple and it is just a

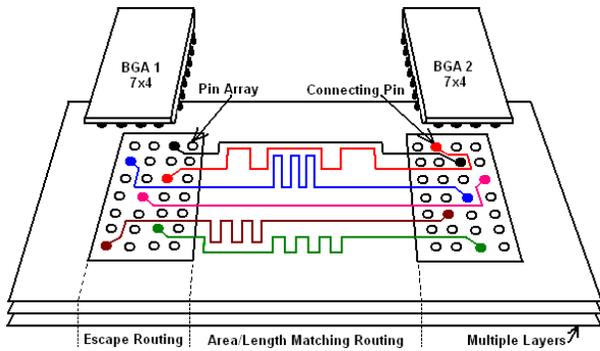


Fig. 1. PCB routing.

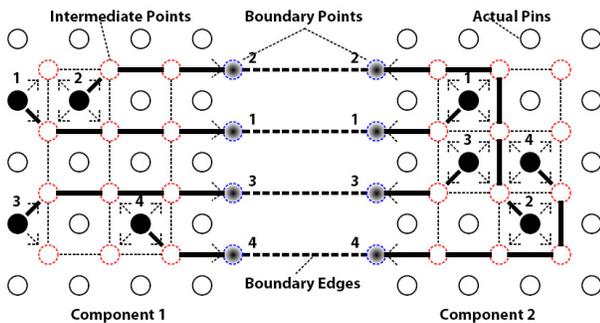


Fig. 2. Simultaneous escape routing.

matter of connectivity of straight wires without any crossing to maintain planarity. The proposed algorithm finds the net order and do ordered escape routing for each component according to that net order. For ordered escape routing optimization model has been used, which helps in detailed routing from a connectivity pin to an escape boundary point. Comparative evaluation shows that the proposed optimization model can connect more number of nets in all tested examples when compared with the well known commercial tools. The proposed optimization model is a contribution towards enhancing the capabilities of PCB designers and can be integrated into commercial software for use. Following are the contributions of this work:

- In the first step an algorithm to find the optimal net order has been proposed. This algorithm helps by converting the complex problem of simultaneous escape routing into simple and single component ordered escape problem. Ultimately this exponentially reduces the complexity of the problem.
- In the second step ILP optimization model for ordered escape routing has been modified, which helps in finding the ordered net escape from a single component.

Collectively these two contributions help in solving SER problem in a much better way.

The rest of the paper is organized as follows: In Section 2, some basic techniques used for planar routing and advancements in this area with a review of the latest literature has been discussed. Section 3 presents proposed algorithm to find a net order. In Section 4, optimization model for ordered escape routing under the constraints of planar routing and escape capacity has been proposed. In Section 5, algorithm/model val-

idation and comparative evaluation of the proposed algorithm has been performed. Section VI concludes this work.

## II. RELATED WORK

Routing of nets on PCB, is divided into two main categories, one is area and the other is escape routing [7]. The objective of escape routing is to establish the planar routes by escaping the nets from pins to the component boundary in an ordered or unordered way, based on constraints on available capacity between adjacent pins. With the escape routes established, the problem of PCB routing is reduced to conventional PCB routing problem of area routing [13], [14].

Escape routing is of three types [15]: 1) Unordered escape routing where there is no particular sequence of pin escaping towards component boundary from a single component [15]–[21]; 2) ordered escape routing where there is a particular sequence (order) of pin escaping towards component boundary from a single component [22] (one of the example of this type of routing is Bus escape which is also called as maximum disjoint subset problem [23]); and 3) simultaneous escape routing where pins from multiple components escaped in the same order. This problem relatively more complex as compare to other escaping problems. To find the solution of SER is time consuming process, particularly for larger problems. This time can be reduced if the escape net order is known in advance. Escape net order is a sequence of nets to escape from the boundary of a component. This prior net ordering information converts the problem of SER into ordered escape routing of a single component. Focus of this work is to find the optimal net order. The following subsections discuss the related literature in detail.

### A. Simultaneous Escape Routing

The initial research work proposed for SER consists of two parts. Unordered escaping of pins proposed in the first component whereas for the second component, ordered escape routing proposed to match the pins of first component to get SER solution. However, this approach might results in an incomplete routing. Ozdal et al. [7] carried out the first well known SER research. Graph based approach was used to find maximal escape nets. This approach gives better results only if the pins of both components are closely aligned to each other. However, complex problems of SER cannot be solved using fixed escape patterns used in this approach. Ozdal et al. [8] also proposed randomized algorithm for large scale problems in their extended work. However, proposed monotonic escape of nets in one direction cannot produce optimal results and also it limits its applications. For routing pattern generation, Ozdal et al. [9] also used congestion driven router, however for large scale SER problems the graph of escape patterns becomes complicated.

Simultaneous Pin Escape [12] is a flow model based approach. Authors show that for a particular problem if construction of planar graph is possible then this approach guarantees planar routing. The existence of a solution can be determined by the relation between maximum flow and number of escape pins of the BGA component [12], [16]. This approach must provide a planar solution if maximum flow is greater or equal to the escape pins; however, for higher inter

pin capacities 'no net crossing' cannot be ensured and exact value of maximal flow is not known. A bus oriented escape routing with consecutive constraints has been proposed for SER on multiple layers [24]. The focus of this research is to reduce the CPU time, however if the pins are randomly placed on each component then its very difficult to form buses and also unable to get optimal routing solution due to escaping of nets in the form of buses.

### B. Net Ordering Algorithms

B-Escape is a routing algorithm proposed by Luo et al. [10]. This algorithm uses dynamic net ordering approach and performs well when compared with Allegro PCB router [10]; however it consumes significant time to find a suitable net order and in reordering of nets again and again. Rout-ability Driven Net Order proposed by Yan et al. [25] is a multi-step approach. An important step is identification of a net ordering, which is derived on the basis of planar bipartite graph theory. A subsequent step uses global and detailed routing technique under the constraints of planar routing and escape capacity. This approach improves the computation time by 54.1% as compared to Kong et al. [12]. The drawback is that the second step is totally dependent on the output of the first step i.e. net ordering. It does not explore all possible paths and may lead to sub-optimal results in complex problem instances.

Escaped Boundary Pins Routing for High-Speed Boards proposed by Chin et al. [26] is an algorithm for finding static net ordering by dynamic pin sequence (DPS). Another work proposed by Kumtong et al. [27] is based on set sequence (SS) technique. While these algorithms can better utilize the routing space and adapt to wire length and shape requirements, their focus is on boundary pin routing and are not suitable for BGA type of PCB components which need escaping from inside the components; their use can also be more challenging, particularly in case of SER.

Most of the research on SER is based on basic constraints of capacity and construction of planar graph. Also most of them are using heuristic approaches to get the routing solution. This research considers optimization theoretic approach. Instead of solving the problem in parts to get sub-optimal solution, this optimization approach is expected to produce optimal results. This approach not only considers capacity and planarity but also considers constraint of power-signal integrity, net route length and net escape order altogether, to solve the SER problem at the cost of higher computational time. However, with high speed computing units, clustering of computational resources to perform complex computational tasks and using net ordering techniques, the drawback can easily be overcome. In this research an algorithm for optimal net ordering has been proposed which reduces the computational time and outperforms the existing algorithms.

### III. ALGORITHM FOR NET ORDERING

This section, first define the terms used in the algorithm and then discuss the tile model and net ordering algorithm in detail. It is also assumed that there are only two components and only one side escape is possible from the side of component, which is facing towards the other component.

$Rows_{Ci}$ : Number of rows of pins for component  $Ci$ .

$Orthogonal_{Cap}$ : This capacity refers to the number of nets that can pass through between two neighboring pins vertically or horizontally.

$Diagonal_{Cap}$ : It refers to the number of nets that can pass through between two neighboring diagonal pins.

$Cap_{Ci}$ : Escape routing capacity for component  $Ci$ .

$SER_{Cap}$ : Simultaneous Escape Routing capacity.

$BoundaryPins_{Ci}$ : Number of pins on Escape boundary line which needs connectivity for the component  $Ci$ .

$URDF\_Mobility_{Pi}$ : Number of nets that can pass from Up, Rear, Down and Front side of pin  $Pi$ .

### A. Net Ordering Algorithm

This proposed algorithm performs four basic functions in sequence as discussed in Algorithm 1, which are  $SER\_Capacity()$ ,  $Initial\_NetOrder()$ ,  $URDF\_Mobility\_Values()$  and  $Final\_NetOrder()$  respectively. This algorithm produces the optimal net order for SER, if it exists. Then escape routing for each component performed separately according to that net order with the help of an optimization model.

---

#### Algorithm 1 Net Ordering Algorithm

---

```
1: procedure NET_ORDERING()
2:    $Routable \leftarrow SER\_Capacity()$ 
3:   if ! $Routable$  then
4:     print 'Complete Routing not possible' and exit
5:    $Pins_{C1}[n] \leftarrow$  'n' Connectivity pins of component 1
6:    $Pins_{C2}[n] \leftarrow$  'n' Connectivity pins of component 2
7:    $I[n] \leftarrow Initial\_NetOrder(Pins_{C1}[n], Pins_{C2}[n])$ 
8:    $M1[n][4] \leftarrow URDF\_Mobility\_Values(Pins_{C1}[n])$ 
9:    $M2[n][4] \leftarrow URDF\_Mobility\_Values(Pins_{C2}[n])$ 
10:   $F[n] \leftarrow Final\_NetOrder(I[n], M1[n][4], M2[n][4])$ 
11:  if  $F[n] = Null$  then
12:    print 'Complete Routing not possible' and exit
13:  else
14:    return  $F[n]$ 
```

---

1)  $SER$  Capacity Function:  $SER\_Capacity()$  function has been discussed in Algorithm 2. First calculate the escape capacity of each component. Here  $BoundaryPins_{Ci}$  has been added in  $Cap_{Ci}$ , because they can escape directly from the component boundary without effecting the other nets or orthogonal capacity. Obviously the  $SER$  capacity must be the minimum escape capacity of both components. Then it is also verified whether the routing with the required number of nets ' $n$ ' is possible or not.

Consider an example instance of two BGA components of size 6x4, having 5 connectivity pins as shown in Fig. 3. Also consider left component as  $C1$  and right component as  $C2$ .  $Rows_{C1}$  and  $Rows_{C2}$  are 6, whereas  $Orthogonal_{Cap}$  is 1. Since there is no boundary connectivity pin, therefore  $Cap_{C1}$  is 5 and  $Cap_{C2}$  is also 5. According to Algorithm 2,  $SER_{Cap}$  is also 5. There are 5 pins which need connectivity (escape routing) and these are less than or equal to the  $SER_{Cap}$ , therefore the Algorithm 2 returns true and move on the next step of the main Algorithm 1.

**Algorithm 2** SER Capacity Algorithm

```

1: procedure SER_CAPACITY()
2:    $Cap_{C1} = (Rows_{C1} - 1) \times Orthogonal_{Cap} +$ 
    $BoundaryPins_{C1}$ 
3:    $Cap_{C2} = (Rows_{C2} - 1) \times Orthogonal_{Cap} +$ 
    $BoundaryPins_{C2}$ 
4:    $SER_{Cap} = \min(Cap_{C1}, Cap_{C2})$ 
5:   if  $SER_{Cap} \geq n$  then
6:     return true
7:   else
8:     return false

```

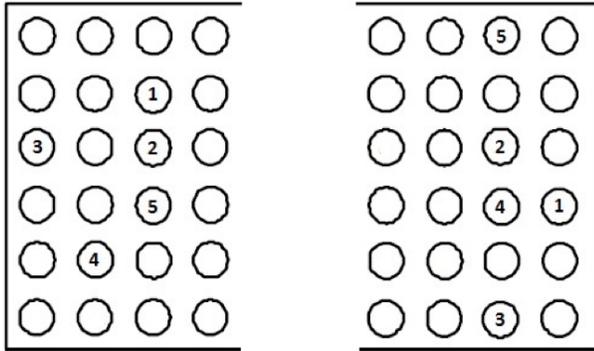


Fig. 3. BGA component of size 6x4.

2) *Initial Net Order*: To find the initial net order according to the position of pins in the grid of the component, *Initial\_NetOrder()* function in Algorithm 3 has been discussed. Also some rules for rearrangement of the initial net order has been defined.

In the first step, convert each row in a vertex for both components as shown in Fig. 4. Then for the rows having multiple connectivity pins apply the rules defined in Step 2 to Step 13 of Algorithm 3. In this example there is Pin-3 on non-escape boundary of left component and also there is one pin above that row whereas there are two escape boundary points above that row. Therefore select Pin-3, first in order. Similarly select Pin-1, first in order in the right component. In the third step connect all pins to their corresponding pins with edges, as shown in Fig. 4.

3) *URDF Mobility Values*: In the third basic function of Algorithm 1, find the *URDF* mobility values of a pin according to the position of pin in the component as shown in Fig. 5.

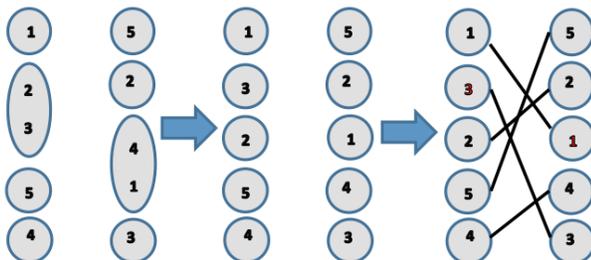


Fig. 4. Three step process to find initial net order.

**Algorithm 3** Initial NetOrder Algorithm

```

1: procedure INITIAL_NETORDER( $Pins_{C1}[n], Pins_{C2}[n]$ )
2:   Read row wise all pins from top to bottom and put
   them in column form, separately for each component.
3:   if there are multiple pins in a single row then
4:     if there are multiple pins in the first row then
5:       order must be from the escape side to non-
       escape side.
6:     else if there are multiple pins in a last row then
7:       order must be from the non-escape side to the
       escape side.
8:     else if there is a pin on the non-escape boundary
       column and the number of escape boundary points are
       greater than the number of pins above that row then
9:       select this pin first and name it as prioritized
       rear boundary pin.
10:    else if In cases, where the same multiple pins are
       in a single row in both components then
11:      select from the escape side to the non-escape
       side.
12:    else
13:      For all other rows order should be according
       to the row number of the corresponding pin in the other
       component.
14:    Now there are two separate columns having number of
       rows equals to the number of pins. This is called initial
       net ordering that facilitates in finding the final net
       ordering. Connect pin in the left column with its
       corresponding pin in the right column with edges as
       shown in the last step of Fig. 4.

```

**Algorithm 4** URDF Mobility Values Algorithm

```

1: procedure URDF_MOBILITY_VALUES( $Pins_{C_i}[n]$ )
2:   Find the Up, Rear, Down and Front (URDF) mobility
   values of a pin by identifying its position in the grid.
3:    $U = i, R = j, D = k$  and  $F = l$  implies that there are
   ( $i/Orthogonal_{Cap}$ ) number of rows on the up side of the
   connectivity pin, ( $j/Orthogonal_{Cap}$ ) number of columns
   on the rear side, ( $k/Orthogonal_{Cap}$ ) number of rows on
   the down side and ( $l/Orthogonal_{Cap}$ ) number of columns
   on the front side (escape side) of the pin respectively, in
   a component. These values help in finding the number of
   nets that can possibly be passed through, from that side
   of the pin.

```

According to the example of Fig. 3, Pin-1 is in second row and third column of the left component, which implies that there are one row on the upside of the pin, four rows on the down side, two columns on the rear side and one column on the front side. Since  $Orthogonal_{Cap} = 1$  is being considered, therefore the *URDF* values are  $U = (\text{No. of rows on Up side} * Orthogonal_{Cap}) = 1, R = 2, D = 4$  and  $F = 1$  respectively as shown in Fig. 5. Furthermore,  $U = 1$  implies that only one net can pass through the upside of Pin-1 and similarly for the other values.

4) *Final Net Order*: The last function of Algorithm 1 finds the final net order as shown in Fig. 6, by eliminating a crossing of edges among two columns. To eliminate the cross, there are four possibilities. For example, consider the first

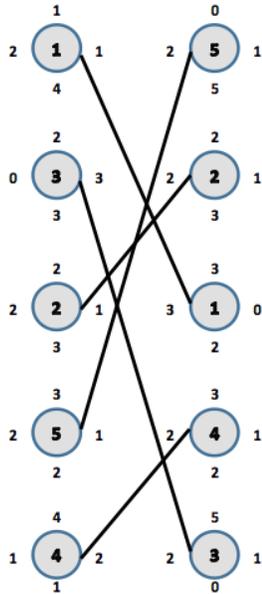


Fig. 5. Finding the URDF mobility values.

crossing between Pin-1 and Pin-5, as shown in Fig. 5. The first possibility is to move the Pin-5 in left column to the top of the Pin-1 in the left column. Second possibility is to move Pin-1 in the left column to the bottom of the Pin-5 in the same column. Similarly, the third possibility is to move the Pin-1 in the right column to the top of Pin-5 of the right column and finally the fourth possibility is to move Pin-5 of the right column to the bottom of Pin-1 in the same column. Try all the four possibilities in some order, till selection of one, which is allowed by *URDF* values as discussed in Algorithm 5. Failure of all the possibilities shows that routing of both nets is not possible in this scenario.

In final net ordering process, initially considers first pin of each column, which is Pin-1 and Pin-5 as shown in Fig. 6. Since the row of Pin-5 in its component is higher (smaller row number), therefore select Pin-5 and try to move its corresponding pin from left component towards upside. But on the way there is Pin-3, which has rear value,  $R = 0$ . This implies that Pin-5 cannot route from rear side of Pin-3. Therefore, first move Pin-3 towards downside of Pin-5. Since the rear and down values of Pin-5 are greater than zero, hence route Pin-3 from the rear side of Pin-5. After the routing of Pin-3, route Pin-5 towards upside at the level of selected Pin-5 of right component as shown in Fig. 6. In this way eliminate all the crossings and get the final nets order if exists. The proposed algorithm not only converts the *SER* problem into much simplified ordered escape routing problem by providing the net order, but also provides the global routing of each net. In the next section optimization model for detailed routing is being used.

#### IV. OPTIMIZATION MODEL FOR ORDERED ESCAPE

Optimization models are very helpful to solve real life problems [28], [29]. These models provide optimal solution by considering all constraints. This section presents the modified Integer linear program based optimization model proposed in

#### Algorithm 5 Final NetOrder Algorithm

```

1: procedure FINAL_NETORDER( $I[n]$ ,  $M_1[n][4]$ ,  $M_2[n][4]$ )
2:   Consider a current variable, which initially points to
   the first pin of each column. In each iteration, the current
   variable moves to the next row. Let Pin-A be from left
   column and Pin-B be from the right column.
3:   Select the pin according to the following rules:
4:   if Escape boundary connectivity pins are already fixed
   with their escape boundary points then
5:     select it first and arrange its counter pin.
6:   else if A and B both have the same position in the
   corresponding ordering of pins of the two components and
   no other pin is being blocked by routing this net then
7:     these pins are already in order. Only need to update
   the mobility values of the left over unordered pins. If, by
   this choice, any pin is being blocked then reselect that pin
   first.
8:   else if A or B is a prioritized rear boundary pin then
9:     select that pin first. If, by this choice, any pin is
   being blocked then reselect that pin first.
10:  else if If A or B is not in the component's first row
   and the next row pin (w.r.t current row) of both columns
   are the same i.e C then
11:    select C
12:  else
13:    select the pin whose row number is lesser.
14:  Route the selected pin (e.g A) by moving its counter
   Pin-A' (or both in case C is selected) towards upside in a
   column to reach at current row level by these rules:
15:  if First pin to move upward then
16:    move the counter Pin-A' (both in case of C) upward
   from the rear side of all the pins, till reaching at current
   row level to eliminate the edge crossing.
17:  else if on the way any pin have  $R=0$  or  $U=0$  then
18:    first move that pin below the counter pin A'
   (counter pin must have  $R>0$  and  $D>0$ ). If there is more
   than 1 pin which faces this situation then order of moving
   down pins is from escape side to non-escape side.
19:  if routing is not possible by selecting Pin-A and still
   Pin-B remains then
20:    select Pin-B and goto Step 15
21:  if routing is not possible by selecting any pin A or B
   then
22:    return false
23:  else
24:    update the mobility values. The value in each
   direction will be either the distance between the pin and
   the net passing through in that direction or connected pin
   whichever is lesser. Move the current variable to the next
   row and goto Step 3.
25:  if eliminate all crossings by reordering pins then
26:    this is our final net order and the function returns
   true.
27:  else
28:    return false

```

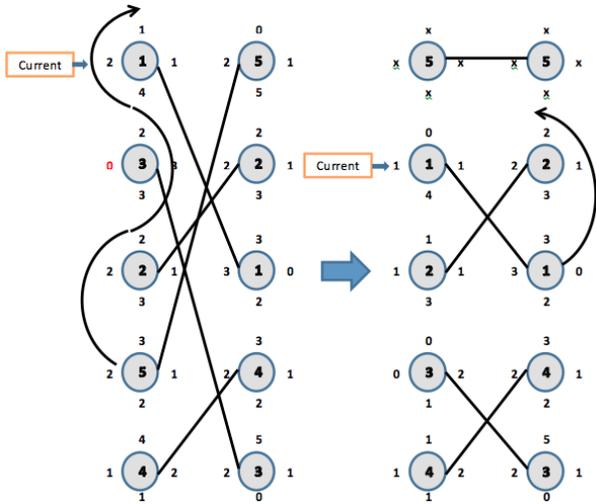


Fig. 6. Finding the final net order.

our earlier work [30]. System constraints from the model has been removed because these are useless once net order has been finalized. The graph  $G(V, E)$  serves as input to this model, where  $V$  is a set of vertices. This set consists of boundary points, intermediate points and connecting pins. All the possible edges are in set  $E$  and the nets use these edges for escape towards the boundary of the component.

Based on the given constraints, the proposed model returns another graph  $G'(V', E')$ , which shows the maximum number of possible escaped nets, with  $G'(V', E') \leq G(V, E)$ , which also implies that  $V' \leq V$  and  $E' \leq E$ . Linear programming model has been formulated that can perform escape routing of nets from a component according to the given net order identified in the previous section. These are the sets used as input in the proposed model:

$N$ = Set of required nets.

$I$  = All intermediate and boundary points.

$V$  = All intermediate points, boundary points and connecting pins.

$BE$ = All edges at the component's boundary.

$E$  = All possible edges of a BGA. The proposed model contains decision variable  $X_{(n_a, e_{ij})}$ , which is a Boolean variable for all nets  $n_a \in N$  and for all edges  $e_{ij} \in E$ . If the solver assign true or 1 to the decision variable then that edge  $e_{ij}$  for a particular net  $n_a$  for escape routing is being used, otherwise the solver assigns false or zero value to the decision variable.

$$X_{(n_a, e_{ij})} = \begin{cases} 1 & \text{if } e_{ij} \text{ edge is being used for net } n_a \\ 0 & \text{if } e_{ij} \text{ is not being used for } n_a \end{cases}$$

By using this optimization model, routing is constrained by the constraints of connectivity, planar graph, net order and system constraint. Each constraint is important in terms of achieving error free routing designs [30]. The following set of equations (1)-(6), completely represents the optimization model.

$$\max. \sum_{n_a \in N} \sum_{e_{aj} \in E} X_{(n_a, e_{aj})} \quad (1)$$

subject to:

$$\sum_{e_{ij} \in E \& i=a} X_{(n_a, e_{ij})} \leq 1 \quad \forall n_a \in N \quad (2)$$

$$\sum_{e_{ij} \in E} X_{(n_a, e_{ij})} = \sum_{e_{jk} \in E} X_{(n_a, e_{jk})} \quad \forall n_a \in N, \forall j \in I \quad (3)$$

$$X_{(n_a, e_{ij})} + X_{(n_a, e_{ji})} \leq 1 \quad \forall n_a \in N, \forall e_{ij}, e_{ji} \in E \quad (4)$$

$$\sum_{n_a \in N} \sum_{e_{ij} \in E} X_{(n_a, e_{ij})} \leq 1 \quad \forall j \in I \quad (5)$$

$$\sum_{e_{ij} \in BE} X_{(n_a, e_{ij})} \cdot j \leq \sum_{e_{kl} \in BE} X_{(n_{a+1}, e_{kl})} \cdot l \quad \forall n_a \in N \quad (6)$$

## V. EVALUATION

Proposed net ordering algorithm has been implemented in C++ and for detailed routing, optimization model has been implemented in AMPL language [31]. Initially, the proposed algorithm and optimization model has been validated in Sub-section V-A and V-B, respectively. Subsequently the performance is being evaluated in Sub-section V-C.

### A. Algorithm Validation

This section verify that the algorithm proposed in Section III, gives the optimal net order if  $SEER$  has a valid routing solution. Furthermore if a particular scenario has no complete routable solution then it return false. Algorithm validated for this purpose on BGA component as shown in Fig. 3. Both components  $C1$  and  $C2$  have 6 rows from  $A$  to  $F$  and 4 columns from 1 to 4. There are 5 connectivity pins in each component that needs escape routing to connect with each other. To simplify the explanation, in this example assume  $Orthogonal_{Cap} = 1$  and  $Diagonal_{Cap} = 2$ . Twelve (12) scenarios has been considered with randomly selected 5 escape pins in each component. In the first scenario, first pin is at location B4 in  $C1$  and location E2 in  $C2$ . Similarly see the location of other pins in Fig. 7. The net order obtained by our proposed algorithm is 1,3,2,4 and 5. Then run the same scenario in commercially available Proteus auto router by Labcenter Electronics Ltd. The result obtained by Proteus is exactly in the same order as obtained by our algorithm as shown in Fig. 7.

Complete validation results can be seen in Table I. First column is the scenario number, second and third columns are the pin positions in each component respectively from Pin-1 to 5. Fourth column is the net order obtained by the C++ program for the proposed algorithm and the fifth column is the net order obtained by the Proteus router. Results validate the algorithm for all scenarios. In some cases the algorithm returns false, which implies that complete routing for all the nets is not possible as shown in Scenario 7 of Table I. Running the same scenario in Proteus, resulted in the un-routable solution as shown in Fig. 8. Routing of Pin-1 from B3 of component 1 to F2 of component 2 is not possible, which validates the results.

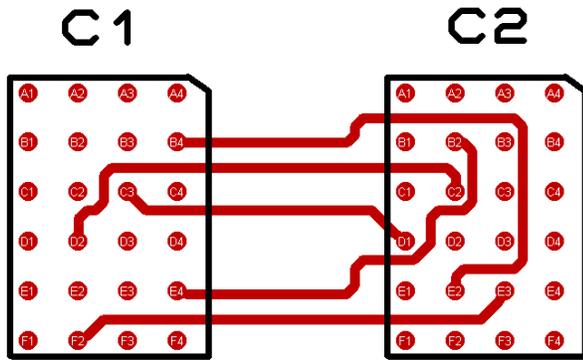


Fig. 7. SER in BGA 24\_1.5.

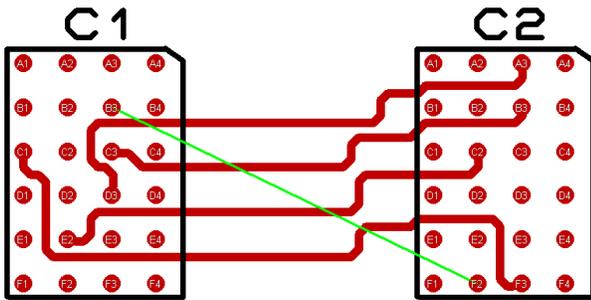


Fig. 8. Proteus incomplete routing.

**B. Model Validation**

Consider Scenario 1 of Table I for the validation of model. The net order obtained by the proposed algorithm is 1, 3, 2, 4, and 5. This net order converts the SER problem into two single components ordered escape routing problems. This example clearly shows that the optimization model efficiently performs ordered escape routing for PCB components based on BGA, as shown in Fig. 9. There are maximum 5 boundary points in the BGA component and model chooses to route all five nets by escaping in a given order. It is evident that this solution ensures planarity by not selecting any vertex or edge twice. The routing of nets is exactly the same as obtained by Proteus auto router as shown in Fig. 7. This example performs comprehensive validation of proposed optimization model and ensures all constraints of connectivity, net order and planar graph.

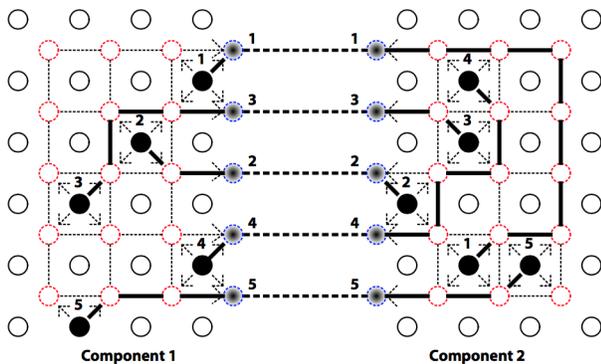


Fig. 9. Ordered escape routing with proposed model.

TABLE I. VALIDATION RESULTS AND COMPARISON WITH PROTEUS

S.No.	C1 Pins	C2 Pins	NetOrder		Accuracy
			Algorithm	Proteus	
1	B4,	E2,	1,	1,	100%
	C3,	D1,	3,	3,	
	D2,	C2,	2,	2,	
	E4,	B2,	4,	4,	
	F2	E3	5	5	
2	B3,	D4,	5,	5,	100%
	B4,	C3,	2,	2,	
	C1,	F3,	1,	1,	
	E2,	D2,	4,	4,	
	D3	A3	3	3	
3	B3,	D4,	5,	5,	100%
	C3,	C3,	1,	1,	
	C1,	F3,	2,	2,	
	E2,	D3,	3,	3,	
	D3	A3	4	4	
4	B3,	D2,	5,	5,	100%
	C3,	C3,	1,	1,	
	C1,	F3,	2,	2,	
	E2,	D3,	4,	4,	
	D3	A3	3	3	
5	C2,	B3,	1,	1,	100%
	C1,	A2,	5,	5,	
	D1,	F2,	2,	2,	
	E2,	E4,	4,	4,	
	B1	D3	3	3	
6	D3,	A3,	1,	1,	100%
	E2,	F2,	5,	5,	
	C3,	C3,	3,	3,	
	C1,	F3,	4,	4,	
	B3	D4	2	2	
7	B3,	D2,	Routing of all nets not possible		100%
	C3,	C3,	Routing of all nets not possible		
	C1,	F3,			
	E2,	D3,			
	D3	A3			
8	B3,	E2,	5,	5,	100%
	C3,	B3,	1,	1,	
	C1,	F3,	4,	4,	
	E2,	C2,	2,	2,	
	D2	A3	3	3	
9	D3,	A3,	5,	5,	100%
	D4,	D3,	3,	3,	
	C1,	C3,	1,	1,	
	D1,	E3,	2,	2,	
	D2	B3	4	4	
10	A3,	E2,	1,	1,	100%
	C3,	D2,	3,	3,	
	D2,	C2,	4,	4,	
	E3,	B2,	2,	2,	
	F2	E3	5	5	
11	D3,	F2,	4,	4,	100%
	C3,	C2,	2,	2,	
	F2,	E2,	5,	5,	
	A3,	B2,	1,	1,	
	C2	D3	3	3	
12	C2,	D4,	Routing of all nets not possible		100%
	F3,	A1,	Routing of all nets not possible		
	B3,	E1,			
	C4,	C2,			
	E1	A4			

**C. Performance Analysis**

The performance comparison in this section is to find the optimal net order by comparing the results of our proposed mobility based net ordering (MBNO) algorithm with the routability driven net ordering (RDNO) algorithm [25]. Consider three different scenarios of the same examples discussed in [25] for the purpose of evaluation. In this example there are 9 rows and 10 columns for both components, having 90 pins in total. There are 14 nets, which need connectivity and escaping of pins for SER. Also single side escape is permitted with  $Orthogonal_{Cap}$  equal to 2 and the  $Diagonal_{Cap}$  equal to 3. Each scenario highlights different aspect of the proposed

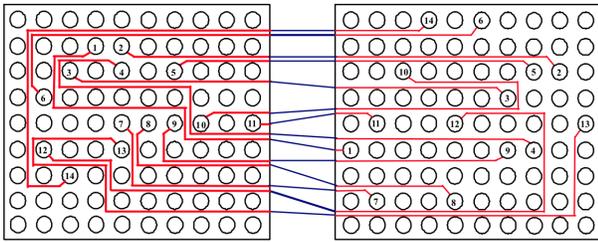


Fig. 10. Routability driven Net ordering [25].

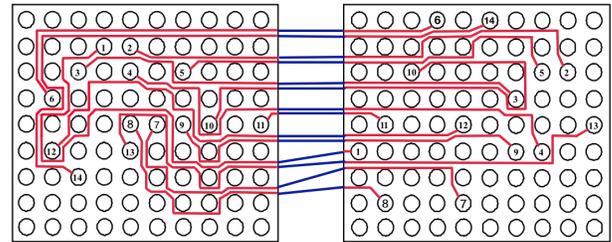


Fig. 13. MBNO routing for swap pin scenario.

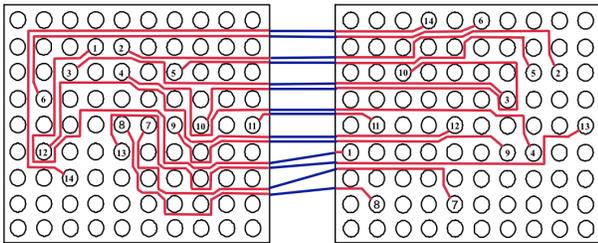


Fig. 11. Proposed Mobility based Net ordering.

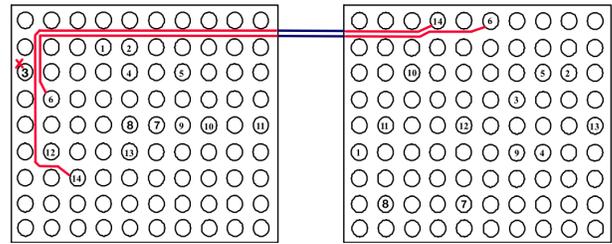


Fig. 14. RDNO routing for back boundary pin scenario.

algorithm as explained in respective sub-section.

1) *BGA\_9x10 Default case Scenario-I*: To validate, results obtained by our proposed MBNO algorithm has been compared with the results obtained by RDNO algorithm. For the default example discussed in [25] both provided the optimal net order. Fig. 10 shows the output of RDNO algorithm and the net order is: 14, 6, 2, 5, 3, 10, 11, 4, 1, 9, 7, 8, 12, 13. Whereas Fig. 11 shows the output of the proposed algorithm and the net order is: 14, 6, 2, 5, 3, 10, 4, 11, 12, 9, 1, 13, 7, 8. Results show that proposed algorithm not only provides optimal net order but also provides more compaction to the escaping nets by utilizing all escape boundary points in a sequence, which can accommodate more nets within the limits of  $SER_{Cap}$ .

2) *BGA\_9x10 Swap pin Scenario-II*: In this scenario swap the pin number 6 and 14 with each other in right component of Fig. 10, whereas rest of the things remain exactly the same. The net order remains the same using RDNO algorithm but after placing net number 14 as a first net, it blocks pin number 6 of the right component as shown in Fig. 12 and ends with incomplete routing. However our proposed MBNO algorithm provides a complete routing as shown in Fig. 13, with the following net order: 6, 14, 2, 5, 3, 10, 4, 11, 12, 9, 1, 13, 7, 8.

3) *BGA\_9x10 Back boundary pin Scenario-III*: This scenario place the pin number 3 at the back boundary of left

component of Fig. 10, whereas rest of the settings remain exactly the same in both components. The net order remains the same using RDNO algorithm but after placing net number 14 and 6 as a first and second net respectively, it blocks pin number 3 of the left component as shown in Fig. 14 and ends with an incomplete routing. However our proposed MBNO algorithm first provides the partial routing to the blocked pin i.e. pin number 3, as shown in Fig. 15. Then the algorithm does routing for the pin number 14 and ultimately provides optimal net order for complete routing: 14, 6, 2, 5, 10, 4, 11, 1, 9, 12, 3, 13, 8, 7.

Detailed routing for the given net order is shown in Fig. 16. Results show that how efficiently proposed algorithm solves the SER problem based on the mobility values for each pin.

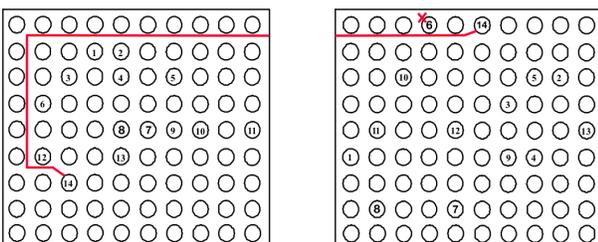


Fig. 12. RDNO routing for swap pin scenario.

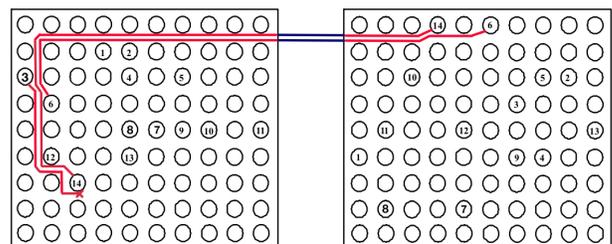


Fig. 15. MBNO partial routing for back boundary pin scenario.

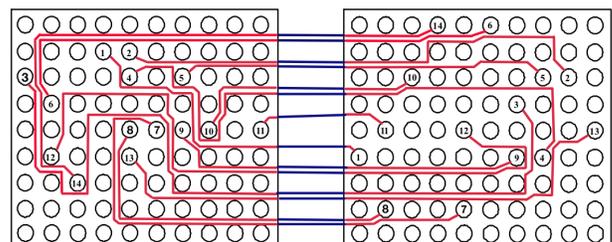


Fig. 16. MBNO complete routing for back boundary pin scenario.

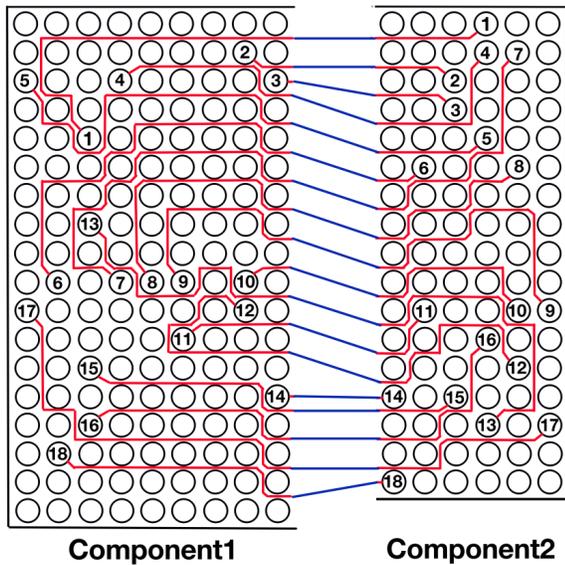


Fig. 17. MBNO results for varying size BGA's scenario.

TABLE II. EXPERIMENT RESULTS AND COMPARISON WITH RDNO ALGORITHM

Scenario No.	C1 Pins	C2 Pins	No. of Nets	Net Routing	
				RDNO	MBNO
I	9x10	9x10	14	14/14	14/14
II	9x10	9x10	14	Incomplete	14/14
III	9x10	9x10	14	Incomplete	14/14
IV	18x9	17x6	18	Incomplete	18/18

4) *Varying Size BGA's Scenario-IV*: The scenarios discuss so far are variations of the same example having two BGA components of the same size. Now consider another example having different size BGA's. Component 1 is of size 18x9 and component 2 is of size 17x6. There are 18 nets that need SER connectivity. Again assume single side escape with  $Orthogonal_{Cap}$  equals to 1 and the  $Diagonal_{Cap}$  is 2. Fig. 17 shows, that the proposed algorithm generates the optimal net order for successful completion of SER for all the nets.

Here, results have been compared by different methods discussed so far. Proposed MBNO algorithm performed 100% routing in all of the four scenarios by generating optimal net order, however RDNO [25] could not route all nets in Scenario-II, Scenario-III and Scenario-IV as shown in Table II. It is important to note that manual routing of even single net left out by an automated method is very difficult to route and is a very time consuming process requiring rearranging of all the routed nets. It is also clear from the results that based on the mobility values of each pin there are better routing results because it avoids any connectivity pin blockage while finding out the next net in the optimal net order.

## VI. CONCLUSION

This research work propose the use of net order, in order to solve the SER problem in a way superior to the previous state of art methods. A mobility based net ordering algorithm has been proposed that finds the optimal net order and converts the problem of SER into two simpler problems of finding an ordered escape routing for each component separately. Once

“connectivity pins” escape from each component, according to the net order obtained by the proposed algorithm, then area routing becomes quite simple by connecting nets of both components in an order preserving way. Net order obtained by proposed algorithm for 12 representative and randomly generated scenarios has been validated and compared it with the net order obtained by Proteus auto router. Proposed ILP optimization model has been updated to get the detail routing results. The detailed routing results are similar with the results obtained by Proteus auto router. The time for SER optimization model is exponential, i.e.,  $O(2^{NE})$ . Where N is the number of nets that needs to be routed and E is the set of all type of edges. This means addition of only one more variable (net or edge), doubles the computation time due to exponential behaviour, i.e.,  $O(2^{NE+1})$ . Similarly by decreasing the edges results in exponential reduction in time. In this case, edges are being reduced to almost half (assuming both components have equal edges). Therefore the new time is  $O(2^{NE/2})+O(2^{NE/2})$  for ordered escape routing of both components. This time is much lesser than  $O(2^{NE})$  due to exponential reduction. Results shows that proposed algorithm finds routable net order for all scenarios whereas previous algorithm fails to route all nets according to their derived net order.

Now a days, there are different types of BGA components based on their pin arrays. These types include, square pin array, triangle pin array, diamond pin array and hexagonal pin array. The proposed models in this research focus only on BGA pins forming square shape grids, in future we can apply this research on other available BGA grid shapes like triangle, diamond and hexagonal to see the increase in capacity for the same size of components. Also in future we are extending the algorithm for higher number of components.

## REFERENCES

- [1] M. M. Ozdal, “Routing algorithms for high-performance vlsi packaging,” Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2005.
- [2] G. Marcantonio, “Ball grid array (bga) integrated circuit packages,” Aug. 18 1998, uS Patent 5,796,170.
- [3] D. Y. Chong, B. Lim, K. J. Rebibis, S. Pan, S. Krishnamoorthi, R. Kapoor, A. Sun, and H. Tan, “Development of a new improved high performance flip chip bga package,” in *Electronic Components and Technology Conference, 2004. Proceedings. 54th*, vol. 2. IEEE, 2004, pp. 1174–1180.
- [4] R. Plieninger, M. Dittes, and K. Pressel, “Modern ic packaging trends and their reliability implications,” *Microelectronics Reliability*, vol. 46, no. 9, pp. 1868–1873, 2006.
- [5] M. Garey and D. Johnson, “Appendix: a list of np-complete problems,” p. 202, 1979.
- [6] H. Xiang, X. Tang, and D. Wong, “An algorithm for simultaneous pin assignment and routing,” in *Proceedings of the 2001 IEEE/ACM international conference on Computer-aided design*. IEEE Press, 2001, pp. 232–238.
- [7] M. M. Ozdal and M. D. Wong, “Simultaneous escape routing and layer assignment for dense pcbs,” in *Proceedings of the 2004 IEEE/ACM International conference on Computer-aided design*. IEEE Computer Society, 2004, pp. 822–829.
- [8] —, “Algorithms for simultaneous escape routing and layer assignment of dense pcbs,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 25, no. 8, pp. 1510–1522, 2006.
- [9] M. M. Ozdal, M. D. Wong, and P. S. Honsinger, “Simultaneous escape-routing algorithms for via minimization of high-speed boards,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 27, no. 1, pp. 84–95, 2008.

- [10] L. Luo, T. Yan, Q. Ma, M. D. Wong, and T. Shibuya, "B-escape: a simultaneous escape routing algorithm based on boundary routing," in *Proceedings of the 19th international symposium on Physical design*. ACM, 2010, pp. 19–25.
- [11] Q. Ma, T. Yan, and M. D. Wong, "A negotiated congestion based router for simultaneous escape routing," in *Quality Electronic Design (ISQED), 2010 11th International Symposium on*. IEEE, 2010, pp. 606–610.
- [12] H. Kong, T. Yan, and M. D. Wong, "Optimal simultaneous pin assignment and escape routing for dense pcbs," in *Design Automation Conference (ASP-DAC), 2010 15th Asia and South Pacific*. IEEE, 2010, pp. 275–280.
- [13] M. Tanaka and Y. Nakagiri, "An approach to pin assignment in printed circuit board design," *ACM SIGDA Newsletter*, vol. 10, no. 2, pp. 21–33, 1980.
- [14] H. N. Brady, "An approach to topological pin assignment," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 3, no. 3, pp. 250–255, 1984.
- [15] T. Yan, Q. Ma, and M. D. Wong, "Advances in pcb routing," *Information and Media Technologies*, vol. 7, no. 2, pp. 535–543, 2012.
- [16] T. Yan and M. D. Wong, "A correct network flow model for escape routing," in *Design Automation Conference, 2009. DAC'09. 46th ACM/IEEE*. IEEE, 2009, pp. 332–335.
- [17] Q. Ma and M. D. Wong, "Np-completeness and an approximation algorithm for rectangle escape problem with application to pcb routing," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 31, no. 9, pp. 1356–1365, 2012.
- [18] Y.-K. Ho, H.-C. Lee, and Y.-W. Chang, "Escape routing for staggered-pin-array pcbs," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 32, no. 9, pp. 1347–1356, 2013.
- [19] K. Wang, S. Dong, H. Wang, Q. Chen, and T. Lin, "Mixed-crossing-avoided escape routing of mixed-pattern signals on staggered-pin-array pcbs," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 33, no. 4, pp. 571–584, 2014.
- [20] J.-T. Yan, "Length-constrained escape routing of differential pairs," *Integration, the VLSI Journal*, vol. 48, pp. 158–169, 2015.
- [21] J. McDaniel, Z. Zimmerman, D. Grissom, and P. Brisk, "Pcb escape routing and layer minimization for digital microfluidic biochips," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 1, pp. 69–82, 2017.
- [22] F. Jiao and S. Dong, "Ordered escape routing for grid pin array based on min-cost multi-commodity flow," in *Design Automation Conference (ASP-DAC), 2016 21st Asia and South Pacific*. IEEE, 2016, pp. 384–389.
- [23] H. Kong, Q. Ma, T. Yan, and M. D. Wong, "An optimal algorithm for finding disjoint rectangles and its application to pcb routing," in *Design Automation Conference (DAC), 2010 47th ACM/IEEE*. IEEE, 2010, pp. 212–217.
- [24] J.-T. Yan, "Layer assignment of escape buses with consecutive constraints in pcb designs," *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 22, no. 3, p. 45, 2017.
- [25] J.-T. Yan, T.-Y. Sung, and Z.-W. Chen, "Simultaneous escape routing based on routability-driven net ordering," in *SOC Conference (SOCC), 2011 IEEE International*. IEEE, 2011, pp. 81–86.
- [26] C.-Y. Chin, C.-Y. Kuan, T.-Y. Tsai, H.-M. Chen, and Y. Kajitani, "Escaped boundary pins routing for high-speed boards," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 32, no. 3, pp. 381–391, 2013.
- [27] W. Kumtong, P. Danklang, and W. Sriborrirux, "Pin set sequence selection guideline routing for printed circuit board routing," in *Knowledge and Smart Technology (KST), 2015 7th International Conference on*. IEEE, 2015, pp. 126–130.
- [28] Q.-D. Ngo, Y. Hady-Said, U. Maulik *et al.*, "Automatic generation of model for building energy management," *International Journal of Advanced Computer Science & Applications*, vol. 1, no. 7, pp. 442–454, 2016.
- [29] E. Ahmed, A. Gani, M. Sookhak, S. H. Ab Hamid, and F. Xia, "Application optimization in mobile cloud computing: Motivation, taxonomies, and open challenges," *Journal of Network and Computer Applications*, vol. 52, pp. 52–68, 2015.
- [30] K. Sattar and A. Naveed, "Ordered escape routing using network flow and optimization model," in *Automation, Robotics and Applications (ICARA), 2015 6th International Conference on*. IEEE, 2015, pp. 563–568.
- [31] R. Fourer, D. Gay, and B. Kernighan, "Ampl: A modeling language for math. programming," *Duxbury Press/Brooks/Cole Publishing*, 2002.

# A Survey on User Interfaces for Interaction with Human and Machines

Mirza Abdur Razzaq  
Department of Computer Science  
Shah Abdul Latif University  
Khairpur, Pakistan

Muhammad Ali Qureshi  
Department of Telecommunication Engineering  
University College of Engineering and Technology  
The Islamia University of Bahawalpur, Pakistan

Kashif Hussain Memon  
Department of Computer Systems Engineering  
University College of Engineering and Technology  
The Islamia University of Bahawalpur, Pakistan

Saleem Ullah  
Department of Computer Science & IT  
Khawaja Fareed University of Engineering and IT  
Rahim Yar Khan, Pakistan

**Abstract**—Interaction with the machines and computers is achieved using user interfaces. Nowadays, with the tremendous growth of technology, the interaction is made more simple and flexible. The study of user interfaces for human-computers and machines interaction is the main focus of this paper. In particular, an extensive overview of different user interfaces available so far is provided. The review covers text-based, graphical-based, and new class of emerging user interfaces to interact with the machines and computers. This work will be helpful for the development of new user interfaces.

**Keywords**—Command line interface (CLI); graphical user interface (GUI); user interface (UI); sixth sense device; natural language interface; brain-computer interface; emerging user interfaces

## I. INTRODUCTION

Human Machine Interaction (HMI) and Human Computer Interaction (HCI) are used for effective control of machines (e.g., automobiles, control panels to operate industrial machines), and computers.<sup>1</sup> Other technologies used in the industry are Operator Interface Terminal (OIT) and Operator Interface Console (OIC). The interaction is achieved using user interfaces to control the input and output. The user interface (UI) is considered as prime ingredient of computer user satisfaction. The UI consists of both hardware and software. The design of an user in UI affects the amount of effort the user must spend to provide input for the system and to interpret the output of the system. Fig. 1 shows a nice example of human-machine interaction using input and output controls. The output interface is used to assist the users by displaying the feedback from the system. HMI is an interface between the human (user) and the machine (or equipment) and it is typically local to one machine or piece of equipment e.g., UI of a mechanical system, a vehicle or an industrial installation (see Fig. 2). Currently, various user interfaces exist to interact with machines and computers [1]. The equipment manufacturing companies are competing with each other by introducing innovative user interfaces to capture large market shares. Seeing the importance of user interfaces, we provide in

this paper, an overview of the existing and future interfaces. We have categorized the user interfaces into text-based, graphical based and emerging user interfaces.

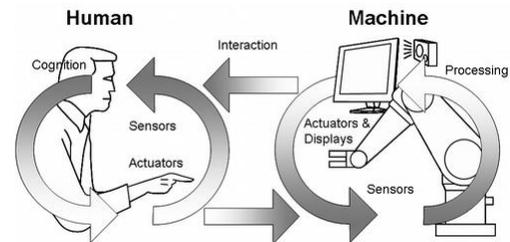


Fig. 1. Example of human machine interaction.



Fig. 2. HMI of sugar industry machine.

The rest of the paper is organized as follows: Sections II, III and IV discuss different categories of user interfaces in detail and finally, the paper is concluded in Section V.

## II. TEXT-BASED INTERFACES

### A. Batch Interfaces

It is the non-interactive interface extensively used in the past during the period 1945 to 1968. In batch interfaces, all jobs or commands are combined in a single batch by the user and executed once instead of executing commands separately.

<sup>1</sup>[https://en.wikipedia.org/wiki/User\\_interface](https://en.wikipedia.org/wiki/User_interface)

All jobs are performed step by step in a given order without any intermediate interaction with the user. The main advantage of this interface is time saving during the interaction period. On the other hand, it is difficult to terminate the execution of running batches until the user completes a job.

### B. Command Line Interface (CLI)

This interface was first time introduced in 1969 and still used by expert programmers. The examples are UNIX, MS-DOS application in Windows operating systems, etc. In CLI, the user enters commands on a prompt line through the terminal to interact with the operating system. The response messages from the system are shown on the display screen. It is very tedious to remember different commands as well as command line switches (see Fig. 3).

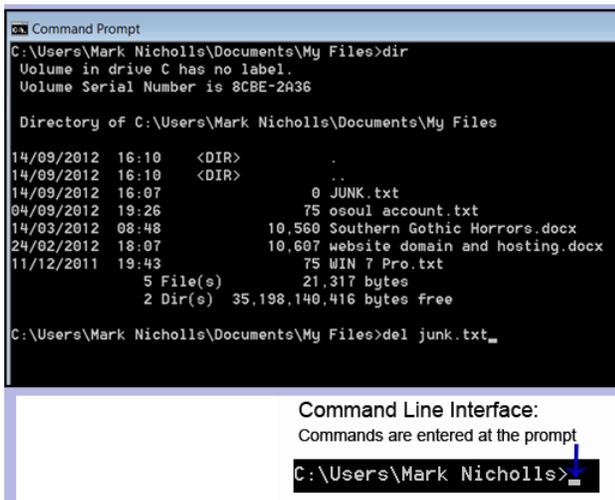


Fig. 3. Example of CLI.

### C. Non-Command User Interface

In this interface, the computer observes the requirement to understand the user without typing of commands. The interface just shows the alphabet options to type instead of typing complete commands (see Fig. 4).



Fig. 4. Example non-command user interface.

### D. Natural Language Interface (NLI)

This type of interface usually used in the search engines. When the user enters a question using verbs, clauses, and phrases, the response is given to the user after searching the web. In the latest research of relational databases, a query in

natural English language is an example of NLI, where the sentence is translated into Structured Query Language (SQL) query using joins, nesting, and aggregation. The SQL query is then entered in relational database management systems and finally, the response is presented to the user. It is convenient for a layman with no knowledge of SQL [2].

## III. GRAPHICAL INTERFACES

### A. Graphical User Interface (GUI)

The first GUI or WIMP (Windows, icons, mouse pointer) was designed by Xerox Corporations Palo Alto Research Centre in the 1970s, but unfortunately, it did not get popularity till 1980. It was popularized for the first time by the Apple company and then extensively used by the Microsoft Windows (see Fig. 5). Using a well-designed GUI, the user can get benefits in terms of fast learning and ease to use. Examples are Microsoft Windows, Apple Macintosh, and Ubuntu operating systems. For UNIX-based systems, both CLI and GUI interfaces are used.

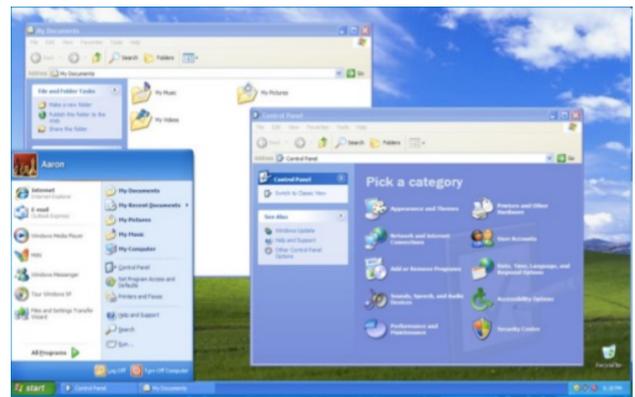


Fig. 5. Example of GUI.

### B. Attentive User Interface (AUI)

In AUI, the notifications or messages are shown to the user and asked to enter the same information to ensure his presence for security purposes before generating related information [3]. The examples are Captcha entry, session expiry, etc. (see Fig. 6).



Fig. 6. Example CAPTCHA: The user is asked to type the printed words before proceeding.

### C. Crossing-Based Interface

In this interface, the cursor or pointer is moved in or out from a graphical object area e.g., hyperlink on mouse hover.

#### D. Intelligent User Interface (IUI)

This interface deals with artificial intelligence having domain knowledge based on reasoning to assist the users [4], [5]. It has the capability “to adapt the user, communicate with the user, and solve problems for the user”, e.g., office assistant or wizard, etc.

#### E. Reflexive User Interfaces

In this interface, the user controls the whole system and can redefine its commands or define new commands by performing actual actions using a rich graphical interface.

#### F. Touch User Interface (TUI)

The TUI is a special kind of GUI using a touch-pad or a touch-screen display as a combined input and output device. In TUI, the display screen is pressure sensitive where the user interacts with the machine by tapping the screen surface using fingers or stylus. On tapping, the system compares the actions stored in the database and executes appropriate commands. Fig. 7 shows an example of TUI. Currently, the TUI is widely used in mobile devices, point of sales, industrial processes, and self-service machines, etc.

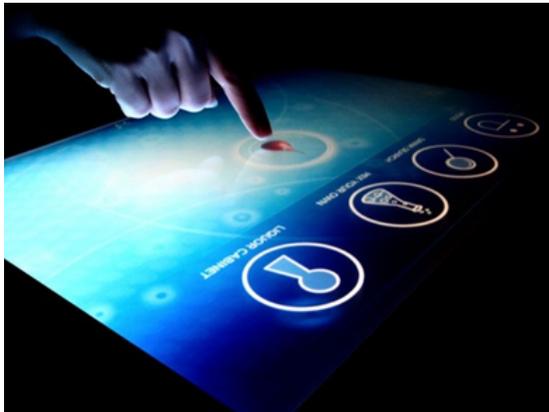


Fig. 7. Example of touch screen user interface.

#### G. Tangible User Interface

In this interface, the user interacts with the digital information using the physical environment [6]. In this way, physical movements are translated into digital information. The simplest example is the mouse movement to control arrow pointer on the display screen. Microsoft has also introduced a tangible user interface-based Windows platform, Microsoft Surface (new name is Microsoft PixelSense) in the year 2000.

#### H. Conversational Interface Agents

In this interface, messages, voices, chatbots with plain English text are presented to the user with the help of animated character, robot or person just like Microsoft Clippy for emulating human to human conversation.<sup>2</sup>

<sup>2</sup><http://www.cbc.ca/news/canada/toronto/toronto-chatbots-1.3581791>

#### I. Multi-Screen Interface

To provide more flexible interaction with the user, multiple screens are used. It is normally used in hand-held markets and computer games to show different statistics at multiple displays of stock markets, etc. (Fig. 8).



Fig. 8. Examples of multiscreen user interface.

#### J. Keyboard Driven Modeless User Interface

In this interface, keyboard shortcuts are used instead of typing command or clicking the options on toolbar or menu bar. Using keyboard shortcuts, the input speed is tremendously increased.<sup>3</sup>

#### K. Zooming User Interfaces (ZUI)

The ZUI is pronounced as zoo-ee in a graphical environment, in which user can zoom in or zoom out to view more or fewer details (see Fig. 9).



Fig. 9. Examples of Zooming UI.

<sup>3</sup><https://blog.mozilla.org/labs/2007/07/the-graphical-keyboard-user-interface/>

### L. Menu Driven Interface

It is a subtype of GUI in which applications have a menu bar with different menu items. By clicking the menu items or by moving the Mouse on these menu items, commands are executed (Fig. 10).

In the next section, we discuss all kinds of latest devices and interfaces where user interactions are not possible using commands or menus, icons, etc.

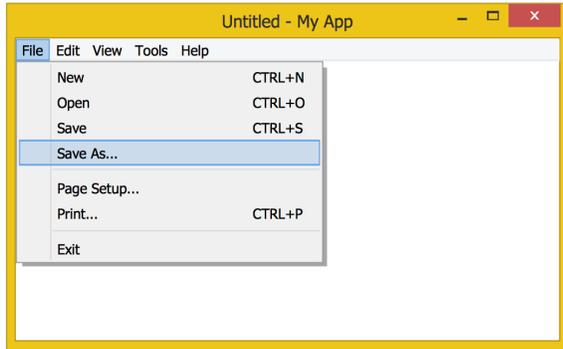


Fig. 10. Example of menu driven interface.

## IV. EMERGING USER INTERFACES (EUI)

There exist situations where the interaction with computers and machines is achieved by observing the user's body relative movements, direction of gaze, or gestures and respond to the specific human action accordingly without executing of commands or mouse clicks. It is achieved with the help of sensors for tracking the positions of body parts. The examples are Sixth Sense device, Google Glass, and gestures of tracking pad, magic mouse.

### A. Eye Movement-based Interface

In this interface, natural eyes are used as input for HCI. It is not using windows icon, menus and is highly interactive. It is natural and needs little conscious efforts. "What You Look At is What You Get". But it needs efficient recognition algorithms for eye-movements.

### B. Brain-Computer Interface

The use of mouse and keyboard for HCI is not suitable for disabled users. It uses brain signals similar to the eye-movements based interaction to communicate with the machines and computers.

### C. Gesture User Interface

Gestures are already being used in human conversations along with the speech for better understanding the conversation. Similarly, to interact with the computers or machines, gestures are also used. In this type of interaction, the input is accepted through human gestures like hand gestures, or mouse gestures received using a computer mouse/stylus [7] (Fig. 11). Gestures are captured with the help of cameras and stored in the database. When gestures are captured using input devices such as scanners or cameras, the processor inside these devices compares with the database and performs action accordingly [8].



Fig. 11. Examples of gesture-based interaction.

### D. Motion Tracking Interfaces

In this interface, human body movements are captured and translated into commands. It was first time introduced by Apple and the latest one uses a camera for this purpose [9].

### E. Voice User Interface

In this interface, the audio voice input is given to the system and after voice recognition, appropriate action is performed followed by voice prompts (Fig. 12). The examples are Siri application (APP) in Apple mobile phone, Google Glass, Cortana, Amazon Echo (Alexa), Google Now, Jibo Assistant.ai, and Ubi. Fig. 13 shows examples of interaction with the computer or machine using voice signals coming from the microphone.

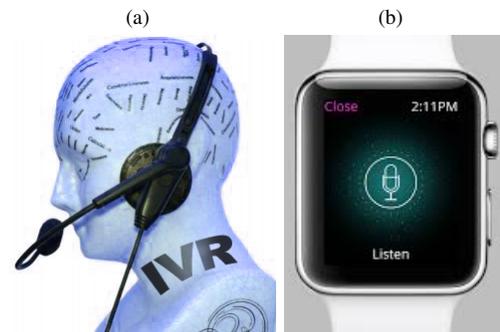


Fig. 12. Examples of voice user interfaces.

### F. Zero-Input Interfaces (ZII)

This interface is used for direct input to electronic circuits from sensors for HCI interaction rather than typing commands by the user.

### G. The SixthSense Device

The SixthSense device is developed by Pranav Mistryis, a PhD student in the Fluid Interfaces Group at MIT's Media Lab. The SixthSense device is used to interact with physical objects using human gestures.<sup>4</sup> If we consider inside the computer as a digital world and outside the computer as a physical world, the interface between the physical and digital worlds provides connections between objects and gestures. The

<sup>4</sup><https://tedsummaries.com/2014/03/18/pranav-mistry-the-thrilling-potential-of-sixthsense-technology/>



Fig. 13. Examples of interaction with the computer or machine using voice signals.

communication with the computer is achieved in a supernatural manner using gestures. The keyboard amazingly appears on hand and touching the figures on hand, the user can access the computer system (see Fig. 14(a)). Amazingly, a person is recognized by its profession as he appeared before the system (see Fig. 14(b)). Observing the time is amazingly achieved by drawing gesture of watch on the forehead and the watch appears (see Fig. 14(c)).

The idea was to combine both physical and digital worlds together. The device is always connected to the internet. Whenever user holds a book, the book's audio version is searched on the internet and the user can listen the audio of the book. Amazingly, playing the live talk of former American President Mr. Obama, just by reading the online news from the internet with the help of SixthSense device. Another interesting idea is playing games on paper (rotating or moving the paper control the car movements) and displaying some pictures on paper as well as printing the hard copy of pictures by touching the print button on paper (see Fig. 15).

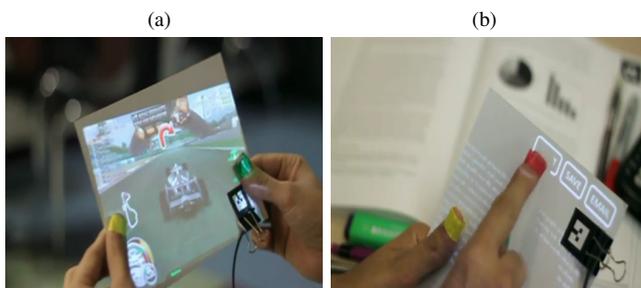


Fig. 15. Example SixthSense device (a) playing games and displaying on paper (b) printing from paper.

#### H. Google Glass

The Google Glass is an amazing wearable small gadget connected with the internet and interacts with the computer without commands or graphical objects. It is head mounted and shows information to the user without touching with hands [10], [11]. It contains a built-in camera, touchpad, and voice recognition software. It can take photographs and record videos as well as displaying to the users in front of their eyes. It has built-in Wi-Fi for surfing the internet and sending emails. It has built-in memory to store media files and messages. Fig. 16 shows an example of wearing Google Glass where the output is displayed in front of the eyes.

The Google Glass helps in internet surfing and searching data by sending verbal messages. It also records the daily activities of the user wearing it, and assist in various circumstances such as suggesting alternate routes in case of traffic congestion on a road. The communication with different languages is also possible using Google translate feature supported in Google Glass.

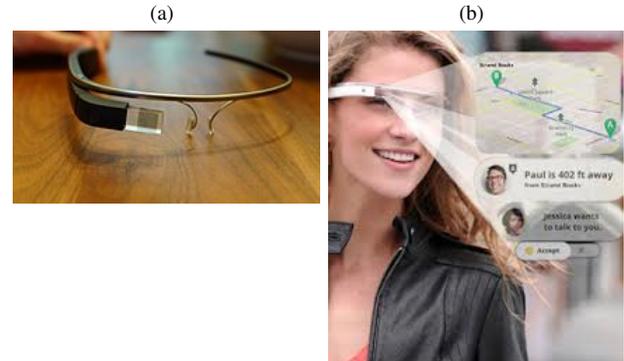


Fig. 16. Example Google Glass User Interface (a) Google Glass Gadget (b) Wearing Google Glass and display the output in front of the eyes.

#### I. Muse Headband

The Muse Headband is a mind reading device containing four sensors to read and extract information from mind waves.<sup>5</sup> The aim is to apply the human thinking in a computer such as changing the text font, playing games, printing documents on computers without physical interactions [12]. This interface may result in reducing the brain stress to remember or executing different commands. Fig. 17 shows the brain waves captured with the help of headband (right side). The text font is changed automatically, the moment, the idea of changing the font comes to the user's mind (left side).

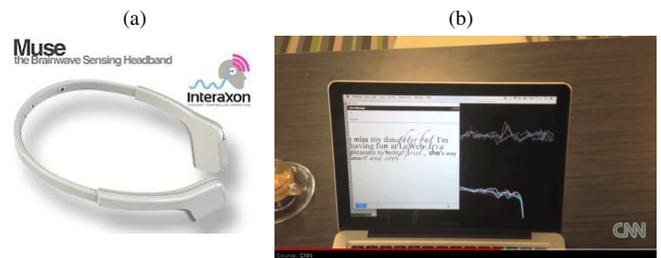


Fig. 17. Example (a) Muse Head Band (b) Muse Head User Interface to change the text font due to thinking only.

#### V. CONCLUSION

In this paper, an overview of existing user interfaces is provided. The user interfaces are grouped into three broad categories i.e., text-based, graphical, and emerging user interfaces. Among different user interfaces, the most commonly used ones are the CLI and the GUI. The CLI requires user training and it is tedious and laborious. Whereas, the GUI is simple and it requires minimal training. Recently, some new devices are invented which do not use the conventional user interfaces

<sup>5</sup><http://edition.cnn.com/2012/12/06/tech/ariel-garten-muse-interaxon/index.html?iref=allsearch>

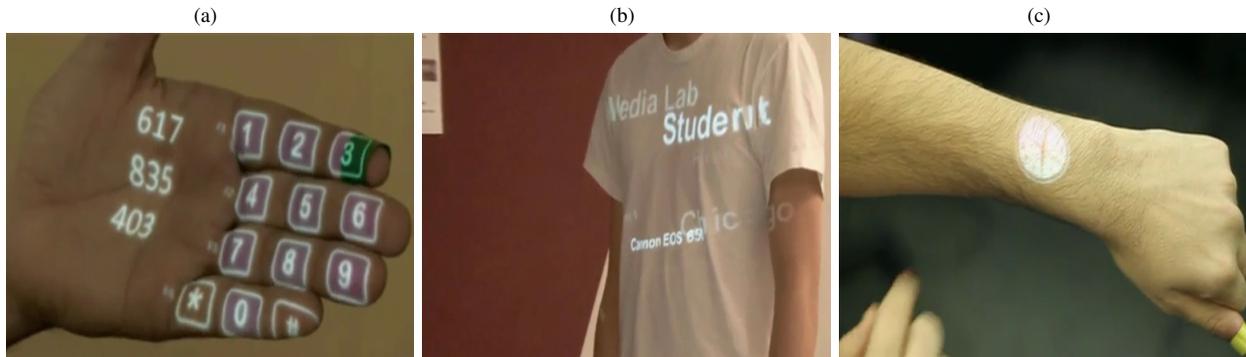


Fig. 14. Example SixthSense device (a) displaying keypad on palm (b) A person coming before this system recognized as student (c) watch on hand.

to interact with the machines and computers. Text based interfaces still being enhanced like Natural language interfaces. Touch screens have played an important role in mobile phone market. Gesture based interfaces have also important impact in the gaming industry. The Emerging user interfaces have a big share in the market of tablets, smartphones, television, gaming consoles, Google Glass, sixth sense devices, etc. In future, it is most likely the EUI may become famous just like the GUI. Using EUI, the integration of information with various devices will help human to be more connected to the physical world. It will eventually help in staying for a long time in front of machines. The survey will be helpful in designing new interfaces for interaction between the machines and computers.

#### REFERENCES

- [1] A. Razzaq, "A new emerging interface: Sorcerous user interface sui," *Researcher*, vol. 7, no. 5, 2015.
- [2] F. Li and H. V. Jagadish, "NaLIR: an interactive natural language interface for querying relational databases," in *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*. ACM, 2014, pp. 709–712.
- [3] B. A. Huberman and F. Wu, "The economics of attention: maximizing user value in information-rich environments," *Advances in Complex Systems*, vol. 11, no. 04, pp. 487–496, 2008.
- [4] M. Maybury, "Intelligent user interfaces: an introduction," in *Proceedings of the 4th international conference on Intelligent user interfaces*. ACM, 1998, pp. 3–4.
- [5] M. T. Maybury and W. Wahlster, *Readings in intelligent user interfaces*. Morgan Kaufmann, 1998.
- [6] H. Ishii, "The tangible user interface and its evolution," *Communications of the ACM*, vol. 51, no. 6, pp. 32–36, 2008.
- [7] M. A. Qureshi, A. Aziz, A. Saeed, M. Hayat, and S. Rasool, "Implementation of an efficient algorithm for human hand gesture identification," in *Electronics, Communications and Photonics Conference (SIEPC), 2011 Saudi International*. IEEE, 2011, pp. 1–5.
- [8] J. Rico, A. Crossan, and S. Brewster, "Gesture-based interfaces: practical applications of gestures in real world mobile settings," in *Whole Body Interaction*. Springer, 2011, pp. 173–186.
- [9] S. Fateh, "Motion detection and tracking system to control navigation and display of portable displays including on-chip gesture detection," Sep. 12 2005, uS Patent App. 11/225,877.
- [10] T. Starner, "Project glass: An extension of the self," *IEEE Pervasive Computing*, vol. 12, no. 2, pp. 14–16, 2013.
- [11] D. Pogue, "Google glass and the future of technology," *The New York Times*, 2012.
- [12] O. E. Krigolson, C. C. Williams, and F. L. Colino, "Using portable eeg to assess human visual attention," in *International Conference on Augmented Cognition*. Springer, 2017, pp. 56–65.

# OSPF vs EIGRP: A Comparative Analysis of CPU Utilization using OPNET

Muhammad Kashif Hanif<sup>\*</sup>, Ramzan Talib<sup>†</sup>, Nafees Ayub<sup>‡</sup>, Muhammad Umer Sarwar<sup>§</sup>, and Sami Ullah<sup>¶</sup>

Department of Computer Science,  
Government College University, Faisalabad, Pakistan

**Abstract**—Routing is difficult in enterprise networks because a packet might have to traverse many intermediary nodes to reach the final destination. The selection of an appropriate routing protocol for a large network is difficult task. The focus of this work is to select and identify the best routing technique for a computer network. In this study, the performance of OSPF and EIGRP routing protocols with respect to CPU utilization is analyzed using OPNET simulator. The results depict EIGRP acquires redundant information which effect CPU utilization.

**Keywords**—Network protocols; topology; OPNET; interior gateway protocols (IGPs); OSPF;

## I. INTRODUCTION

In this present era, computer networks are growing rapidly day by day. Communication technologies provide user convenient services such as file sharing/transferring, printer sharing, video streaming and video/voice conferencing. Internet is the global network of interconnected computers. Internet plays an important role in today's communication networks which are based on technology that provides the technical infrastructure. In this technical infrastructure routing protocols are used to find an efficient route to transmit packets across the internet.

In WAN, IP packets are being forwarded by the routers. For this purpose the routing devices use the routing protocols which determine and then select the best path to forward the packets [1]. Communication among different routing protocols depends on routing algorithms which base on the nodes to determine the route to forward the packet over networks [2]. Routing in internet plays an important role so it is the heartbeat of the internet. Routing protocols are comprised into two diverse categories: Interior gateway protocols (IGPs) and Exterior gateway protocols (EGPs). The IGPs such as OSPF and EIGRP work in an autonomous system routing whereas the EGPs (such as BGP) work for routing among multiple autonomous systems. Our effort in this research is to investigate the IGPs so the purpose of this investigation is to examine the OSPF and EIGRP regarding traffic-load and CPU utilization by nodes.

As the use of internet growing day by day, ISPs are trying to encounter the traffic demands with new technologies and enhanced existing resources. Network utilization depends on routing of data packets on a network because a packet follows a path to reach its destination. Intra-domain internet routing protocol mostly uses the OSPF to determine the best path for packets [3].

OSPF uses a topology to determine the short path. For this purpose it produces the link-state packets from each router in a network and these packets contain the updated information for routers. So OSPF uses this information to determine the path. If any change occurs in a network then recalculation process occurs [4].

Extended IGRP is an improved form of IGRP (interior gateway routing protocol). This improvement is the resultant of changes in routing and changes in demands of internet works. In this improvement, abilities of link-state protocols have been integrated into distance-vector protocols. Extended IGRP also consists of some vital protocols that increase its working competence than other different routing protocols. It uses DUAL that enables it to find whether the advertised path by a neighbor router is looped or loop-free. DUAL also permits the EIGRP to determine another route without waiting for updated routing information from other routers [5].

The objective of this work is to find routing protocol which has better CPU utilization and enhanced performance. Network administrators can use this study to select a protocol for computer network. In this work, protocols are analyzed with respect to the CPU utilization and link-state advertisements by considering various scenarios. The remainder of the paper is organized in different sections. Section II discusses related work. Section III describes the research methodology. The results are presented in Section IV. Section V concludes the outcomes.

## II. RELATED WORK

Agarwal et al. examined the effect of CPU utilization of Border Gateway Protocol (BGP) from multiple routers in the Sprint IP network. They correlated BGP with SNMP data to measure the CPU Utilization of 200 routers. The results showed BGP uses the majority of CPU cycles for short time slice. This is due to increased size of BGP routing table [6]. Shah and Rana analyzed the convergence and traffic for RIP and OSPF within network using OPNET simulator. Convergence time of OSPF single area is greater than OSPF multi area and OSPF multi stub area [7]. Nazumudeen and Mahendran compared the OSPF, EIGRP, RIPv1 and RIPv2 routing protocols to determine which protocol is suitable for a network. EIGRP performs better than RIP and OSPF due to speedy convergence process, great handling, and improved scalability of routing loops [8].

Kudtarkar et al. compared IGPs protocols using WFQ

technique by establishing dissimilar scenarios in OPNET. EIGRP perform better for non-real time applications while OSPF and IGRP have better performance for real time applications [9]. Vetrivelan et al. analyzed IGRP, EIGRP, RIP, and OSPF protocols to evaluate the performance for slip8\_gateway routers. Simulations showed OSPF has minimum transmission cost and IGRP has maximum overhead. Moreover, OSPF and EIGRP has better for maximum throughput and link utilization, respectively [10].

Patel and Pandey discussed the necessity to advertise route among routers for multiple routing protocols and autonomous systems in hybrid network. Moreover, they analyzed the the route summarization and redistribution for OSPF and EIGRP protocols. The route distribution and summarization reduces the memory, CPU utilization, and network traffic [11]. Masood et al. compared OSPF, EIGRP, and RIP using NSFnet topology over 14 nodes. OSPF showed better convergence while RIP has low CPU Utilization [12]. Stankoska et al. studied OSPF and EIGRP with respect to the convergence time, end-to-end delay, packet loss, jitter and throughput for video streaming and voice conferencing. EIGRP achieved better convergence, throughput and less packet loss. Researchers have not evaluated the routing protocols for spoke-to-hub topology. In this work, we study the OSPF and EIGRP for spoke-to-hub topology [13].

### III. METHODOLOGY

Different network scenarios were designed for spoke-to-hub topology to analyze the impact of CPU utilization for OSPF and EIGRP protocols. OPNET simulator was used to design the network scenarios and investigate the performance of these protocols over the designed scenarios. Spoke-to-hub network topology was designed using Cisco 2600, 7500 routers and switches. Within this topology, Local Area Networks (LAN), 100 Base T Ethernet link and Point to Point T1 link are used. The single area OSPF is used for this study.

Fig. 1 and 2 shows the scenarios where the EIGRP and OSPF protocols are deployed and configured. The scenario for this work has five spoke sites and one central hub to connect these sites. Ping request sent by one spoke site to other is represented by dotted line. Sample scenario for ping request from one spoke site to another can be as follows:

- Node 15 of site A sending a ping request to node 15 of site B.
- Node 3 sending a ping request to the core hub.
- Node 15 of site E sending the ping request to the node 15 of site D.

OSPF and EIGRP protocols are deployed on all routers in the network. The autonomous system number used for this network is 150. Each spoke site has multiple routers and switches. Multiple LANs exist in each spoke site which are connected with routers. Routers are linked to a central router through the switch. All spoke sites have same network structure. Eleven routers and one core switch are used at each spoke site which are directly linked to the central hub. Each spoke site consists of 10 LANs and multiple nodes. Fig. 3 shows a sample spoke site configured in this study.

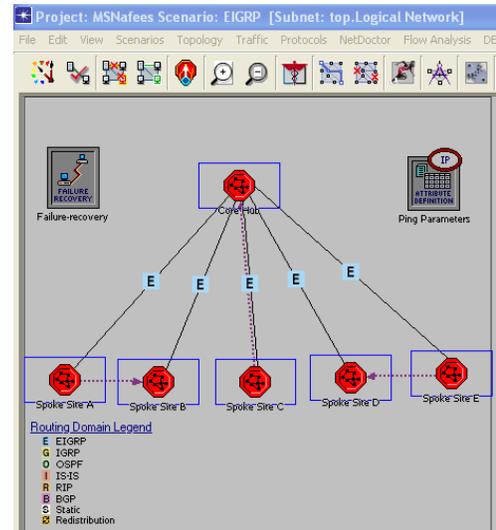


Fig. 1. EIGRP scenario.

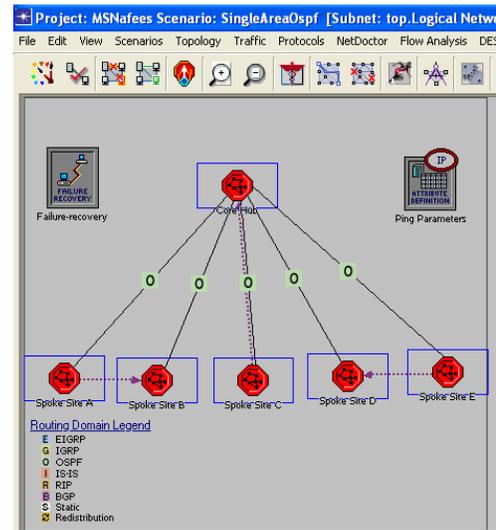


Fig. 2. OSPF scenario.

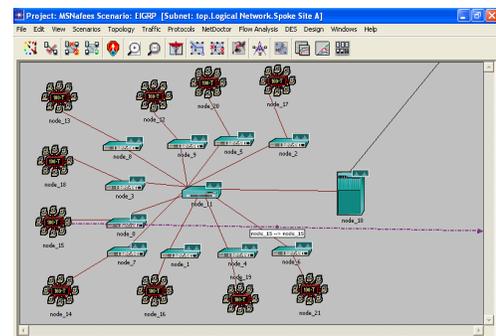


Fig. 3. Spoke Site A.

Central hub behaves like an intermediary among all spokes sites of network. Spoke sites are connected through this central hub (Fig. 4). Cisco 7500 series router is used as a central hub.

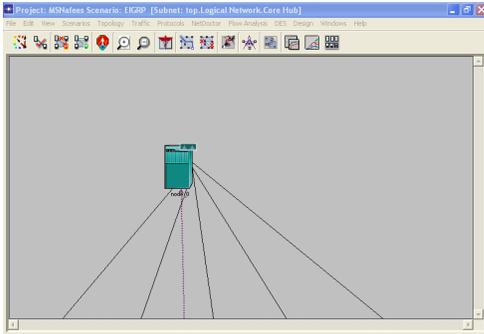


Fig. 4. Core Hub.

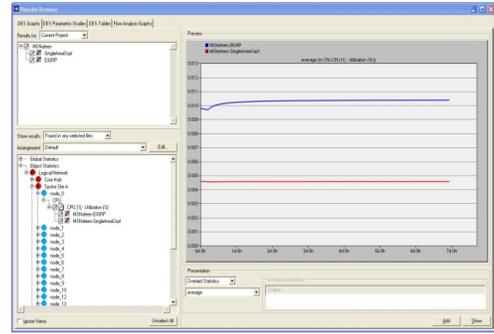


Fig. 7. Results browser for CPU utilization of Spoke Site A.

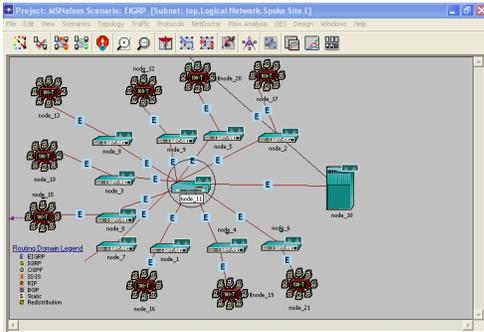


Fig. 5. Spoke Site E (EIGRP).

A. EIGRP Scenario

Fig. 5 shows that EIGRP is deployed on the routers of designed network. After the deployment of EIGRP the DES statistics are selected to examine the working of EIGRP.

B. OSPF Scenario

Fig. 6 shows a network that uses the OSPF in all of its routers. After the deployment of OSPF in this network the DES statistics are selected to evaluate the working of OSPF.

C. Node Description

Node models are created using Node Editor and then these node models are used for creating instances of node existing in network. A node is defined by connecting different modules

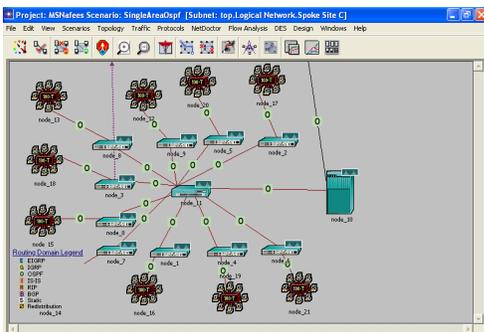


Fig. 6. Spoke Site C (OSPF).

with packet streams and statistics. This connection among different modules allows packets and status to be exchanged. The modules in this connection serve for a specific purpose like producing packets, line up the produced packets, processing on packets and exchanging these packets. In our topology we use router 7609 as a core hub. The other nodes used in our topology are Cisco router 7513 and switch 3000.

IV. RESULTS AND DISCUSSION

The graphs in this section are the resultant of simulation. We compared the graphs of CPU Utilization for node\_0 for each of the spoke sites. The resultant full window of CPU Utilization of node\_0 of spoke site A is shown in Fig. 7. As shown in the figure, the CPU Utilization for both EIGRP and OSPF are checked in the window's left side. Then the right side shows the graph for both protocols. The result browsers for all the spoke sites are same as shown for spoke site A but the resultant graphs are different those are explained in this section.

The axis description for all the graphs is as follows:

- Time duration for CPU Utilization is represented on X-axis.
- Percentage for CPU Utilization at some specific time is represented on Y-axis.

A. CPU Utilization of Node\_0 in all Spoke Sites

Fig. 8 shows graph for CPU Utilization of Node\_0 in all of the Spoke Sites. As shown in the figure the EIGRP curve in all sub figures changes its behavior along y-axis and then after some time the curve goes parallel along x-axis. But on the other hand the OSPF curve goes parallel along x-axis from the start of the simulation. As shown in the graph the behavior of OSPF curve is better than the EIGRP curve. This performance of OSPF is because of exchanging less routing updates than EIGRP. EIGRP needs the same information again and again so EIGRP wastes time and resources more than OSPF.

As shown in the graphs the behavior of the OSPF is better than the behavior of the EIGRP because the EIGRP is less intelligent than OSPF. The curve for EIGRP is illustrating the different behavior in each graph but the curve for OSPF is same in behavior.

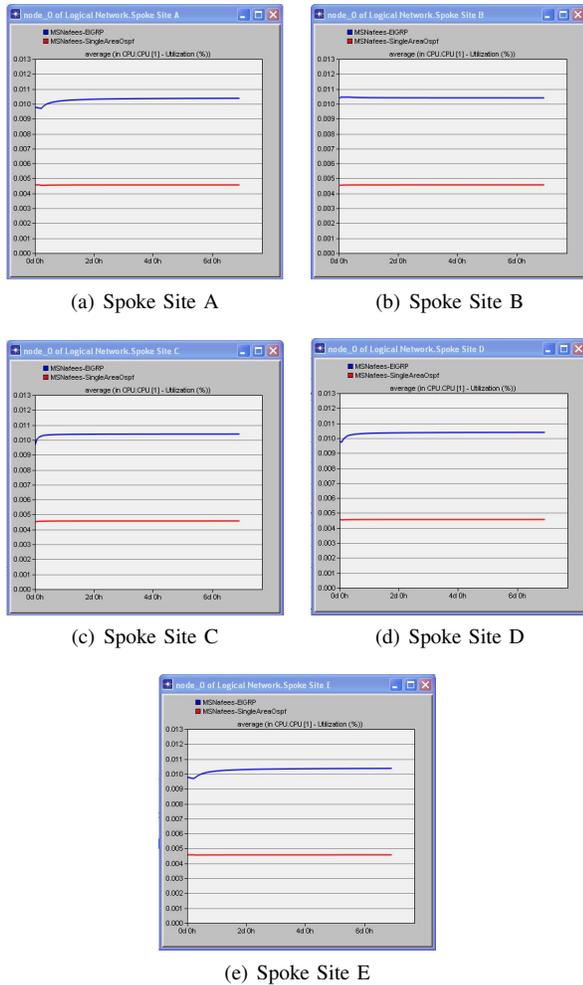


Fig. 8. CPU utilization of Node\_0 in Spoke Sites.

## V. CONCLUSION

Amongst the several findings of our working the most important was about CPU Utilization. So, in this paper the investigating results showed that EIGRP deals with more routing information like additions, deletions and updates than the OSPF. As EIGRP requires updated information again and

again and OSPF relies on previous information so the repeated steps in OSPF are reduced. The resultant investigation of our research describes that EIGRP acquires same information again and again so being less intellectual than OSPF it wastes time and resources like CPU.

## REFERENCES

- [1] D. Xu and L. Trajkovic, "Performance analysis of rip, eigrp, and ospf using opnet," 2011.
- [2] J. Singh and D. R. Mahajan, "Simulation based comparative study of rip, ospf and eigrp," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 8, 2013.
- [3] M. Ericsson, M. G. C. Resende, and P. M. Pardalos, "A genetic algorithm for the weight setting problem in ospf routing," *Journal of combinatorial optimization*, vol. 6, no. 3, pp. 299–333, 2002.
- [4] H. Pun, "Convergence behavior of rip and ospf network protocols," Ph.D. dissertation, Simon Fraser University, 2001.
- [5] A. Frantsi and H. Venalainen, "Testing demands and routing protocols," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, pp. 1–5, 2006.
- [6] S. Agarwal, C.-N. Chuah, S. Bhattacharyya, and C. Diot, "Impact of bgp dynamics on router cpu utilization," in *International Workshop on Passive and Active Network Measurement*. Springer, 2004, pp. 278–288.
- [7] A. Shah and W. Rana, "Performance analysis of rip and ospf network using opnet," *ICSI*, vol. 10, no. 6, pp. 256–265, 2013.
- [8] N. Nazumudeen and C. Mahendran, "Performance analysis of dynamic routing protocols using packet tracer," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, 2014.
- [9] A. Kudtarkar, R. Sonkusare, and D. Ambawade, "Performance analysis of routing protocols for real time application," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 3, no. 1, pp. 5072–5072, 2014.
- [10] V. Vetrivelvan, P. R. Patil, and M. Mahendran, "Survey on the rip, ospf, eigrp routing protocols," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 2, pp. 1058–1065, 2014.
- [11] H. N. Patel and R. Pandey, "Extensive reviews of ospf and eigrp routing protocols based on route summarization and route redistribution," *Int. Journal of Engineering Research and Applications*, vol. 4, no. 9, pp. 141–144, 2014.
- [12] M. Masood, M. Abuhelala, and I. Glesk, "A comprehensive study of routing protocols performance with topological changes in standard networks," *International Journal of Electronics, Electrical and Computational System (IJECS)*, vol. 5, no. 8, pp. 31–40, 2016.
- [13] E. Stankoska, N. Rendevski, and P. Mitrevski, "Simulation based comparative performance analysis of ospf and eigrp routing protocols," 2016.

# Anonymized Social Networks Community Preservation

Jyothi Vadisala

Dept. of Computer Science & Systems Engineering  
College of Engineering(A), Andhra University  
Visakhapatnam, India

Valli Kumari Vatsavayi

Dept. of Computer Science & Systems Engineering  
College of Engineering(A), Andhra University  
Visakhapatnam, India

**Abstract**—Social Networks have been widely used in the society. Most of the people are connected to one another, communicated with each other and share the information in different forms. The information gathered from different social networking sites is growing tremendously in large volumes of various research, marketing and other purposes which is creating security and privacy concerns. The gathered information contains some sensitive and private information about an individual, such as the relationship of an individual or group information. So, to protect the data from unauthorized users the data should be anonymized before publishing. In this paper, we study how the  $k$ -degree and  $k$ -NMF anonymized methods preserve the existing communities of the original social networks. We use an existing heuristic algorithm called Louvian method to identify the communities in social networks. We conduct the experiments on real data sets and compare the performances of the two anonymized social networks for preservation of communities of the original social networks.

**Keywords**—Community; anonymity; degree; social network

## I. INTRODUCTION

Social networks are ubiquitous these days and are widely used for communication. The people are connected, whether near or far, anyone can be connected through social networks to anyone they want to and share the information like images, videos and text, etc. This data is published for various research purposes. Facebook, Twitter, Goggle are the best examples of social media where people share their information. These social networks must provide the privacy to their members and a privacy policy regarding how the collected data is used and published for various purposes. To protect the privacy of individuals the data must be anonymized before publishing the data. There are different anonymization algorithms which anonymizes the data. Most of the social network data are represented by graphs so there is no standard anonymization method which protects the privacy of individuals. In general, the privacy protection either identity of individuals, the relationship of individuals and the node content of their network. There are different anonymization methods and are applicable for appropriate privacy risks like anonymization via modification of the original graph, anonymization via clustering and differential privacy, etc. In this paper, we study how well the anonymized networks preserve the existing communities of the initial networks.

The communities of a social network mean groups of nodes which have similar characteristics or properties. There

are different community detection algorithms presented in the paper [1]. In this paper, we use a heuristic algorithm called a Louvian method [2] based on modularity optimization. The modularity function has two values either positive or negative. The positive values indicate the presence of community structure possibilities. We follow a two steps to study how well the anonymized networks preserve the communities of the initial social networks. First, the initial network is anonymized by the two approaches, i.e  $k$ -degree anonymization and  $k$ -NMF anonymization. Second, we apply Louvian method to detect the communities from the anonymized networks and compare the two methods of preservation of communities of the initial networks by conducting experiments on real data sets.

## II. RELATED WORK

Several studies address the need of anonymizing the social networks to protect the privacy of individuals. Most of the prior work focus on preserving the structural properties between the original and anonymized social networks. The complete survey of existing social networks anonymization methods and the other privacy issues of the social networks is covered in [3]. An another important study of social networks is that of identifying communities in the network. Generally, communities are groups where we can identify the groups of interacting the nodes and the relationship between them like the friends group who studied in the same school or working in the same company, etc. There are so many papers which discussed how to detect the communities from the social networks. There are different algorithms are used to detect the communities. In Moradi and Olovsson et al. [4] used large e-mail networks to experimentally evaluate the qualitative performance of several community detection algorithms. In Malliaros and Vazirgiannis [5] suggested a methodology-based taxonomy to classify the different community detection approaches for directed graphs. The Ruan and Zhang [6] proposed a modularity measure to assess the quality of community structures. To compare the different community structures the modularity measure is well used. A larger modularity value means stronger community structures. The optimization of modularity measure is proposed in Newman [7], Duch and Arenas [8].

In this paper, we study the two graph modification approaches ( $k$ -degree and  $k$ -NMF anonymizations) and we focus on how these methods preserve communities of the original social network. To conduct experiments we consider the publicly available data sets and compare the results for the both

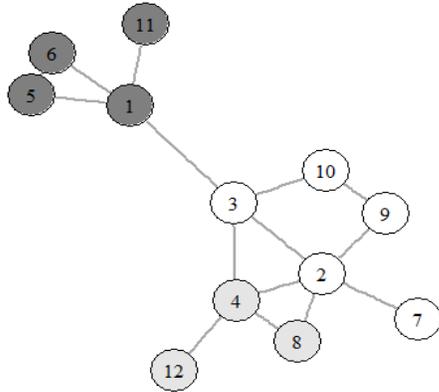


Fig. 1. Initial social network( $G_1$ ).

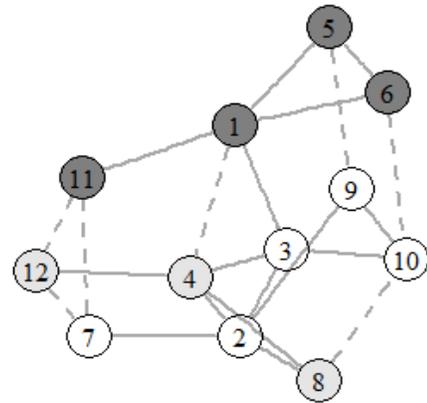


Fig. 2.  $k$ -Degree anonymized social network( $G'_2$ ).

anonymized methods.

### III. MODELS FOR SOCIAL NETWORK ANONYMITY

In this section we present the two anonymization techniques  $k$ -degree anonymity and  $k$ -NMF anonymity and we focus on the preservation of communities based on the structure of social networks. The process of anonymization is also based on the social network structural properties. Generally, a social networks are modelled as a graph  $G = (V, E)$  where  $V$  is the set of vertices and  $E$  is the set of edges which represents the relationship between these vertices. In this paper, we consider the social network  $G$  as a simple undirected graph and the intruder knows the structure of the network and able to identify the individuals along the sensitive information due to the unique structure of the social network data. Fig. 1 shows the example of a social network which has 12 nodes and 15 edges.

#### A. $k$ -degree anonymity

$k$ -degree anonymity is the extension of well-known  $k$ -anonymity model where the intruder has the knowledge of the vertex degree to breach the identity of vertices. This method is a vertex based anonymization technique where there is at least  $k - 1$  other vertices have the same degree. Liu, K.Terzi et al.[9] created an initial algorithm and proposed a  $k$ -degree anonymous network based on the degree property of the network. In this paper, we consider a Fast  $k$ -degree Anonymization Algorithm (FKDA) which is proposed by Lu et al.[10].

FKDA is a greedy algorithm in which the social network is anonymized by edge addition to the network until the network is  $k$ -degree anonymous. FKDA is a two step process, in Step 1 the vertices of original network is separated into several groups. Step 2, select each group and anonymize by adding edges to the vertices of the same group until all the vertices have the same degree in that group. If the group does not achieve the anonymization by edge creation, then it adds the edges by the relaxed edge addition method in which the vertices in that group are anonymized by connecting to other vertices in the graph rather than the same group. But the relax

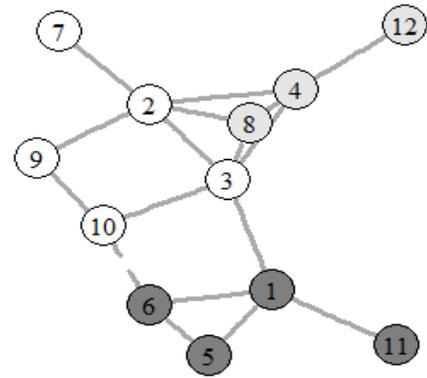


Fig. 3.  $k$ -NMF anonymized social network( $G'_3$ ).

edge addition method may destroy the previously anonymized groups and the whole process will be restarted. The worst case time complexity of performing this approach is  $\mathcal{O}(V^2)$  where  $V$  is the total number of vertices in the network. The Fig. 2 shows the anonymization of the graph  $G'_1$  using FKDA technique. The dashed edges represent the newly added edges by FKDA algorithm. The network has 3 nodes has degree 5, 2 nodes have degree 4 and 7 nodes has degree 3 so the network satisfies the 2-degree anonymity where  $k = 2$ . The details of the algorithm are specified in [10].

#### B. $k$ -NMF anonymity

In this method, we anonymize the original graph only by edge addition. The intruder has the background knowledge of the number of common vertices of an edge. This method is an edge based anonymization technique where there are at least  $k$  edges in a group has the same count of the number of common vertices. In  $k$ -NMF anonymization [11] first, we group the more than  $k$  edges and second, each edge of this group is anonymized by the breadth first search method. The edges are updated dynamically in the edge list because some new edges are added into the network. Therefore, when adding one edge will affect the count of the number of common friends of another edge or more edges. So it has to follow the anonymized triangle preservation principle which aims to preserve the already anonymized edges neither creating

some additional anonymized triangles by edge addition or destroy by the edge deletion. This preservation leads to avoid repeatedly anonymizing the same edges. The Fig. 3 shows an example of  $k$ -NMF anonymization process for the graph  $G_1$ .

The  $k$ -NMF anonymization problem can be seen as a parallel of  $k$ -degree anonymization problem. As in  $k$ -degree anonymization process needs more number of edge additions than  $k$ -NMF anonymity so most of the structural properties of a graph will be preserved by the  $k$ -NMF anonymization algorithm. This significant difference in the privacy protection of individuals between the two methods leads us to the  $k$ -NMF anonymity will preserve the communities of the original social networks than the  $k$ -degree anonymity model. This can be explained by conducting experiments on real data sets.

#### IV. COMMUNITY DETECTION

In this paper, we focus on preserving communities by anonymized social networks. Generally, identifying a community in complex networks is a universal problem and has consequently been raised in many domains, leading to different solutions. Most of the community detection methods rely on Newman's modularity to assess the quality of their results. The community detection algorithms are grouped as hierarchical, optimization and others. In hierarchical approach the result is a tree of the communities which is represented as a dendrogram. The hierarchical method consists of two approaches, i.e. Agglomerative and Divisive. The optimization-based approaches use a Newman's modularity measure to calculate the quality of a network partition. The algorithm consists of two steps. In Step 1, processing several partitions of the network either randomly or by a fitting function and in Step 2 based on quality measure choose the best nodes and this algorithm is modified to get the better quality. Most of the optimization algorithms have used a modularity measure because it is a costly measure to process [12], [13], [14]. Other algorithms use a clustering principle [15] [16] and also find the overlapping communities, i.e. one node may be a part of several communities at once [17]. In Derneyi et al. [18] used the latent space approach to process the probability for a node to belong to a community. In this paper, we use a Louvain community detection method for detecting the communities from the original and anonymized social networks. In this method each node is assigned to one community. Then the modularity gain of each community is maximized by moving nodes between those communities. This step is stopped when there is no change in modularity gain with the movement of nodes. After this process the network obtained from the first step is used and a weighted network is created. In this weighted network, one node represents a community from the original network, and weights are added to edges to represent the number of original edges that are collapsed into a super edge. After the completion of this step again the first step is implemented. This process repeated iteratively until the modularity gain is maximized. Communities obtained by the Louvain method for a graph  $G_1$  are shown in Fig. 1. The color of a vertex represents the community they belong.

TABLE I. PRESERVATION AT COMMUNITY LEVEL ( $k$ -DEGREE ANONYMITY)

Community	Communities in $G_1$	Communities in $G'_2$	Preservation of Community(%)
1	{1,5,6,11}	{1,5,6,9,10}	33%
2	{4,8,12}	{2,3,4,8}	66%
3	{2,3,7,9,10}	{7,11,12}	20%

TABLE II. PRESERVATION AT COMMUNITY LEVEL ( $k$ -NMF ANONYMITY)

Community	Communities in $G_1$	Communities in $G'_2$	Preservation of Community(%)
1	{1,5,6,11}	{1,5,6,11}	100%
2	{4,8,12}	{2,3,4,7,8,12}	100%
3	{2,3,7,9,10}	{9,10}	40%

#### A. Community Preservation

In this section, we estimate the preservation of communities by the anonymized social networks and compare with the communities of the original social networks. We compute the communities of the anonymized social networks and original networks using Louvain method and compare the results between the anonymized and original networks using two different approaches.

1) *Preservation at Community Level (PCL)*: In this, we count how many vertices have remained in the same community after the anonymization process. The preservation of communities by two anonymization methods is shown in Table I and Table II. The percentage of preservation of communities is calculated for each community of the graph  $G_1$  with the corresponding community of Graphs  $G'_2$  or  $G'_3$  that contain the maximum number of vertices from the original community. The *PCL* value for a network will be calculated by the average of the results for the percentage of preservation of each community.

The percentage of preservation of each community in the initial social network and anonymized social networks is shown in Tables I and II. For example the percentage preservation for the second community from Table I i.e. {4,8,12}, the best match is the community {2,3,4,8} and the percentage of preservation is  $\frac{2}{3}$ . To measure the percentage of preservation for the network is the sum of all results of the percentage of preservation of communities divided by the total number of communities. The *PCL* values for the two anonymized social networks is given below:

- $PCL(G_1, G'_2) = 39.66\%$
- $PCL(G_1, G'_3) = 80\%$

2) *Preservation of Community at Node Level (PCNL)*: In this section, we estimate the preservation of communities at each node individually. We compare the community of each node at original network and anonymized network. Consider the initial social network as  $G = (V, E)$  and the anonymized social networks as  $G' = (V, E')$ . The set of nodes  $V = v_1, v_2, \dots, v_n$  and the  $Com(v_i)$  and  $Com'(v_i)$  represent the node community at original and anonymized social networks respectively. The community preservation for each node  $v_i$  is

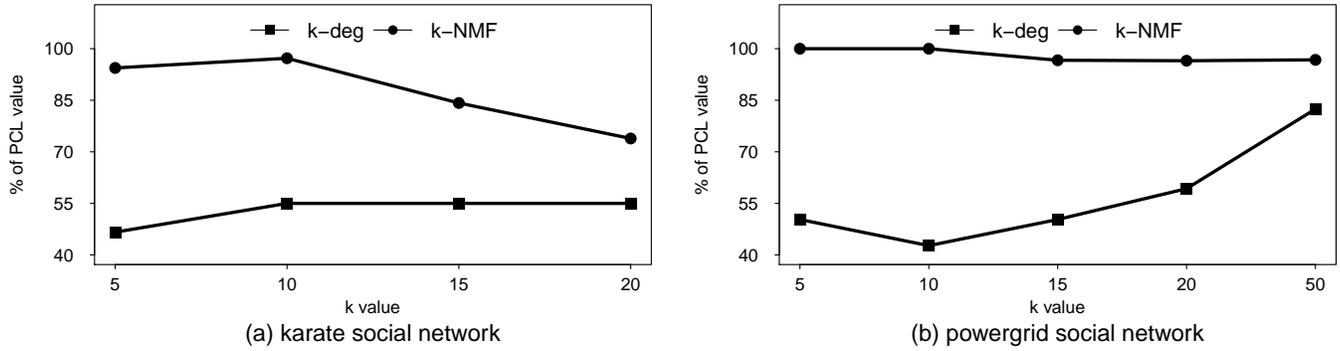


Fig. 4. Percentage of preservation of community level (%PCL).

TABLE III. PRESERVATION OF COMMUNITY AT NODE LEVEL (*k*-DEGREE ANONYMITY)

Node	$Com(v_i)$	$Com'(v_i)$	$PCNL(v_i)$
1	{1,5,6,11}	{1,5,6,9,10}	50%
2	{2,3,7,9,10}	{2,3,4,8}	28.57%
3	{2,3,7,9,10}	{2,3,4,8}	28.57%
4	{4,8,12}	{2,3,4,8}	40%
5	{1,5,6,11}	{1,5,6,9,10}	50%
6	{1,5,6,11}	{1,5,6,9,10}	50%
7	{2,3,7,9,10}	{7,11,12}	14.28%
8	{4,8,12}	{2,3,4,8}	40%
9	{2,3,7,9,10}	{1,5,6,9,10}	25%
10	{2,3,7,9,10}	{1,5,6,9,10}	25%
11	{1,5,6,11}	{7,11,12}	16.66%
12	{4,8,12}	{7,11,12}	20%

TABLE IV. PRESERVATION OF COMMUNITY AT NODE LEVEL (*k*-NMF ANONYMITY)

Node	$Com(v_i)$	$Com'(v_i)$	$PCNL(v_i)$
1	{1,5,6,11}	{1,5,6,11}	100%
2	{2,3,7,9,10}	{2,3,4,7,8,12}	37.5%
3	{2,3,7,9,10}	{2,3,4,7,8,12}	37.5%
4	{4,8,12}	{2,3,4,7,8,12}	50%
5	{1,5,6,11}	{1,5,6,11}	100%
6	{1,5,6,11}	{1,5,6,11}	100%
7	{2,3,7,9,10}	{2,3,4,7,8,12}	37.5%
8	{4,8,12}	{2,3,4,7,8,12}	50%
9	{2,3,7,9,10}	{9,10}	40%
10	{2,3,7,9,10}	{9,10}	40%
11	{1,5,6,11}	{1,5,6,11}	100%
12	{4,8,12}	{2,3,4,7,8,12}	50%

the number of nodes common in both  $Com(v_i)$  and  $Com'(v_i)$  divided by the at least one of these two communities.

$$PCNL(v_i) = \frac{|Com(v_i) \cap Com'(v_i)|}{|Com(v_i) \cup Com'(v_i)|} \quad (1)$$

Where  $|V|$  represents the number of elements in set  $V$ . The final  $PCNL$  value is calculated as the sum of all individual preservation of community node values divided by the total number of nodes in the network is shown below:

$$PCNL(G, G') = \frac{\sum_{i=1}^n PCNL(v_i)}{n} \quad (2)$$

The preservation of community at node level for the two anonymized networks of Fig. 2 & 3. is shown in Tables III & IV. To illustrate this computation, let us consider the node 4 from Table III. The initial community for the node 4 is {4,8,12} and the  $k$ -degree anonymized community is {2,3,4,8}. By observation, there are two nodes common in these two sets i.e., {4,8} and 5 nodes in their union of sets i.e {2,3,4,8,12}. So the  $PCNL$  value for node 5 is  $\frac{2}{5}$ . The final preservation of community at node level for each anonymized social network with respect to original network is shown below:

- $PCNL(G_1, G'_2) = 32.34\%$
- $PCNL(G_1, G'_3) = 61.87\%$

## V. EXPERIMENTAL RESULTS

In this section, the following publicly available data sets are used for the preservation of communities between original and anonymized social networks.

- Zacharys karate club is a small undirected friendship relation social network. It has 34 nodes, 78 edges and 4 communities.
- A Power grid is an undirected, unweighted network representing the topology of the western states power grid of the united states. This network consists of 4,941 nodes, 6,594 edges and 40 communities.

We performed different steps to measure the preservation of communities. In step1, first we consider the above initial networks, and anonymize these networks by the two anonymization methods (FKDA,  $k$ -NMF) using several anonymity values of  $k$  i.e 5, 10, 15, 20 and 50. Next we calculated the communities of the initial networks, and each anonymized network using a Louvian method in R programming. Finally, we compute the preservation of communities using  $PCL$  and  $PCNL$  approaches and plot the average results of  $PCL$  and  $PCNL$  values for the above networks.

Fig. 4 represents the percentage of preservation of communities of karate and power grid social networks. In both the networks, a  $k$ -NMF method preserves the communities

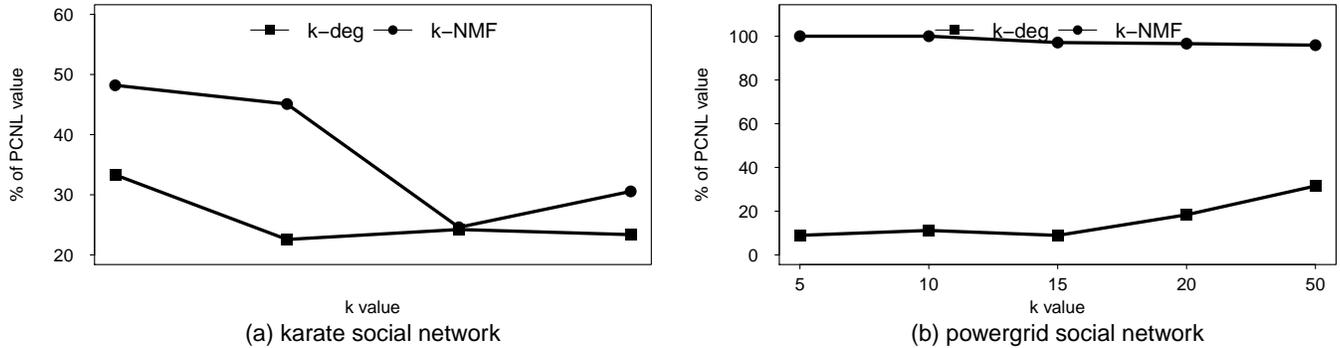


Fig. 5. Percentage of preservation of community at node level (%PCNL).

of original network very well. In  $k$ -NMF anonymization preserved the communities of the original network well when  $k$  is small but as the  $k$  value increases the preservation of communities is decreased for the karate social network dataset. Similarly, Fig. 5 represents the percentage of preservation of communities at node level is also well preserved by  $k$ -NMF anonymization method for both the data sets. But at the  $k = 15$  the PCNL value decreased for karate network. Based on the above results, we conclude that the  $k$ -NMF algorithm preserves the communities of original network very well using both PCL and PCNL measures.

## VI. CONCLUSION

In this work, we focused on how well the anonymized social networks preserve the communities of the original social networks. We analyzed  $k$ -degree anonymization model where the adversary identifies a vertex based on the degree of a vertex as a background knowledge, whereas a  $k$ -NMF model, the adversary identifies an edge, based on the number of common friends of the connected edge as a background knowledge. Our results show that the  $k$ -NMF model preserves the very well communities than the  $k$ -degree model. However, there are several future directions have to be considered. First, while anonymizing the social networks if the number of vertices are increased, then how well the anonymized networks will preserve the communities of the original network. Second, while anonymizing if the number of communities are increases or decreases how the communities are preserved in large complex networks. Our method does not discuss these situations, therefore we plan in future to create a more robust way of comparing community preservation.

## REFERENCES

- [1] Gunce Orman, Vincent Labatut. A Comparison of Community Detection Algorithms on Artificial Networks. Discovery Science (DS), 2009, Porto, Portugal. Springer, 5808, pp. 242-256, 2009.
- [2] Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." Journal of statistical mechanics: theory and experiment 2008.10 (2008): P10008.
- [3] Jyothi Vadisala and Valli Kumari Vatsavayi, Challenges in Social Network Data Privacy, International Journal of Computational Intelligence Research (IJCIR), vol -13, pp. 965-979, 2017.
- [4] F Moradi, T Olovsson, P Tsigas, An of community detection algorithms on large-scale email traffic, in: SEA. Berlin/Heidelberg: Springer; 2012; 283294.
- [5] F.D Malliaros, M Vazirgiannis, Clustering and community detection in directed networks: a survey, CoRR, abs/1308.0971, 2013.
- [6] J. Ruan and W. Zhang, An Efficient Spectral Algorithm for Network Community Discovery and Its Applications to Biological and Social Networks, Proc. Seventh IEEE Intl Conf. Data Mining (ICDM 07), pp. 643-648, Jan. 2007.
- [7] M. Newman, The Structure and Function of Complex Networks, SIAM Rev., vol. 45, no. 2, pp. 167-256, 2003.
- [8] Jordi Duch, Alex Arenas, Community detection in complex networks using extremal optimization, Phys. Rev E72, 027104, August 2005.
- [9] Liu, Kun, and Evimaria Terzi. "Towards identity anonymization on graphs." Proceedings of the 2008 ACM SIGMOD international conference on Management of data. ACM, 2008.
- [10] Lu, Xuesong, Yi Song, and Sthpne Bressan. "Fast identity anonymization on graphs.", proceedings of the 23rd International Conference on Database and Expert Systems Applications. Springer Berlin/Heidelberg, pp. 281-295, 2012.
- [11] C. Sun, P. S. Yu, X. Kong and Y. Fu, "Privacy Preserving Social Network Publication against Mutual Friend Attacks," 2013 IEEE 13th International Conference on Data Mining Workshops, Dallas, TX, 2013, pp. 883-890.
- [12] Duch, J., Arenas, A.: Community detection in complex networks using extremal optimization. Physical Review E 72 (2005) 027104 (3).
- [13] Reichardt, J., Bornholdt, S.: Detecting Fuzzy Community Structures in Complex Networks with a Potts Model. Physical Review Letters 93 (2004) 218701.
- [14] Clauset, A., Newman, M.E.J., Moore, C.: Finding community structure in very large networks. Physical Review E 70 (2004) 066111.
- [15] Falkowski, T., Barth, A., Spiliopoulou, M.: DENGGRAPH: A Density-based Community Detection Algorithm. IEEE/WIC/ACM International Conference on Web Intelligence (2007) 112-115.
- [16] Liu, Y., Wang, Q., Wang, Q., Yao, Q., Liu, Y.: Email Community Detection Using Artificial Ant Colony Clustering. Advances in Web and Network Technologies, and Information Management. Springer, Berlin / Heidelberg (2007) 287-298.
- [17] Hoff, P., Raftery, A., Handcock, M.: Latent space approaches to social network analysis. Journal of the American Statistical Association 97 (2002) 1090-1098.
- [18] Derenyi, I., Palla, G., Vicsek, T.: Clique percolation in random networks. Physical Review Letters 94 (2005).

# A Comparative Study for Performance and Power Consumption of FPGA Digital Interpolation Filters

Tim Donnelly, Jungu Choi, Alexander V. Kildishev, Matthew Swabey, Mark C. Johnson  
Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907

**Abstract**—The development of FPGA-based digital signal processing devices has been gaining attention. Researchers seek to reduce power consumption and enhance signal processing quality in these devices with given resources and spatial limits. Hence, there is a need to investigate both the capability and the power consumption associated with the various digital filtering schemes commonly used in FPGA-based devices. We carry out a set of performance and power consumption measurements of interpolation filters using an FPGA and other basic signal processing building blocks. We compare the signal processing performance with theoretical prediction, and measure the power consumed by the filters. Our experimental measurements also confirm the accuracy of the numerical tools used for predicting FPGA power consumption. This paper is aimed at providing a framework to accurately test basic signal processing across various interpolation schemes and compare the respective schemes' software-side contributions to power consumption and filtering quality.

**Keywords**—Digital signal processing; digital interpolation filters; FPGA

## I. INTRODUCTION

FPGA filters are widely used in broad applications, including radio-frequency sensors [1]–[3], imaging [4]–[7] and wearable medical devices [8]–[14], due to its flexible programmability. Although versatile, the FPGA-based filters are known to consume considerable power as a draw-back. Previously, researchers have looked into the performance and power estimates of FPGA-based digital filters [15], [16] and these are still relevant issues as consequence of the growing demand for quality signal processing with low power consumption. It is therefore of great interest to understand how the capability of such devices are influenced by the nature of the energy constrained environment in which they operate. In the industry, the effort to optimize power consumption in FPGA-based devices heavily extends to a hardware side such as clock gating [17] and word-length [18]. And power estimation methodology was developed [19] to identify high-power-consuming groups and improve on FPGA power dynamic models.

Taking wearable medical devices as an example, consider how in [20]–[22] remote devices for monitoring a patient's brain-wave were developed and assessed for their efficiency and power consumption. The intra-body sensor in [2] showed clear trade-offs between power consumption and signal processing quality. Likewise, the increasing digital signal processing (DSP) capability of portable heart monitoring devices [23], [24] must be accounted for as part of the available energy budget. Overall, a big challenge biomedical researchers face is to make devices such as these comfortable for patients so as not to compromise their daily activities while still implementing the advanced signal processing algorithms. That is, the design

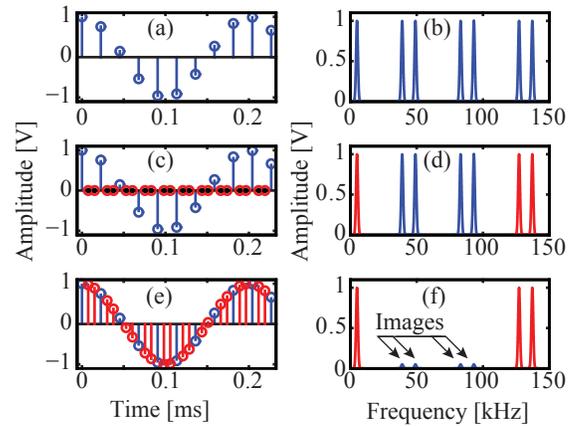


Fig. 1. Depiction of digital interpolation. (a) Original signal with (b) its spectral representation; (c) zeros are added between sampled data and (d) the effective sampling rate increases; (e) interpolation is performed and new data are created (red) and (f) low frequency sampling images are eliminated (significantly attenuated).

of a comfortable device for patients is at odds with the large (heavy) battery needed to meet the power requirements of such a device.

In addition, we note that numerous devices utilize hybrid-filter configurations to optimize performance and reduce power consumption. Most hearing aids contain filter-banks for certain frequency components of sound to serve an individual's hearing needs [8]–[12]. A novel electrocardiogram (ECG) design [14] by Hong et al. contains high-pass FIR filter followed by linear interpolation for quality signal processing with low power. An FPGA-based accelerometer [25] by Ramaesh et al. has FIR filters with cascaded-integrator comb (CIC) filters for highly efficient filtering. In addition, an ultrasound digital transcranial Doppler system (digiTDS) [7] monitors intracranial vessels and update the structure profile in the continuous time domain. Various FPGA-based filters in this system, including the CIC and FIR filters, serve different purposes at the expense of varying degrees of computation resources. These issues motivate us to examine several different digital interpolating filters to compare ideal (theoretical) performance and actual measurements. More specifically, different approaches for the interpolation (upsampling) of signals covering the human audible range have been implemented. Although applicable to wearable medical devices in its own right [26]–[29], the multirate filters used here also serve as a prototypical example for the general study of comparing signal processing performance and digital filter power consumption.

This work considers the power consumption and filter



Fig. 2. Block diagram for our measurement setup. See text for more details.

performance for a range of filter designs, including First Order Hold (FOH), Finite Impulse Response (FIR), and Cascaded Integrator-Comb (CIC). Mathematical analysis of the various interpolation schemes is provided and our empirical results are in good agreement with the analytical solutions. The power and FPGA resource utilization for each interpolation scheme is detailed. The measured power consumptions were also in great agreement with the values predicted by the simulation tools. Our designs outlined in the block diagram in Fig. 2 along with the experimental setup described in Section III-A provides a framework for graduate students and researchers to accurately test signal processing theories. For this purpose, we have also uploaded our HDL files with detailed documentation to an online<sup>1</sup> repository.

## II. THEORY

In the following sub-sections, we cover the theory behind several different types of interpolation filters.

### A. Interpolation Overview

The well-known sampling theorem states that a continuous band-limited signal can be perfectly reconstructed from its samples, given a sufficiently high sampling rate. This is accomplished by passing the samples through an ideal low-pass filter; the frequency domain transfer function of which has an infinitely sharp cut-off characteristic as in (1).

$$H(f) = \text{rect}\left(\frac{f}{f_s}\right) \quad (1)$$

Here,  $f_s$  is the sampling frequency and  $\text{rect}(x)$  is a unit function that becomes 1 when the absolute value of  $x$  is less than  $1/2$  and 0 when the absolute value of  $x$  is greater than  $1/2$ . The corresponding time domain impulse response is a sinc function as in (2). Thus, convolving a sampled signal with (2) reconstructs the original continuous signal.

$$h(t) = \text{sinc}\left(\frac{t}{T}\right) \quad (2)$$

where  $T$  is the sampling period.

In order to change the sampling rate of a signal digitally, one must evaluate the result of the convolution of a sampled signal with (2) at the new sampling interval [30]. Of course, (2) cannot be evaluated perfectly, on account of its infinite duration. Thus, the problem of sample rate conversion - or upsampling/interpolation as is the primary concern here - becomes a problem of approximating (1) through filter design. Fig. 1 illustrates how interpolation can be used to reduce the performance requirements of a post-DAC analog reconstruction filter - spacing out the sampling images allows for a looser approximation of (1) to be used as the reconstruction filter.

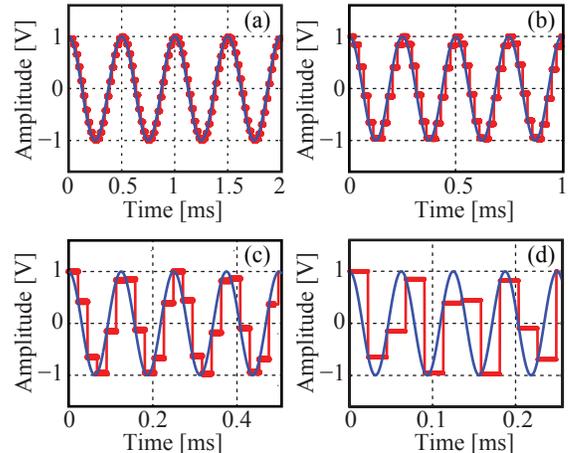


Fig. 3. Various sinusoidal waves at (a) 2 KHz, (b) 4 KHz, (c) 8 KHz, and (d) 16 KHz. The blue curves are the original signals and the red curves represent signals reconstructed via zero-order hold (ZOH) from the sampling at 44.1 KHz of the original signals.

### B. Zero-order Hold (ZOH)

It is common for digital-to-analog converters (DAC) to inherently perform a zero-order hold (ZOH) operation on sampled inputs. That is, DAC's typically hold an analog equivalent output of the previous digital input until a new value is presented. Thus, no digital signal processing is required for this simple interpolation scheme. The transfer function of the ZOH operation has the form of a sinc function. This can be shown by considering a ZOH impulse response, as in (3).

$$h_0(t) = u(t) - u(t - T). \quad (3)$$

Here,  $u$  is the step function and  $T$  represents the sampling period of the discrete signal. The Fourier transform is thus:

$$H_0(f) = e^{-j\pi f T} \text{sinc}(fT) \quad (4)$$

Comparing the frequency domain transfer functions, one finds the sinc function of (4) only loosely approximates the rect function, (1), needed for perfect reconstruction.

A more detailed look at the ZOH operation can be found by considering the case of a sinusoidal input. As shown in Fig. 3, the ZOH scheme is characterized by a stair-case like profile, which becomes more evident as the frequency of the signal increases. Using MATLAB, we evaluated (4) for a sine wave with  $f \simeq 5$  KHz. In Fig. 4(a), we show the sampled signal reconstructed via ZOH scheme. In Fig. 4(b), we show the fast Fourier transform (FFT) of the reconstructed signal, and verified that the frequency spectrum of the reconstructed signal matches well with the coefficients obtained from evaluation of (4).

<sup>1</sup>[https://github.com/donnnelly92/FPGA\\_Interpolation](https://github.com/donnnelly92/FPGA_Interpolation)

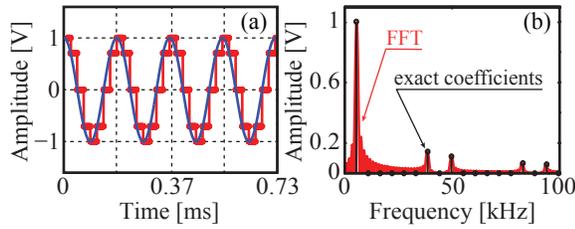


Fig. 4. The original signal (blue) is a 5.5 KHz cosine wave. (a) The reconstructed signal (red) and (b) the FFT spectral representation.

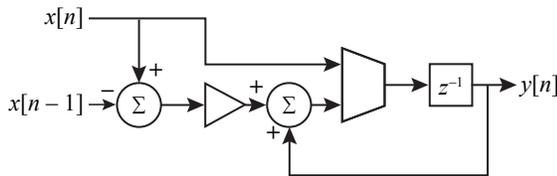


Fig. 5. The block diagram for the linear interpolation scheme. Its structure is relatively simple due to the lack of multiplication operations.

### C. Linear Interpolation

Linear interpolation, or first order hold (FOH), is a familiar operation, even outside the realm of digital signal processing. The impulse response of a FOH filter is a triangle (or tent) function [30].

$$h_1(t) = \frac{1}{T} \text{tri} \left( \frac{t}{T} \right) \quad (5)$$

The function  $\text{tri}(t)$  is defined as when  $t$  is zero, the function becomes 1. As  $t$  increases (decreases), the function slopes towards the  $t$  axis with unity slope until they intersect at  $\pm 1$ . The corresponding transfer function is thus of the form sinc squared as shown in (6).

$$H_1(f) = \text{sinc}^2(fT) \quad (6)$$

While  $h_1(t)$  as a continuous impulse response would yield perfect linear interpolation, a digital interpolating filter must operate with discrete samples. This is accomplished by first sampling the impulse response  $h_1(t)$  at a rate corresponding to the up-sample factor, then using this now discrete impulse response to convolve a ‘zero-stuffed’ (formally, upsampled) input signal. This is the same method behind any general finite impulse response (FIR) upsampling filter, which we shall discuss later in more detail. Unlike FIR interpolation, which requires multiplication and coefficient storage (see Section II-D), the linear interpolation scheme requires far less computational resources and its architecture is depicted in the block diagram in Fig. 5. Recall that the DAC inherently applies a zero-order-hold between samples, so the discrete linearly interpolated signal is converted into a continuous signal through this operation.

Fig. 6 shows a few examples of signals upsampled 16 times via linear interpolation. With the original signal at relatively low frequencies, 2 KHz and 4 KHz in Fig. 6(a) and (b), respectively, the interpolated signal matches the original reasonably well. At higher frequencies, 8 KHz and 16 KHz in Fig. 6(c) and (d), respectively, there is a noticeable mismatch between the interpolated and original signals. In Fig. 7(a),

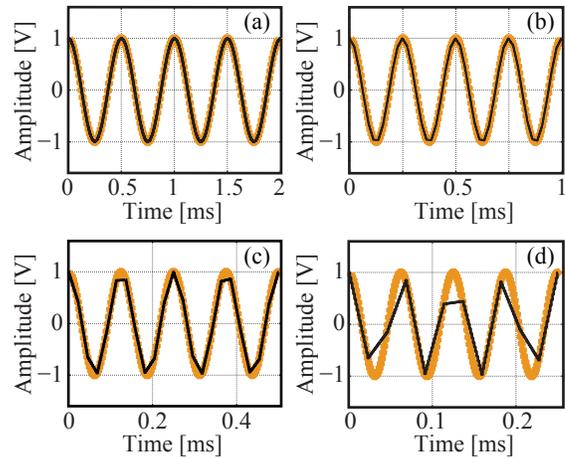


Fig. 6. Various sinusoidal waves at (a) 2 KHz, (b) 4 KHz, (c) 8 KHz, and (d) 16 KHz. The (orange) markers represent the original signals and the black curves represent signals reconstructed via linear interpolation at a 16 time upsampling frequency (i.e.  $16 \times 44.1$  KHz).

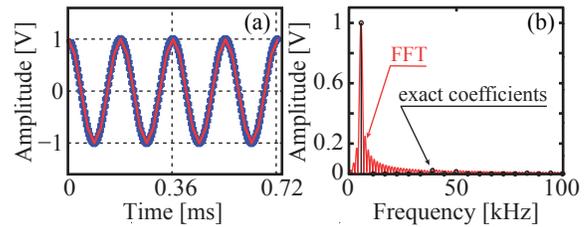


Fig. 7. The original signal (blue) is a 5.5 KHz cosine wave. (a) The interpolated signal (red) obtained via linear interpolation with a upsampling factor of 16 (b) The FFT results of the sampled signal and the Fourier coefficients for the Fourier expansion of the sampled signal.

we plotted linearly interpolated signal using a sine wave of frequency ( $f \simeq 5$  KHz). As with ZOH, the FFT routine was performed and is shown in Fig. 7(b). We validate the spectrum with the exact Fourier coefficients obtained by evaluating (6).

### D. Finite Impulse Response (FIR) Interpolation

The structure of the finite impulse response scheme is shown in Fig. 8. Here,  $x[n]$  represents data samples, which are strung together via unit delays so as to create a sequence of previous values. These samples are then multiplied by coefficients and summed. The coefficients  $b_0, b_1, \dots, b_n$  for a low-pass filter are samples of a sinc function. As mentioned, since (2) is infinite, for it to be used as a basis for the FIR filter coefficients it must first be truncated. One method to truncate it is to apply a window such that it has nonzero only in a certain range, as in (7) and (8) for which  $n \leq N$  and  $N$  is the FIR filter length and corresponds to the number of the  $b_n$  coefficients.

$$h[n] = h_{\text{sinc}}(nT)w[n] = \frac{\sin(\omega_c(n - \frac{N-1}{2}))}{\pi(n - \frac{N-1}{2})}, \quad (7)$$

$$H(\omega) = H_{\text{sinc}}(\omega) * W(\omega) \quad (8)$$

Here,  $w[n]$  represents a rectangular window that truncates the *sinc* function to a length of  $N$  and  $W(\omega)$  is the Fourier Transform of said window:  $|W(\omega)| = \frac{|\sin(\omega N/2)|}{|\sin(\omega/2)|}$ . Other

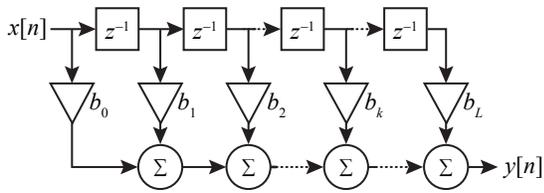


Fig. 8. The block diagram for the finite impulse response (FIR) scheme.

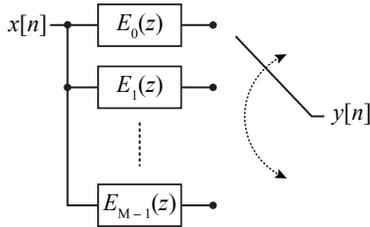


Fig. 9. The block diagram for the polyphase multi-rate signal processing.

truncating windows, such as Hamming, Hanning, or Blackman could alternatively be used for a trade-off between transition width and stop-band ripple (peak sidelobe). The rectangular window has the sharpest transition band - but considerable stopband ripple; whereas the other windows mentioned have wider main lobes in exchange for reduced peak sidelobes.

An alternative to applying a window as a means of truncation is to derive optimal coefficients constrained to a given number of terms [31]. That is, for a given filter length, coefficients can be determined such that the approximation error between the desired frequency response and the actual frequency response is minimized in a least squared error (LSE) fashion. This particular method will be referred to as an LSE FIR scheme.

An interpolation filter can be optimized by using a polyphase structure where sub-filters are multiplexed together. Instead of actually performing upsampling followed by filter-

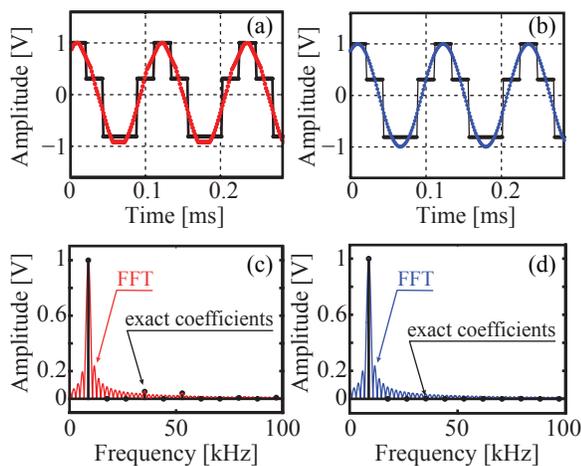


Fig. 10. 9 KHz sine wave sampled at 44.1 KHz with (a) ZOH reconstruction (black), FIR upsampling with the filter length  $L = 21$  and (c) its spectral analysis and (b) ZOH reconstruction (black), FIR upsampling with the filter length  $L = 201$  with (d) its spectral analysis. Both cases have the upsampling factor  $v = 16$ .

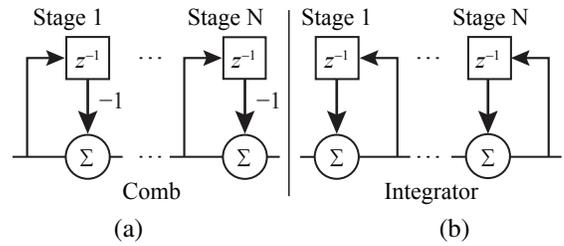


Fig. 11. Block Diagram for (a) N integrator stages and (b) N comb stages.

ing, since many of the samples are known to be zero (and thus multiplication by a coefficient is also zero and would not contribute to the summation tree), the FIR filter can be divided into sub-filters as in Fig. 9. Fig. 10(a) and (c) show, respectively, the time domain interpolation and frequency domain match between the FFT and analytical coefficients (obtained by evaluating (8) with a sine input) for a FIR filter with  $N = 21$ . Fig. 10(b) and (d) shows the same for a FIR filter with  $N = 201$ .

### E. Cascaded Integrator-Comb (CIC)

An efficient interpolation low-pass filter due to lack of multiplication and coefficient storage can be constructed with a Cascaded Integrator Comb (CIC) Filter [32]. In this section, we apply the  $Z$ -transform, where discrete samples are converted to the  $z$ -domain. A comb has a transfer function of the form

$$H_{\text{comb}}(z) = 1 - z^{-RM}. \quad (9)$$

The CIC filter integrator stage has a transfer function that can be defined as:

$$H_{\text{integ}}(z) = \frac{1}{1 - z^{-1}} \quad (10)$$

where  $M$  is the differential delay per comb stage ( $M = 1$  is used throughout) and  $R$  is the rate change factor. For an interpolating filter, the comb stage at  $f_s$  are followed by the integrator at a higher frequency  $R \times f_s$ .  $N$  sequential integrator and comb stages can be represented as

$$H_{\text{CIC}}(z) = H_{\text{comb}}(z)H_{\text{integ}}(z) = \frac{(1 - z^{-RM})^N}{(1 - z^{-1})^N}. \quad (11)$$

## III. RESULTS

In this section, we first compare the transfer function measurements with the analytical expressions. We then discuss the power and FPGA resource utilization for each interpolation scheme.

### A. Filter Measurements

Our measurement setup is represented in the block diagram in Fig. 2. To measure the transfer functions, we generated uniformly distributed white noise from the desktop computer and obtained the frequency spectrum of the interpolated white noise using the FFT capability of the oscilloscope. A printed circuit board (PCB) was produced that incorporated the Texas Instrument PCM 2706 and PCM 1702 ICs The USB interface chip (PCM 2706) connects to a desktop computer, which is responsible for streaming the audio data, and produces a digital

data output in the form of  $I^2S$ . An FPGA development board (featuring the Xilinx 7 series Artix FPGA) receives this  $I^2S$  output and performs the signal processing (interpolation). The signals are then sent to the digital-to-analog converter (PCM 1702) output of which (after being passed through a trans-impedance amplifier) is measured with a Keysight 3000 X-series oscilloscope.

The transfer functions shown in Fig. 12 illustrate how digital interpolation can be used to reduce the performance requirements of analog reconstruction filters. If no upsampling is performed, as in Fig. 12(a), the analog reconstruction filter must compensate for attenuation in the passband as well as have a steep transition band to remove sampling images. On the other hand, when digital interpolation is used the analog reconstruction filter is only responsible for removing sampling images that appear far away from the passband, centered around the new sampling frequency.

For the FOH scheme in Fig. 12(b), while it has a steeper transition band compared to the ZOH scheme, it too suffers from attenuation in the passband and in general does a poor job of removing the sampling image near the cutoff frequency. However, the computational cost can be quite small.

Fig. 12(c) shows the transfer function of an FIR scheme that used a rectangular window to truncate the ideal low pass filter impulse response. As expected, the transition band appears rather steep but ripples in both the passband and stop band are prominent. In Fig. 12(d) a more preferred transfer function was observed using the least squared error FIR scheme discussed in Section. II-D. In this case, the same computational resources were required as in the rectangularly windowed FIR scheme (i.e., the two schemes have the same filter length). The measured transfer function is rather noisy in the stopband region. After analysis of the setup, we concluded that the noise floor of the DAC on the PCB contributes significantly to the deviation between the measurement and the theoretical yield. This is evident when the filter length is reduced to  $L = 64$  in Fig. 12(e). The noise floor no longer appears to affect the interpolation performance, thus producing a transfer function in good agreement with the analytical solution.

Lastly, Fig. 12(f) shows the measured transfer function of the CIC interpolation scheme. The three stage, 16x upsampling CIC filter ( $R = 16$ ,  $N = 3$ ) used here has the benefit of relatively small computational cost (see III-C). However, it suffers from attenuation in the passband, much like as was the case in the ZOH and FOH interpolation schemes, as well as a slow transition band such that low frequency sampling image content is not fully eliminated.

### B. FPGA Implementation

For the FOH scheme a single 16 bit full adder was synthesized, along with three 16 bit registers to store intermediate values. No explicit multiplication resources were required, since the scaling block shown in Fig. 5 was implemented by bit-shifting (which was possible in this case because the upsampling factor was a power of two).

In the polyphase FIR implementations, Fig. 8 and 9, multiplication blocks needed to be synthesized. In order to

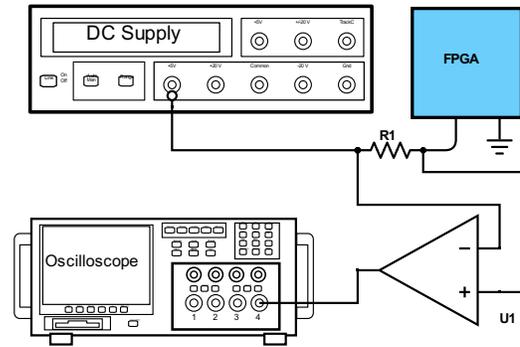


Fig. 13. The power measurement setup consists of a DC power supply, an oscilloscope and the resistor  $R1 = 20 \text{ m}\Omega$ .

meet timing constraints, pipeline registers were also used. It should be noted that the number of multiplications needed to implement the polyphase scheme is  $L/R$ , where  $L$  is the filter order (length) and  $R$  is the interpolation factor and thus 4 multipliers are needed in the case for  $L = 64$ ,  $R = 16$  while 20 multiplications are needed for the case  $L = 200$ ,  $R = 10$ . Longer filter lengths require a more-than-proportional length adder tree. The adder tree and coefficient storage detail are not shown explicitly in Fig. 8, but a similar implementation was used to what is detailed in [33].

Lastly, the CIC scheme uses six 28 bit adders, three for addition and three for subtraction, and 12 similar width registers. Six of these registers are for the storing previous values for the integrator and comb stages as in Fig. 11, and the remaining six are for pipelining each arithmetic operation (not shown in figure). The adders and registers are wider than the input data (16 bits), to account for bit growth as described in [32].

### C. FPGA Resource Usage

In this section, we discuss FPGA resource requirements for each interpolation scheme. The FPGA resources consumed by the interpolation filter designs are of two types: lookup tables (LUT) and flip-flops. Taken together, the power consumed by the FPGA was both predicted and measured. Although total power consumption, including other FPGA resources and external analog circuitry, is an important aspect of signal processing, we omit this measure from the results since all interpolation schemes are equivalent in this regard. Fig. 13 shows the schematic diagram for power measurements. We first isolated all peripherals that may consume power and therefore our measurement is due only to the digital filters in the FPGA. The voltage across a current sensing resistor ( $\sim 20 \text{ m}\Omega$ ) was monitored using the oscilloscope. Since one interpolation filter consumes very little power and its measurement would therefore be sensitive to noise, we implemented a number of filters in parallel inside the FPGA and derived an average power for a single filter from this measurement. We set the clocking frequency and other parameters of the FPGA constant for all filters so that differences in power consumption would be purely due to the complexity of the filters. We used the Xilinx software package Vivado to estimate circuit current and

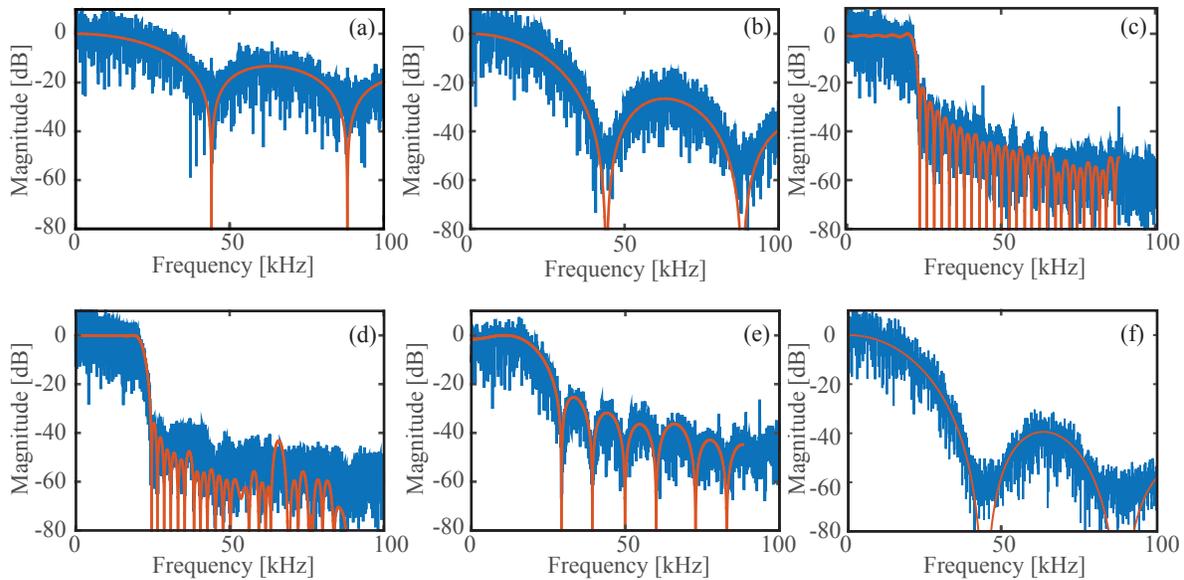


Fig. 12. Measurements (blue) and analytical (red) expressions for the transfer functions. (a) zero-order hold (ZOH), (b) first-order hold (FOH), (c) rectangularly windowed FIR ( $L = 200$ ), (d) least square error (LSE) FIR ( $L = 200$ ), (e) least square error (LSE) FIR ( $L = 64$ ), and (f) CIC ( $R = 16, N = 3$ ).

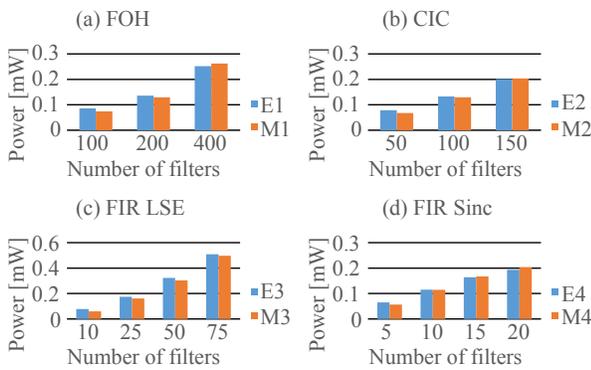


Fig. 14. The estimated (orange) and measured (blue) current for each interpolation scheme. The horizontal axis is the number of filters that are run in parallel in the FPGA.

our measured data shows good agreement with this estimation. Fig. 14 shows the predicted values of FPGA current draw compared to the measurements.

Table I summarizes the FPGA resource requirements for each interpolation scheme. Due to the varying complexity of each scheme, some filters draw considerably more power than others. In the case of the ZOH scheme, all logic LUTs and flip-flops are used to perform data flow control from the input source ( $I^2S$  from PCM 2706) to the DAC (PCM 1702). Since resources utilized in this way do not contribute to the signal processing operation, they are not included in the power measurements. All other interpolation schemes require additional resources on top of those required by the ZOH scheme; the FPGA resources and power consumption results listed in Table I only account for these additional resources.

TABLE I. FPGA LOGIC AND POWER UTILIZATION

	LUT	Flip-flops	$P_E$ [mW]	$P_M$ [mW]
ZOH	101	73	N/A	N/A
FOH	71	88	0.66	0.63
CIC	267	315	1.36	1.34
FIR ( $L=64$ )	761	293	6.64	6.79
FIR ( $L=200$ )	2670	1336	10.18	9.65

$P_E$  represents the estimated power consumption and  $P_M$  represents the actual measurements.

#### IV. CONCLUSION AND FUTURE WORK

We have studied various digital signal interpolation schemes using an FPGA. Our measurements were in good agreement with analytical expressions, although noise effects contributed to a slight mismatch for some cases. The power measurement confirms the power consumption rate predicted by the simulation tools. We emphasize that studies of power consumption and filter performance across a wide range of FPGA-based devices and applications, are crucial. This paper presents implementation details of various filtering schemes to test digital filter theory in a compact system for graduate students and researchers in the digital signal processing fields. In addition, our work covers some basic hardware-side contributions to power consumption in the digital filtering schemes. In the future, we plan to incorporate more hardware-oriented FPGA performance into the filter designs. This will allow us to explore optimization of filter quality while minimizing power consumption and other resources. Specifically, we will study applications with portable digital signal processing units such as wearable medical devices. Such devices may often hamper patients' daily activity due to the physical dimensions and available battery life. More in-depth work on the hardware contributions to the FPGA-based filter performance will complement our current software-side evaluation in designing the optimization scheme for wearable medical devices.

#### ACKNOWLEDGMENT

The authors would like to thank C. A. Bouman, N. Conrad, M. Y. Shalaginov, and D. S. Elliott for helpful advice and discussions. Also, technical help from Kyle Woodworth is acknowledged.

#### REFERENCES

- [1] Z. Szadkowski and D. Gas. Adaptive linear predictor fir filter based on the cyclone v fpga with hps to reduce narrow band rfi in radio detection of cosmic rays. *IEEE Transactions on Nuclear Science*, 63(3):1455–1462, June 2016.
- [2] G. E. Santagati and T. Melodia. Sonar inside your body: Prototyping ultrasonic intra-body sensor networks. In *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pages 2679–2687, April 2014.
- [3] R. P. S. Rathore. Reconfigurable digital radar receiver implemented in fpga using under-sampling, direct iq generation, multi-rate filter and pulse compression. In *2014 IEEE International Microwave and RF Conference (IMARC)*, pages 174–177, Dec 2014.
- [4] A. Gabiger-Rose, M. Kube, R. Weigel, and R. Rose. An fpga-based fully synchronized design of a bilateral filter for real-time image denoising. *IEEE Transactions on Industrial Electronics*, 61(8):4093–4104, Aug 2014.
- [5] F. E. Sagcan, A. Bayram, and B. Sen. Fpga implementation of delta sigma modulation based all-digital rf transmitter for parallel magnetic resonance imaging. In *2015 European Microwave Conference (EuMC)*, pages 494–497, Sept 2015.
- [6] K. Hu, X. Wang, Y. Yao, X. Gao, and G. Jin. Real-time data acquisition for single photon imaging detector. *IEEE Transactions on Nuclear Science*, 63(2):1076–1082, April 2016.
- [7] M. Lewandowska, M. Walczaka, P. Karwata, B. Witeka, A. Nowickia, and P. Kar lowiczb. Research and medical doppler platform. In *Acoustics 2012 Nantes Conference*, pages 853–858, 2012.
- [8] N. M. Ghamry. An FPGA implementation of hearing aids based on wavelet-packets. *Journal of Computers*, 7(3), 2012.
- [9] A. R. Buzdar, A. Latif, and L. Sun. FPGA prototype implementation of digital hearing aid from software to complete hardware design. *International Journal of Advanced Computer Science and Applications*, 7(1), 2016.
- [10] V. Hanson and K. Odame. Real-time embedded implementation of the binary mask algorithm for hearing prosthetics. *IEEE Transactions on Biomedical Circuits and Systems*, 8(4):465–473, Aug 2014.
- [11] S. G. Kambalimath, P. C. Pandey, P. N. Kulkarni, S. S. Mahant-Shetti, and S. G. Hiremath. Fpga-based design of a hearing aid with frequency response selection through audio input. In *2016 29th International Conference on VLSI Design and 2016 15th International Conference on Embedded Systems (VLSID)*, pages 579–580, Jan 2016.
- [12] L. Bendaouia, H. Salhi, S. M. Karabernou, L. Kessal, and F. Ykhlef. Fpga-implementation of a bio-inspired medical hearing aid based dwt-ola. In *2014 International Conference on Audio, Language and Image Processing*, pages 806–811, July 2014.
- [13] Geng Yang, Jian Chen, Ying Cao, Hannu Tenhunen, and Li-Rong Zheng. A novel wearable eeg monitoring system based on active-cable and intelligent electrodes. In *HealthCom 2008 - 10th International Conference on e-health Networking, Applications and Services*, pages 156–159, July 2008.
- [14] Y. Hong, I. Rajendran, and Y. Lian. A new eeg signal processing scheme for low-power wearable eeg devices. In *2011 Asia Pacific Conference on Postgraduate Research in Microelectronics Electronics*, pages 74–77, Oct 2011.
- [15] M. Knieser A. Hayim and M. Rizkalla. Dsp/fpgas comparative study for power consumption, noise cancellation, and real time high speed applications. *Journal of Software Engineering and Applications*, 3(4):391–403, 2010.
- [16] G. C. Cardarilli, A. Del Re, A. Nannarelli, and M. Re. Power characterization of digital filters implemented on fpga. In *2002 IEEE International Symposium on Circuits and Systems. Proceedings (Cat. No.02CH37353)*, volume 5, pages V–801–V–804 vol.5, 2002.
- [17] J. P. Oliver, J. Curto, D. Bouvier, M. Ramos, and E. Boemo. Clock gating and clock enable for fpga power reduction. In *2012 VIII Southern Conference on Programmable Logic*, pages 1–5, March 2012.
- [18] K. I. Kum and W. Sung. Word-length optimization for high-level synthesis of digital signal processing systems. In *1998 IEEE Workshop on Signal Processing Systems. SIPS 98. Design and Implementation (Cat. No.98TH8374)*, pages 569–578, Oct 1998.
- [19] R. Jevtic and C. Carreras. Power measurement methodology for fpga devices. *IEEE Transactions on Instrumentation and Measurement*, 60(1):237–247, Jan 2011.
- [20] A. J. Casson and E. Rodriguez-Villegas. Toward online data reduction for portable electroencephalography systems in epilepsy. *IEEE Transactions on Biomedical Engineering*, 56(12):2816–2825, Dec 2009.
- [21] T. J. Chen, C. Jeng, S. T. Chang, H. Chiueh, S. F. Liang, Y. C. Hsu, and T. C. Chien. A hardware implementation of real-time epileptic seizure detector on fpga. In *2011 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 25–28, Nov 2011.
- [22] A. Page, C. Sagedy, E. Smith, N. Attaran, T. Oates, and T. Mohsenin. A flexible multichannel eeg feature extractor and classifier for seizure detection. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 62(2):109–113, Feb 2015.
- [23] C. I. Jeong, P. I. Mak, C. P. Lam, C. Dong, M. I. Vai, P. U. Mak, S. H. Pun, F. Wan, and R. P. Martins. A 0.83- $\mu$  W QRS detection processor using quadratic spline wavelet transform for wireless eeg acquisition in 0.35- $\mu$ m cmos. *IEEE Transactions on Biomedical Circuits and Systems*, 6(6):586–595, Dec 2012.
- [24] C. F. Zhang and T. W. Bae. Vlsi friendly eeg qrs complex detector for body sensor networks. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2(1):52–59, March 2012.
- [25] Addanki Purma Ramesh, G.Nagarjuna, and G.Siva Raam. Fpga based design and implementation of higher order fir filter using improved da algorithm. *International Journal of Computer Applications*, 35(9):0975–8887, 2011.
- [26] A. Tobola, F. J. Streit, C. Espig, O. Korpok, C. Sauter, N. Lang, B. Schmitz, C. Hofmann, M. Struck, C. Weigand, H. Leutheuser, B. M. Eskofier, and G. Fischer. Sampling rate impact on energy consumption of biomedical signal processing systems. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 1–6, June 2015.
- [27] A. Tobola, C. Espig, F. J. Streit, O. Korpok, B. Schmitz, C. Hofmann, M. Struck, C. Weigand, H. Leutheuser, B. M. Eskofier, and G. Fischer. Scalable ECG hardware and algorithms for extended runtime of wearable sensors. In *Medical Measurements and Applications (MeMeA), 2015 IEEE International Symposium on*, pages 255–260, May 2015.
- [28] R. Brennan and T. Schneider. An ultra low-power DSP system with a flexible filterbank. In *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, volume 1, pages 809–813 vol.1, Nov 2001.
- [29] C. Qian, J. Shi, J. Parramon, and E. Snchez-Sinencio. A low-power configurable neural recording system for epileptic seizure detection. *IEEE Transactions on Biomedical Circuits and Systems*, 7(4):499–512, Aug 2013.
- [30] R. W. Schafer and L. R. Rabiner. A digital signal processing approach to interpolation. *Proceedings of the IEEE*, 61(6):692–702, June 1973.
- [31] T. Parks and J. McClellan. A program for the design of linear phase finite impulse response digital filters. *IEEE Transactions on Audio and Electroacoustics*, 20(3):195–199, Aug 1972.
- [32] E. Hogenauer. An economical class of digital filters for decimation and interpolation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(2):155–162, Apr 1981.
- [33] C. Chou, S. Mohanakrishnan, and J. B. Evans. FPGA implementation of digital filters. In *In Proceedings of the Fourth International Conference on Signal Processing Applications and Technology*, pages 80–88, 1993.

# A Survey of Datasets for Biomedical Question Answering Systems

Muhammad Wasim, Dr. Waqar Mahmood, Dr. Usman Ghani Khan Al-Khawarizmi Institute of Computer Science  
University of Engineering and Technology, Lahore

**Abstract**—The massively ever increasing amount of textual and linked biomedical data available online poses many challenges for information seekers. So, the focus of information retrieval community has shifted to precise information retrieval, i.e. providing exact answer to a user question. In recent years, many datasets related to Biomedical Question Answering (BioQA) have emerged which the researchers can use to evaluate the performance of their systems. We reviewed these biomedical datasets and analyzed their characteristics. The survey in this paper covers these datasets for BioQA and has a two fold purpose: to provide an overview of the available datasets in this domain and to help researchers select the most suitable dataset for benchmarking their system.

**Keywords**—Biomedical; QA system; review; survey

## I. INTRODUCTION

The massive amount of textual data on web makes finding information a challenging task. Although *Information Retrieval* (IR) systems are developed to cater this problem, such systems just provide a list of documents instead of precise information [1]. The information seeker thus needs to process the provided document list to filter the required information. To overcome this problem, the research and development in the area of question answering is in progress [2]–[4]. To accelerate research in this area, many research initiatives including Text Retrieval Conference (TREC)<sup>1</sup> have been taken to develop benchmark datasets for open domain question answering [5]–[12]. Early research focused on heuristic based methods such as patterns for specific question types or redundancy of data over web to answer questions mostly from textual documents [13], [14]. The current focus in question answering systems is to use statistical learning approaches and neural networks to answer questions [15]–[19] both in textual and linked data. Moreover, the trend in question answering is changing from open domain to restricted domains such as biomedical [20], [21]. Focusing on restricted domains makes finding solution to specific answer types easy as different patterns can be used to determine question and answer types and domain knowledge can be incorporated [22].

The biomedical data on web can be broadly divided into two main categories i.e. textual data and linked data [23]. Textual data includes scientific literature published by biomedical journals and may include other authentic websites on biomedical domain. Linked data provides information about how different biomedical entities are related to each other and provides mechanism to inference and derive new knowledge.

Linked biomedical data is one of the most important types of data in linked data world as around one-tenth of all linked data available online is biomedical data [10]. It may include drugs, compounds, diseases, genomics, proteomics or other nomenclature/lexical ontologies. Fig. 1 depicts the general categorization of available biomedical data. Biomedical experts usually publish their research both in the form of research articles and submit their experimental results to linked data repositories. To validate their hypothesis during any further experimentation, biomedical experts need to consult both textual and linked data sources manually. Biomedical question answering systems should be able to derive such answer automatically by combining evidence from both textual and linked data sources.

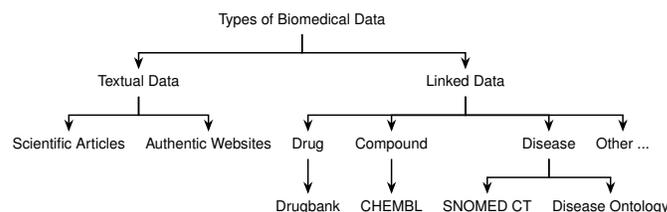


Fig. 1. Information sources in biomedical domain.

To tap the real potential of this heterogeneous data, techniques need to be developed and biomedical question answering datasets play an important role in this process as researchers can use it to evaluate the performance of their systems. Question answering datasets usually contain two main components: 1) *Information Need* or question in natural language and 2) the *corpus or linked data* which will help in answering the questions to fulfill some information need. Questions may also have additional information such as question types such as passage, factoid, list or summary. The focus of *passage* type question is to return relevant passages. The *factoid* type questions deal with single entity such as genes, proteins, disease, symptoms etc. List questions typically return list of factoids. The answer of *yes/no* type question is either yes or no. Finally, the aim of the *summary* question is to provide answer in a summarized form and may include definitional questions. Such questions typically start with *What is* phrase. There are also datasets available where question type is *multiple choice question*. The focus of such type of questions is to improve the passage comprehension and inferencing capabilities of question answering systems. Question types may help a QA system to focus on one particular answer

<sup>1</sup><http://trec.nist.gov/data/qa.html>

strategy [24] and following are some typical questions asked in biomedical domain:

- Passage Question
- Factoid Question
- List Question
- Multiple Choice Question
- Yes/No Question
- Summary Question

Biomedical experts have certain information needs. They might be interested in knowing how genes interact with organ functions or what role some genes play in a particular disease. They might be further interested in *proteins, protein-protein interactions, mutations, drugs, adverse effects, cell or tissue types and signs or symptoms*. Other entities of interest might be *chemicals, species, pathways, genetic variations, and patient characteristics*. Heterogeneous biomedical sources are normally required to fulfill an information need as shown in Fig. 2. Identifying them early in the question answering pipeline helps narrow down candidate answers in later stages [24].

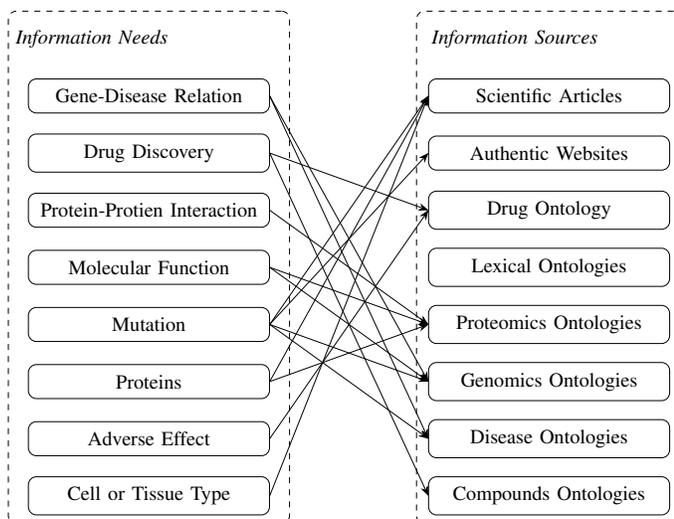


Fig. 2. Mapping biomedical information needs to sources.

This paper is organized as follows. Section II describes different biomedical question answering datasets and their specific characteristics. A comparative study of such datasets is presented in the next section and finally the paper is concluded in the last section.

## II. DATASETS FOR BIOMEDICAL QUESTION ANSWERING

This section enlists the biomedical question answering datasets in chronological order.

### A. TREC Genomics Track

Genomics corpus provided by TREC was one of the pioneer datasets developed for Biomedical Question Answering. The dataset was prepared for TREC Genomics track held

in 2006 and 2007 [20], [21]. The challenge was to retrieve relevant passages against a topic question from a corpus of 162,259 documents. The corpus was collected by crawling 49 biomedical journals and covered various biomedical categories. A set of 28 and 36 topic questions was assembled in 2006 and 2007 respectively. Some example questions from both collections are shown in Tables I and II. The biomedical entities covered in these two years included: *genes, proteins and gene mutations* in 2006; and *antibodies, cell or tissue type, disease, drug, gene, molecular function, mutations, pathways, proteins, symptoms, toxicities, and tumor types* in 2007.

TABLE I. TREC 2006 GENOMICS TRACK TOPIC QUESTIONS

Example Questions
How does L2 interact with L1 to form HPV11 viral capsid?
How does p53 affect apoptosis?
How do mutations in the Pes gene affect cell growth?

TABLE II. TREC 2007 GENOMICS TRACK TOPIC QUESTIONS

Questions
What serum [PROTEINS] change expression in association with high disease activity in lupus?
What [MUTATIONS] in the Raf gene are associated with cancer?
What [DRUGS] are associated with lysosomal abnormalities in the nervous system?

The dataset was focused on scientific articles only and limited number of questions were not sufficient to evaluate the performance of large-scale QA system. Moreover, no annotations were provided with the corpus or questions so machine learning algorithms could not be effectively used.

### B. Question Answering for Machine Reading Evaluation (QA4MRE): Biomedical Text about Alzheimer's Disease

QA4MRE<sup>2</sup> for biomedical data [25] differs from TREC datasets because the focus of the dataset is on passage comprehension and multiple answers are already provided with each question. Corpus is used as a *background knowledge* i.e. it is used to acquire knowledge instead of directly answering the question. Collection developed as a *background knowledge* for the task was named Alzheimer's Disease Literature Corpus (ADLC corpus) was collected from different sources: 66,222 pubmed abstracts were from *PubMed*, 8249 open access full articles from *pubmed central*, 1041 full text articles from *pubmed central* in HTML and text format. Moreover, 379 full text articles and 103 abstracts were added in this collection from a list of popular articles on Alzheimer disease. The test set is composed of four reading tests where each test consists of one document and 10 questions related to that document and five answer choices per question. Table III shows some sample questions from the collection. The systems had to either select answer from the given list of possible answers or leave the question unanswered. The background collection, test documents and questions were annotated with *word, lemma, chunk, part of speech, named entity, parent node in the dependency tree, dependency syntax label, UMLS entity name, and named entity*.

<sup>2</sup><http://celct.fbk.eu/QA4MRE/>

TABLE III. QA4MR ABOUT BIOMEDICAL TEXT ON ALZHEIMER'S DISEASE

Questions	Options
Which CLU isoform has a consistently higher gene expression?	CLU2 ribosomal protein L13A CLU1 allele PNGase
Which hormone can control the expression of CLU isoforms?	real-time PCR cDNA AD rs11136000 androgen
What effect do androgens have on CLU2 gene expression?	association repression inhibition activation expression

These annotations were performed automatically using state of the art tools such as *GDep parser*, *CLIPS NE Tagger*, and *ABNER tagger*. The named entities considered for the task were *genes* and *gene products*, *chemicals*, *drugs*, *symptoms*, *experimental methods*, *species*, *pathway*, *cellular*, *genetic variation*, *adverse effect*, *dose*, *timing*, *patient characteristics* etc. The challenge was more about checking the inferencing capabilities of the system and multiple choice question proposed a new dimension of research. Moreover background collection was also different in the perspective. There was no retrieval sub-system and linked data was not used as a source in the dataset.

### C. QALD-4: Biomedical Question Answering over Interlinked Data

Question Answering over linked data (QALD)<sup>3</sup> started in 2011 viewing the wide spread of linked data over the web and focused on converting information need into standard semantic query processing and inferencing. The objective of the dataset is to establish a standard against which question answering systems over structured data can be evaluated and compared. The fourth year competition had three tracks i.e. *multilingual question answering*, *biomedical question answering over inter-linked data*, and *hybrid question answering*. Biomedical data was selected as there are many structured datasets available in this domain and answer to information need can be satisfied only if evidence is combined from multiple sources. Three biomedical datasets were combined for this competition i.e. 1) *SIDER*, which describes drugs and it's side effects; 2) *diseasome*, which provides information about diseases and genetic disorders; and 3) *drugbank*, which provides FDA-approved active compounds of medication. There were a total of 25 training questions and 25 similar test questions. Some example questions from the dataset are shown in Table IV.

There are very limited number of test and training questions and the only target of dataset is on three ontology sources. The questions were prepared in a manner that multiple sources needed to be inquired to find answer to a question. The systems using this dataset should address the problem of converting natural language questions to SPARQL queries efficiently.

TABLE IV. SOME QUESTIONS FROM QALD-4 BIOMEDICAL DATASET

Questions
Which genes are associated with Endothelin receptor type B?
Which genes are associated with subtypes of rickets?
Which drug has the highest number of side-effects?
List drugs that lead to strokes and arthrosis.
Which drugs have a water solubility of 2.78e-01 mg/mL?

### D. BioASQ large-scale Biomedical Semantic Indexing and Question Answering

The focus of BioASQ challenge has been to assemble information from multiple heterogeneous sources in order to answer real-life biomedical experts' questions. The competition is being held every year since 2013. The challenge consists of two tasks 1) Biomedical semantic indexing; and 2) Semantic question answering; semantic answering is further divided into two tasks i.e. retrieval of relevant documents, snippets and triplets (Phase A) and finding the precise answer of the question (Phase B). The focus of the challenge is on drugs, targets, disease and covers both textual and linked data. The selected sources for these categories are shown in Table V.

TABLE V. THE RESOURCES THAT WERE INCLUDED IN BIOASQ CHALLENGE

Focus	Resources
Drugs	Joche
Targets	Gene Ontology, UniProt
Diseases	Disease Ontology
General Purpose	MeSH
Document Sources	PubMed, PubMed Central
Linked Data	LinkedLifeData

The challenge involves text/passage retrieval, RDF triplet retrieval, QA for exact answer, multi document summarization and natural language generation. The answer of any question may be factoid or passage depending on the type of question. The system could provide exact answer or ideal answer where ideal answer is paragraph-sized summary of the answer. The dataset contained development and test questions. There are four type of questions in dataset 1) Yes/No; 2) factoid; 3) list; and 4) summary. All questions expect exact and ideal answer except summary question where only ideal answer is expected. Table VI shows example questions for each type. The training and test questions for each year are: 29, 282 in 2013; 310, 500 in 2014; 810, 500 in 2015; 1307, 500 in 2016; and 1799, 500 in 2017 [26].

TABLE VI. QUESTIONS TYPES IN BIOASQ DATASET

Question Type	Example Question
Yes/No	Is miR-21 related to carcinogenesis?
Factoid	Which is the most common disease attributed to malfunction or absence of primary cilia?
List	Which human genes are more commonly related to craniosynostosis?
Summary	What is the mechanism of action of abiraterone?

## III. DISCUSSION

Biomedical data is available in both textual and linked formats. Therefore, the question answering datasets should also provide good coverage of both of these sources. Fig. 3

<sup>3</sup><http://qald.sebastianwalter.org/index.php?x=home&q=1>

shows the number of datasets which provide linked, textual and heterogeneous biomedical data. Only one dataset provides linked data source. Three datasets provide textual data out of which two use the same corpus with different question sets. The third textual dataset uses background collection as data source. The four datasets provided by BioASQ provide heterogeneous data. The dataset mostly targets open and publicly available textual and linked datasets and provides good number of training examples. Textual datasets can be searched using keywords based IR techniques. Linked data requires specific query format conversion supported by linked data repositories known as SPARQL. Heterogeneous datasets require both keyword based queries and linked data specific query formats. Moreover, to produce answer in natural language, answers from linked data are normally required to pass through natural language generation module. The datasets providing heterogeneous sources are most effective as both sources are equally important in the context of biomedical domain.

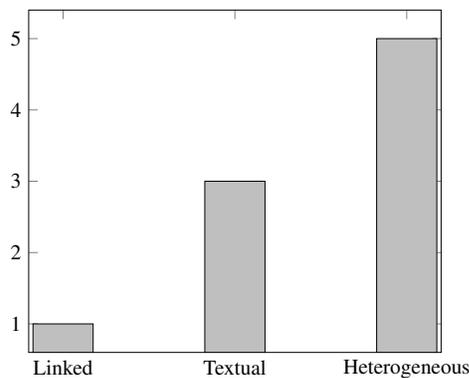


Fig. 3. Datasets with respect to type of data.

The number of questions provided with each dataset have increased over years. Initially, only 28 questions were provided in TREC Genomics track with no training set. Training question were first introduced in QALD-4 with 25 training questions. The trend from then has only increased and current year's BioASQ track contained 1307 training questions. This increase in information needs of real world biomedical experts' needs is essential to build practically usable QA systems. More training data may greatly aid in building machine learning based algorithms. Fig. 4 shows the number of training and test questions in all presented datasets.

Table VII shows a comparative table comparing all the datasets presented in this paper. Each dataset's strength and weakness are also mentioned to help researchers select appropriate dataset for their research. To summarize:

- If the QA system is to be built upon IR system, TREC dataset for 2006 and 2007 with a total of 64 questions provide a good starting point.
- To build and test a QA system which queries multiple linked data sources, QALD dataset can be used.
- State-of-the-art dataset is provided by the recent BioASQ challenge. The dataset can be used to evaluate:

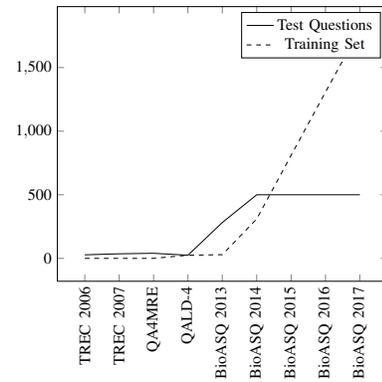


Fig. 4. Datasets - training vs test questions.

- Performance of system on heterogeneous data sources.
- Inferencing capabilities of a QA system.
- Ability to generate natural language answer.
- The quality of summarized answers.

#### IV. CONCLUSION

Question answering aims to provide a practical solution to the information overload problem. The availability of biomedical dataset highlights both the challenges and opportunities present in this domain. The availability of training data provides the opportunity to tailor systems in learning from examples. Moreover, the initiatives from BioASQ on heterogeneous dataset is paving the way towards better datasets to evaluate the effectiveness of biomedical QA systems on a larger scale. BioASQ provides the opportunity to exploit heterogeneous sources, perform inference, and produce summary for ideal answers. All the datasets have their strengths and weakness but overall BioASQ provides heterogeneous sources and more training data. In future, we shall investigate the systems (and their characteristics) which work best for each dataset.

#### REFERENCES

- [1] O. Ferret, B. Grau, M. Hurault-Plantet, G. Illouz, L. Monceaux, I. Robba, and A. Vilnat, "Finding an answer based on the recognition of the question focus." in *TREC*, 2001.
- [2] V. Lopez, V. Uren, M. Sabou, and E. Motta, "Is question answering fit for the semantic web?: a survey," *Semantic Web*, vol. 2, no. 2, pp. 125–155, 2011.
- [3] S. K. Dwivedi and V. Singh, "Research and reviews in question answering system," *Procedia Technology*, vol. 10, pp. 417–424, 2013.
- [4] G. Zayaraz *et al.*, "Concept relation extraction using naive bayes classifier for ontology-based question answering systems," *Journal of King Saud University-Computer and Information Sciences*, vol. 27, no. 1, pp. 13–24, 2015.
- [5] E. M. Voorhees, "The trec question answering track," *Natural Language Engineering*, vol. 7, no. 04, pp. 361–378, 2001.
- [6] N. A. Smith, M. Heilman, and R. Hwa, "Question generation as a competitive undergraduate course project," in *Proceedings of the NSF Workshop on the Question Generation Shared Task and Evaluation Challenge*, 2008.
- [7] P. Rajpurkar, J. Zhang, K. Lopyrev, and P. Liang, "Squad: 100,000+ questions for machine comprehension of text," *arXiv preprint arXiv:1606.05250*, 2016.
- [8] M. Iyyer, J. Boyd-Graber, L. Claudino, R. Socher, and H. Daumé III, "A neural network for factoid question answering over paragraphs," in *Empirical Methods in Natural Language Processing*, 2014.

- [9] J. Berant, A. Chou, R. Frostig, and P. Liang, "Semantic parsing on freebase from question-answer pairs." in *EMNLP*, vol. 2, no. 5, 2013, p. 6.
- [10] V. Lopez, C. Unger, P. Cimiano, and E. Motta, "Evaluating question answering over linked data," *Web Semantics Science Services And Agents On The World Wide Web*, vol. 21, pp. 3–13, 2013.
- [11] K. M. Hermann, T. Kočiský, E. Grefenstette, L. Espeholt, W. Kay, M. Suleyman, and P. Blunsom, "Teaching machines to read and comprehend," in *Advances in Neural Information Processing Systems (NIPS)*, 2015. [Online]. Available: <http://arxiv.org/abs/1506.03340>
- [12] Y. Yang, W.-t. Yih, and C. Meek, "Wikiqa: A challenge dataset for open-domain question answering," in *Proceedings of EMNLP*. Citeseer, 2015, pp. 2013–2018.
- [13] D. Ravichandran and E. Hovy, "Learning surface text patterns for a question answering system," in *Proceedings of the 40th annual meeting on association for computational linguistics*. Association for Computational Linguistics, 2002, pp. 41–47.
- [14] A. Ittycheriah, M. Franz, W.-J. Zhu, A. Ratnaparkhi, and R. J. Mammone, "Ibm's statistical question answering system." in *TREC*, 2000.
- [15] X. Yao and B. Van Durme, "Information extraction over structured data: Question answering with freebase." in *ACL (1)*. Citeseer, 2014, pp. 956–966.
- [16] A. Fader, L. S. Zettlemoyer, and O. Etzioni, "Paraphrase-driven learning for open question answering." in *ACL (1)*. Citeseer, 2013, pp. 1608–1618.
- [17] M. Iyyer, J. L. Boyd-Graber, L. M. B. Claudino, R. Socher, and H. Daumé III, "A neural network for factoid question answering over paragraphs." in *EMNLP*, 2014, pp. 633–644.
- [18] T. Khot, N. Balasubramanian, E. Gribkoff, A. Sabharwal, P. Clark, and O. Etzioni, "Markov logic networks for natural language question answering," *arXiv preprint arXiv:1507.03045*, 2015.
- [19] A. B. Abacha and P. Zweigenbaum, "Means: A medical question-answering system combining nlp techniques and semantic web technologies," *Information Processing & Management*, vol. 51, no. 5, pp. 570–594, 2015.
- [20] W. Hersh, A. Cohen, L. Ruslen, and P. Roberts, "Trec 2007 genomics track overview."
- [21] W. R. Hersh, A. M. Cohen, P. M. Roberts, and H. K. Rekapalli, "Trec 2006 genomics track overview." in *TREC*, 2006.
- [22] D. Mollá and J. L. Vicedo, "Question answering in restricted domains: An overview," *Computational Linguistics*, vol. 33, no. 1, pp. 41–61, 2007.
- [23] G. Balikas, A. Krithara, I. Partalas, and G. Paliouras, "Bioasq: a challenge on large-scale biomedical semantic indexing and question answering," in *Multimodal Retrieval in the Medical Domain*. Springer, 2015, pp. 26–39.
- [24] X. Li and D. Roth, "Learning question classifiers," in *Proceedings of the 19th international conference on Computational linguistics-Volume 1*. Association for Computational Linguistics, 2002, pp. 1–7.
- [25] R. Morante, M. Krallinger, A. Valencia, and W. Daelemans, "Machine reading of biomedical texts about alzheimers disease 1," 2012.
- [26] G. Wiese, D. Weissenborn, and M. Neves, "Neural domain adaptation for biomedical question answering," *arXiv preprint arXiv:1706.03610*, 2017.

TABLE VII. COMPARING BIOMEDICAL QUESTION ANSWERING DATASETS

Dataset	Questions (Train+Test)	Textual Data	Linked Data	Factoid	List	Paragraph/Summary	Yes/No	No. of Entities
TREC 2006 Genomics	28	✓	×	×	×	✓	×	4
TREC 2007 Genomics	36	✓	×	×	×	✓	×	10
QA4MRE about Alzheimer's disease	40	✓	×	✓	✓	×	×	11
QALD-4	25+25	×	✓	✓	×	×	×	3
BioASQ 2013	29+282	✓	✓	✓	✓	✓	✓	∞
BioASQ 2014	310+500	✓	✓	✓	✓	✓	✓	∞
BioASQ 2015	810+500	✓	✓	✓	✓	✓	✓	∞
BioASQ 2016	1307+500	✓	✓	✓	✓	✓	✓	∞
BioASQ 2017	1799+500	✓	✓	✓	✓	✓	✓	∞

# A Comprehensive Analysis on the Security Threats and their Countermeasures of IoT

Abdul Wahab Ahmed  
Department of Computer Science  
COMSATS Institute of Information Technology  
Islamabad, Pakistan

Mian Muhammad Ahmed  
Department of Computer Science  
COMSATS Institute of Information Technology  
Islamabad, Pakistan

Omaid Ahmad Khan  
Department of Computer Science  
COMSATS Institute of Information Technology  
Islamabad, Pakistan

Munam Ali Shah  
Department of Computer Science  
COMSATS Institute of Information Technology  
Islamabad, Pakistan

**Abstract**—Internet of Things referred as a pervasive network architecture which provides services to the physical world by processing and analyzing data. In this modern era Internet of Things has been shown much significance and rapidly developing by connecting heterogeneous devices with various technologies. By this way interconnectivity of large number of electronic devices connected with the IoT network leads the risk of security and confidentiality of data. This paper analyzes different security issues, their counter measures and discusses the future directions of security in IoT. Furthermore, this paper also discusses essential technologies of security like encryption in the scenario of IoT for the prevention of harmful threats in the light of latest research.

**Keywords**—Internet of things; security threats; countermeasures; privacy

## I. INTRODUCTION

The term Internet of things (IoT) was firstly used by Kevin Ashton in 1999 [1] in the term of supply chain management but now it is used in a general perspective. We only do not get information from the internet but it follows the protocols the internet use to store information. It is estimates that in 2020 there will be 50 billion smart objects and devices as shown in Figure 2 so each person will have 6.6 physical devices which are very large in number [2]. Due to use of modern technology like RFID and Greenhouse monitoring etc a very rapid development arises in IOT but there is issue regarding privacy and security in different layers. Figure 1 gives an overview of IoT with its connectivity.

The enhancements of wireless sensor network are widely popular by some of its prospects and new discoveries. IOT refers to the communication between physical devices like smart phones and some other smart objects that exchange data and give useful services via internet [3]. Some applications (For Example: Greenhouse monitoring, Smart meter and grids) evolve much popularity through IoT. Mainly IoT is generated by some of its important components that contain sensing, varied access, processing of information (RFID, GPS etc) and some other components like its security.

The objective of IoT is to make interconnection between machines. Thus IoT surrounds and connects the real world through these physical devices which are embedded with different types of sensors [4]. The word “things” in IoT cover a wide range of physical objects and also includes several electronic devices including RFIDs, GPS and NFC etc.

The security of essential information on IoT should incorporate into different features such as identification, data privacy and confidentiality etc. So with the rapid development and a mixture of heterogeneous devices, it formulates very large scale of IoT infrastructure [5]. So it is predicted that IoT is feared to be under threats on its versatile technology and future capabilities [6]. The security threat to IoT such as Denial of Service, Brute Force, Man in the middle attacks and many other attacks are envisaged in the interconnected network. These attacks occurs because of weak password, no encryption, personal information leakage etc so storage of such confidential data on cloud in quite alarming.

If such security attacks are not solved to some safe level then this weak security services can be harmful for the market of IoT. It not only involves such security issues but also have some access control issues, authentication of various network and some information store problems [7]. This problem needs to have a well defined security infrastructure that can address these problems and reduce the security challenges [8].

### A. Physical Layer

Physical layer deals with the physical environment and collects all the data obtained from real world with the help of sensor nodes and other physical devices. This layer is responsible for communication between various physical devices. The objective of this layer is to provide services to the network and authentication of devices. The main devices [9] in physical layer includes Arduino, ZigBee, Barcodes, RFID and all other type of sensors. Each device in IoT system must have a unique tag which allows strong connection to the network and mostly Universally Unique identifiers (UUID) are used in the whole network by various devices. Uniformly a device can

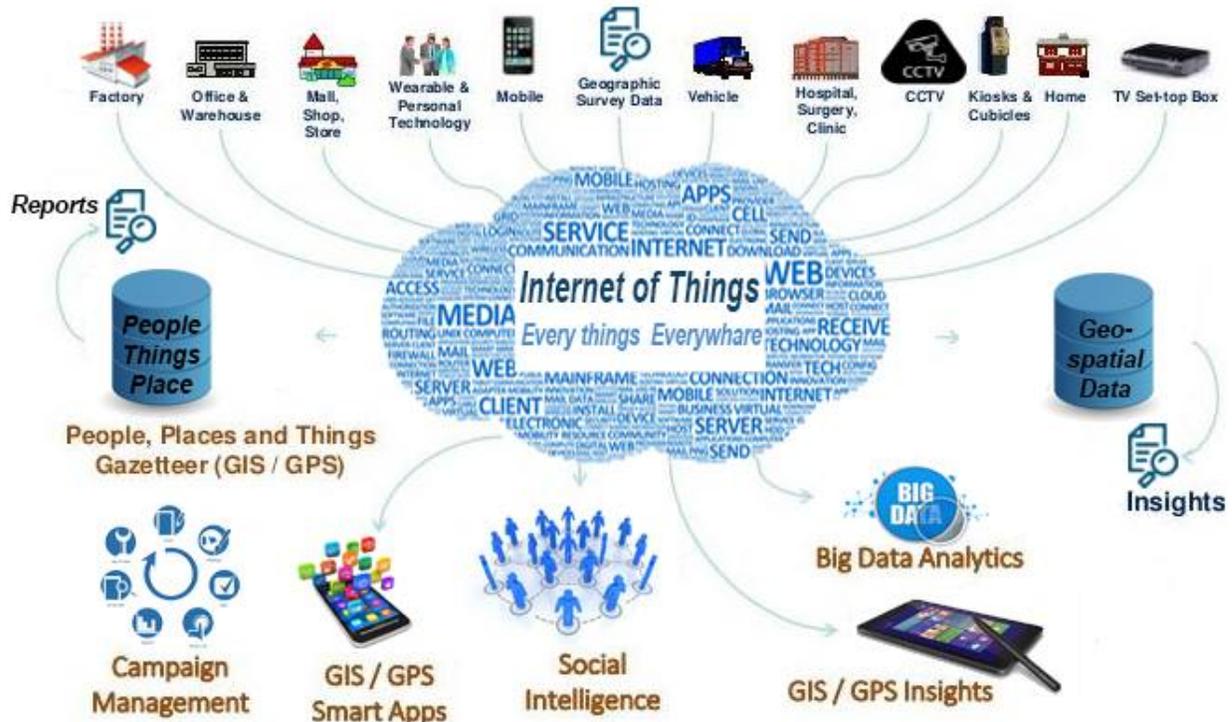


Fig. 1. IoT Diagram

be connected with many sensor nodes with unique ID because of uniquely identification of objects. Network layer carries the collection of transmitted information and transferred to central processing system.

### B. Network Layer

Network layer is responsible for communication between different physical devices, management of network and also for maintenance of information through many communication protocols in an IoT system. There is not yet any fix protocol for IoT but most common protocol now a day used is MQTT 3.1 and CoAP (Constrained Application Protocol). With the help of Wireless Sensors, the main objective of the network layer is that it gathers information which is obtained from physical layer which is further transferred to information processing unit. Every device in the IoT network sends its private information with the help of wireless sensors [10]. The network layer carried transferring of information on the network of IoT. Hence reliable and secure transfer of data is done by this layer from physical layer to other layers.

### C. Processing Layer

The working of processing layer is to combine the network and physical layer. Due to large amount of data it is very essential to store and process this data by associating with database for storage capabilities. Processing layer can automatically evaluate information and process data onto the basis of intelligent computing. Hence all the ubiquitous computing and intelligent processing function performed in this layer which

is initial technology in this layer, so future technologies of this layer will more suitable for IoT. For this reason, the innovation of future technologies of this layer will be helpful for evolution of IoT system.

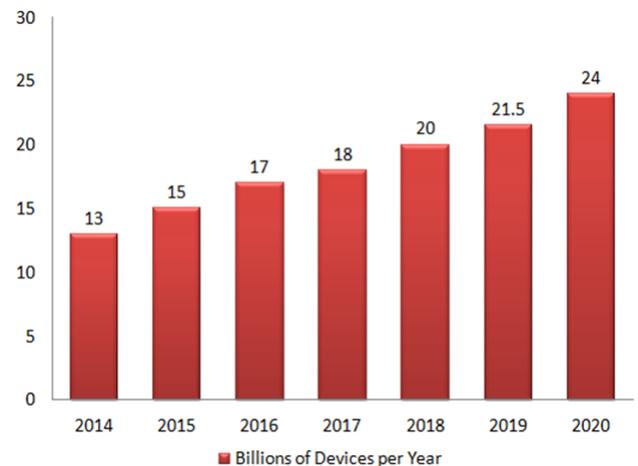


Fig. 2. Year Wise Growth in IoT Devices

### D. Application Layer

The application layer is service oriented layer which provides context-aware services between connected devices in a pervasive way for end users. The processed information on

processing layer gives a platform to application of IoT which facilitates the user needs in various ways like transportation, communication and smart hospitals etc

This paper aims to discuss security of four layered architecture of IoT as shown in Figure 4. This paper also discusses different security features, security challenges of these layers and on the bases of former research different security aspects has discuss like cryptography, communication security, protecting sensor data and outline the challenges briefly. Rest of the paper is organized as follows. Section II provides a brief overview of the security threats to each layer and their countermeasure. Section III describes the countermeasures against each attack. Section IV provides the performance evaluation is done on the basis of the literature and in Section V gives the description work done in the paper is concluded.

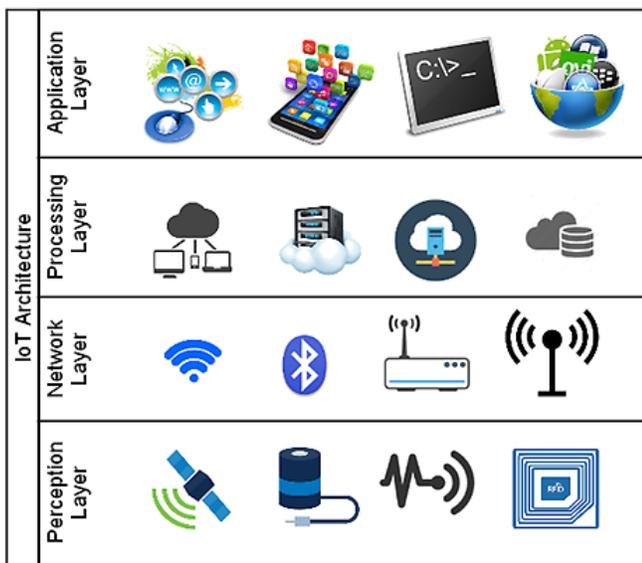


Fig. 3. IoT Layered Achitecture [86]

## II. ATTACKS AT DIFFERENT LAYERS

In this section various security threats which threaten the confidentiality of data and their possible countermeasures on each layer which are suggested recently are briefly discussed as shown in figure 3.

### A. Physical Layer

Physical layer composed of various enabling sensor technologies such as Bluetooth, GPS and Zigbee which are unprotected to different types of attacks. This type of attack is implemented [11] on the hardware parts of the IoT network and the adversary needs to be close to the IoT systems. Table 1 analyzes briefly on physical layer attacks.

1) *Node Tempering*: This type of attack may destroy the sensor node or cause damages by physically sending and receiving complete node or component of hardware or even electronically examine the nodes to get access and change sensitive information [12]. For Example:-shared cryptographic keys or influence the process of higher communication layer.

2) *Jamming of node in Wireless Sensor Network*: Node Jamming is much more popular in wireless sensor networks and is similar as in RF Interference attack explained above. This type of attacker gets involved in radio frequencies of wireless sensor nodes [13] and then afterwards it blocks the signals which stop the communication of nodes. When attacker successfully handles the blockage in key sensor nodes then it can stop service to IoT [14]. DoS attack can disturb RF signals by sending a large number of noisy signals which disrupt the network which in turn causes RF jamming.

3) *RF interface on RFID*: Dos attack can be impose on any tag of RFID. Denial of service attack implemented by sending noisy signal across radio frequency signal [15] when these noisy signals implemented on RFID then it stops communication.

4) *Malicious Node Injection*: This type of attack is also known as man in the middle attack. The attacker can actually set up a new malicious node between the sender and receiver node by this mechanism it controlled all the data [16] from one end to another in IoT system.

5) *Physical Damage*: The attacker can damage the network of IoT by attacking on the devices for its own purpose. This type of attack deals with the security that hosts by IoT system. This type of attack is different from Node Tempering attack [17] because in this attack attacker tries to directly damage the IoT services.

6) *Social Engineering*: In this type of attack the adversary can exploit the user of IoT system, to get useful and secret information and to achieve task by extracting that type of private information [18]. This type of attack is categorized into physical attack because the attacker physically communicates with the network of IoT to serve his task.

7) *Sleep Deprivation Attack*: Many sensor nodes are activated and perform its functions by replaceable batteries in IoT system and these sensor nodes are programmed to follow some functions such as sleep routines for the enhancement of their battery lifetime. This type of attack keeps the sensor nodes busy all the time [19] and will result in more battery consumption.

8) *Malicious Code Injection*: In this attack the adversary can physically insert a malicious program into a node and by implementing this attack into a node it would get access of the whole IoT system [20]. For Example: An attacker inserts any plug and play device into a node with harmful virus then it would gain full access of that node and control all the IoT system.

9) *Unauthorized Access to the Tags*: In this type of attack the adversary can get access to any of the tag without any authorization. This can be done due to the inadequacy of proper authentication procedure in RFID system [21]. The attacker cannot only access the data but can modify or even delete the complete information or data.

10) *Tag Cloning*: In IoT system, tags are deployed on various physical objects which are visible and thus data can be read and also modified [22] by some hacking techniques. So the crucial data can be easily accessed by any cybercriminal that can discover duplicate tag and hence the user cannot distinguish between duplicate and original data.

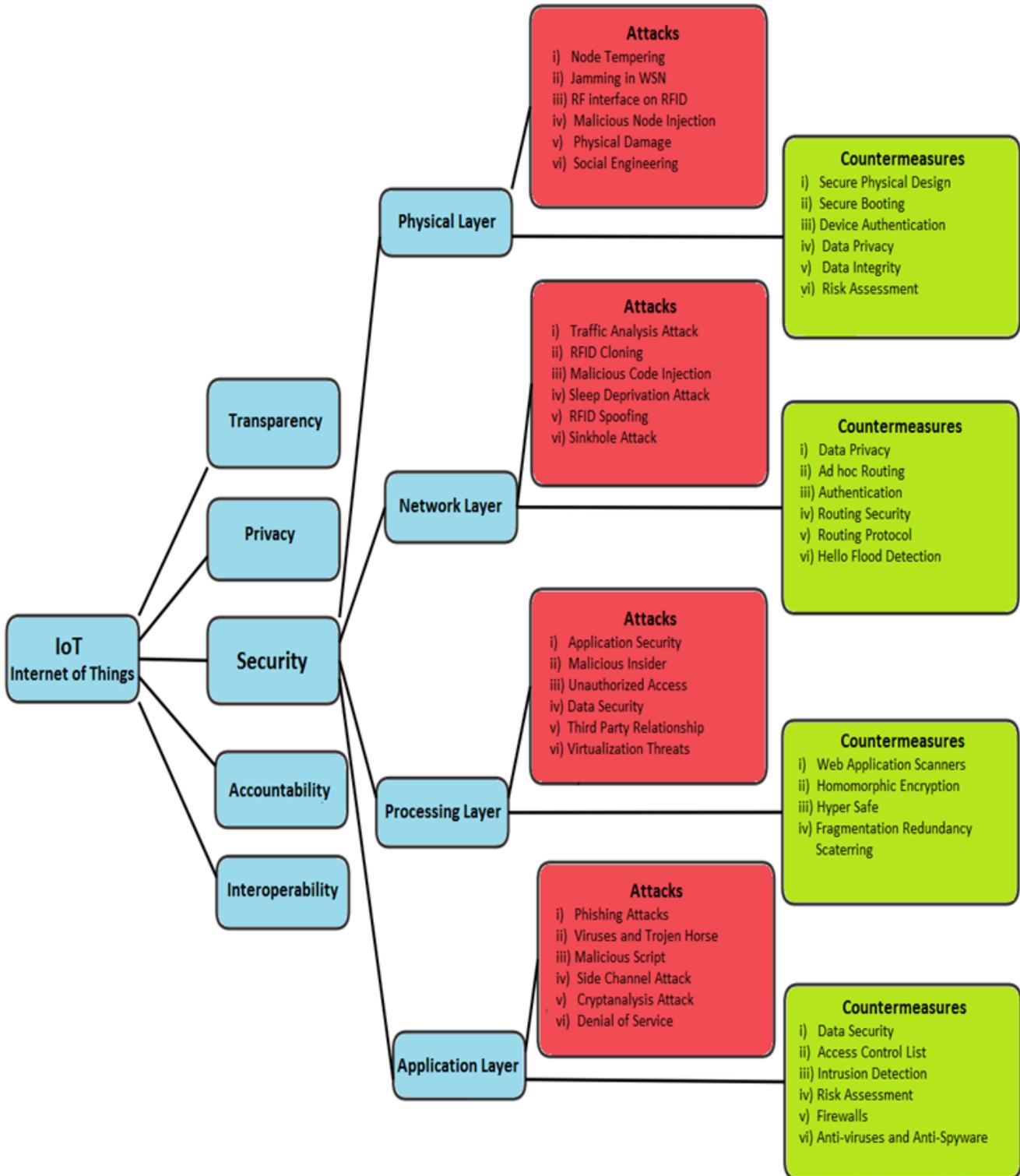


Fig. 4. Attacks and CounterMeasures on the Layers of IoT

*11 Eavesdropping:* In this type of attack the attacker can easily get confidential information such as password or some other data which are flowing from tag to user or user to tag [23]. This type of attack can happen because RFID has wireless characteristics.

*12 Spoofing:* In spoofing the adversary spreads false information on the RFID system and assumes this as original and makes that the data is appearing from original source [24]. Hence by this the attacker captures information and gets complete access to network.

*13 Timing Attack:* Another threatening attack of the confidentiality of the system is timing attack in which the attacker can get access of encryption key by analyzing the time which is required to do the [25] encryption task. Side Channel is also type of timing attack in which the adversary attacks on the encryption devices when there is leakage of information on the duration of device operation [26] like power consumption, processing or electromagnetic radiation etc.

*14 Node Capture Attack:* In node capture the attacker captures all private data and information by completely controls the node [27]. The adversary can add duplicate node to the network and by sending malicious data it threatens the confidentiality of the data.

*15 Replay Attack:* The privacy of the perception layer can be easily exploited by this type of attack. The adversary alters or replays the node by spoofs the information like identity and location etc of the node in the IoT system [28].

*16 Routing Threats:* The attacker can generate routing loops by altering and false routing information, [29] blocks the transmission of network and enlarge the network path by sending lot of error messages hence it increase point to point delay etc.

## B. Network Layer Attacks

In the network attack the adversary needs to concentrate on the network of the IoT system and the attacker does not need to be close to the network of IoT. Table 2 analyzes briefly on the network layer attacks.

*1) Traffic Analysis Attack:* Traffic analysis attack is the main security attack on the network layer when using any web browser. The adversary can access secret information and other useful data which are from RFID technology because of its wireless attribute. Before applying this attack the attacker initially captures information and data about the connected network [30]. This work is accomplishing by using some sniffing operations such port scanning applications, packet sniffer applications etc.

*2) RFID Cloning:* In this type of attack the adversary can access of useful data by mimic RFID and copying data from valid RFID to another RFID tag [31]. This type of technique does not physically simulate an original RFID tag and differentiate between composed and original, dissimilar in the case of spoofing attack on RFID.

*3) Malicious Code Injection:* This type of attack causes severe effect to the network of IoT or even may block the complete network. In this attack, [30] the adversary injects a malicious code in to a system by comprising a node. So the attacker gains full control on IoT network.

*4) Sleep Deprivation Attack:* In Wireless sensor network the sensor nodes are charged with batteries which are not compatible because life time of such batteries is not so efficient so sleep routine procedure is used to the nodes to enhance the lifetime of battery [32]. In sleep deprivation attack the adversary keep battery awake which result in more battery consumption and at last it shut down the sensor nodes.

*5) RFID Spoofing:* In this attack the adversary captures the transmission of data by spoofing the signals of an RFID. Then making it to be authentic the attacker transmits his own data which have original ID [33] of RFID tag, hence by showing to be the actual source the attacker can access the IoT system.

*6) RFID Unauthorized Access:* In RFID systems, getting access of tags is very easy for anyone because mostly in the RFID system it lacks the established procedure or any system of authentication [34]. Thus it clearly means that attacker can change, read or delete the information of sensor nodes.

*7) Sinkhole Attack:* The attacker generates a sinkhole and tempts all traffic which is from the nodes of wireless sensor network. In the sinkhole attack it harms the confidentiality and privacy of data and by stopping the transmission of packets [35] rather than sending to its destination it denies the resource to the network.

*8) Man In the Middle Attack:* The adversary can have access to confidential data, breaching the privacy between nodes by controlling, monitoring [36] the network and cause interference in the communication between two sensor nodes. Dissimilar to the types of physical attack, the attacker not need to be physically close but in network layer it must concentrate on the communication of network protocol between one node to another in an IoT system.

*9) Denial of Service:* In denial of service attack the adversary can attack on the network of IoT by sending much traffic data. It controls all the data leads to well settle denial of attack. In this type of attack the user is unable to utilize its resource over the network [37].

*10) Routing Information Attack:* In this type of attack the attacker spoofs and changes the information about routing. Arise complexity of the network build a routing loops, sending false messages, sending errors, separate the network and drop traffic signals which result failure of sending data onto its destination [38]. Hello Attack is the example of such type of routing information attack.

*11) Sybil Attack:* It is a type of malicious attack in which a neighboring node in wireless sensor network accepts false information. This type of network layer [33] attack (Sybil Attack), it claims to hold the identification of large number of nodes. For Example: A Sybil node voted by many nodes rather than one node in the wireless sensor network.

*12) Wormhole attack:* Relocation of bits can be done from the original place of bits in network [39]. The mechanism of relocation is done from that channel of bits where there is link with low latency.

*13) Hello flood attack:* In hello flood attack the attacker sends useless messages from one node and causes a traffic jamming and block the channel in the network. Only a single malicious [40] node can do this and cause blockage of entire network by creating large number of traffic.

Attack Name	References	Effects	Countermeasures	Countermeasure Reference	Countermeasure Description
<b>Node Tempering</b>	[10]	Alter sensitive information by damaging sensors	Physically secure Design	[76]	Physically Secure Designing of devices should not be changeable and not be of high quality.
<b>Jamming node in WSN</b>	[40]	Communication blockage between the nodes	IPSec Security channel	[77]	Node tempering and eavesdropping can be stopped by encryption and authentication which ensures confidentiality of data
<b>RF interface on RFID</b>	[12]	Stop communication by distortion in signals	Device Authentication	[41]	A new physical device before sending and receiving of data the device should authenticate itself
<b>Malicious node injection</b>	[13]	Create interruption in transmission process	Secure booting	[42]	Secure booting is done by cryptographic hash algorithm which checks the software on the devices by digital signature
<b>Physical Damage</b>	[14]	Attacking on devices and cause damage to the IoT network	Risk Assessment	[64]	provides confidentiality of data and avoiding security breaches in an IoT network
<b>Social Engineering</b>	[15]	Leakage of private information	Data privacy	[47]	when data is sending to the destination it avoids the attacker to access the essential data
<b>Sleep Deprivation Attack</b>	[16]	Shutdown of nodes	Device Authentication	[41]	Without any authentication the device cannot enters or connect with other node in the IoT system
<b>Unauthorized Access to the Tags</b>	[36]	Modify or delete the entire information	Device Authentication	[41]	With the help of Device Authentication, unknown device cannot communicate in the IoT network

Table 1. Physical Layer Analysis

14) *Selective forwarding*: In selective forwarding only compromised node can transmit data onto its destination. The attacker selects and restricts the nodes to achieve his malicious purpose and hence some nodes cannot forward the data packet [41].

### C. Processing Layer Attacks

The processing layer consists of different type of technologies like data storage and data processing. Cloud attack is the most significant kind of attack in IoT system and the security threats in this layer which makes network vulnerable are analyzed in Table 3

1) *Unauthorized Access*: Processing layer provides data storage and various functionalities in applications processing task [42]. In this attack, the adversary can easily access services of the system in authorizing manner and deleting the crucial data which can cause lots of damage to the IoT network.

2) *Malicious Insider*: This is insider attack in which the attacker from inside the [43] organization attacks by altering the data because of his own purpose. In this attack the data can be easily modified and extracted from purpose of the inside user.

3) *Application security*: In context of application security, Software as a service (SAAS) provides available software and

Attack Name	References	Effects	Countermeasures	Countermeasure Reference	Countermeasure Description
Sinkhole Attack	[22]	Data leakage from the nodes	security aware ad hoc routing	[78]	Stops inside attacks from the network of IoT and the adversary is dropped from the network
Traffic Analysis Attack	[18]	Leakage of secret information about the network of IoT	Routing Security	[49]	Routing security is used for data confidentiality. In this technique transmitted data is stored in packets after the analysis of data it then sent to the processing
RFID Cloning	[19]	Access useful data by mimic RFID	Authentication	[48]	With the help of proper authentication mechanism. Cloning of RFID can be prevented
RFID Spoofing	[20]	Controls transmission process and data manipulation	GPS system technique	[79]	Encounter the spoofing attack
Wormhole Attack	[31]	Relocation of bits in the network	Routing protocol	[80]	Routing protocol is used to produce the multiple paths between the sender and receiver and checks the presence of route.
Hello flood Attack	[32]	traffic jamming and channel blockage	Hello flood Detection cum Prevention	[81]	a node sends hello message to check the strength of signal if strength is similar as in radio range then receiver accept the message
Routing Information Attack	[25]	Destruction of network by routing loops	Encrypting Routing Tables	[51]	OWAS identifies different security issues on web by encryption process in rout.

Table 2. Network Layer Analysis

data on cloud through internet. The adversary in IoT system can easily steal data [44] and can operate malicious activities by using internet. Their security problems are much different than normal network security problems. Open Web Application Security Project (OWASP) has identified many web services and security issues in SAAS.

4) *Data security*: To provide and ensure data security to the user is a major responsibility for SAAS provider. Many security problems occur to the backup of data onto the service provider [45] because data backup is performed by other party which can cause data theft.

5) *Underlying infrastructure security*: In Platform as a Service (PaaS), the developers cannot access the lower layer and the security of this layer is the responsibility for service

providers [46]. The objective of developer is to maintain a secure application of IoT but security of the lower layer remains unprotected and cause vulnerability.

6) *Third-party relationships*: PaaS can also provide many third party components like mashups [47]. There is combination of many sources of mashups so it increases security issues of data and network.

7) *Virtualization threats*: Security of virtual machine is very important as other machines and the occurrence of any damage to machine affects the other. In this layer virtualization is very insecure about many kind of attacks [48].

8) *Shared Resources*: Same resource sharing and utilization in virtual machine can cause a various security threats in IoT network [49]. The adversary controls all the resources

Attack Name	References	Effects	Countermeasures	Countermeasure Reference	Countermeasure Description
Virtualization threats	[57]	Damaging the resource	Hyper Safe	[70]	Hyper Safe used for protection of the memory pages from being altered
Shared Resources	[58]	Unauthorized user can control the resources	Homomorphic encryption	[68]	cipher text is allowed to compute immediately without decryption
Application security	[51]	Data theft	Web Application Scanners	[50]	Discovery of various threats which is present on the front end of web
Data Security	[54]	Leakage of confidential data because data on Cloud	Fragmentation redundancy scattering	[67]	data on cloud is splits and allocates in to various fragments for storage in servers
Underlying infrastructure security	[55]	lower layer remains unprotected	Fragmentation redundancy scattering	[67]	Data divides and allocate to different fragments for storage
Third-party relationships	[56]	Data leakage	Encryption	[69]	In Encryption, Data is firstly encrypted and then sent to the cloud

Table 3. Processing Layer Analysis

which are shared between virtual machine by using covert channels. So sharing of data might threaten by data theft.

#### D. Software Layer Attacks

Software attacks are the major challenges arises in the IoT system. Software attacks are used to damage the system resources by using harmful viruses and attacks such as Trojan horse, worms, spyware etc that can breaches [50] the confidential data, altering data, damage the IoT devices and get access to useful information. Table 4 discussed its effects on IoT.

1) *Phishing Attack*: In this type of attack the adversary can capture useful information [51] and access of private data by spoofing authentication authorization of user. These attacks used to steal login credentials, information of credit card etc.

2) *Virus, Worms, Trojan Horse and Spyware*: The attacker affects the system of IoT by injecting malicious software in the system [52] which results in varying outcomes. These types of attacks harm the system by denying its services, altering data and get access to confidential data.

3) *Malicious Scripts*: In the IoT system usually devices are connected and communicating with each other via internet. The system occurs to a complete shutdown [53] when user monitors the gateway and runs the active-X script. This type of scripting occurs to web applications and is use to control the access and theft of data.

4) *Denial of Service*: The adversary can affect all users in a network of IoT system by injecting denial of service attack of the network of IoT by application layer hence unauthorized user can get access to systems information [50]. This type of attack also blocks the authorized users for communication with application layer. The attacker can get full access to the application layer.

5) *Data Protection and Recovery*: Privacy of user is involved in the communication with data. By improper procedure and algorithm of data processing the confidential data can be lost or may even cause a catastrophic damage [54].

#### E. Encryption Attacks

In IoT system these types of attacks is entirely using for breaking the procedure of encryption techniques.

1) *Cryptanalysis Attack*: The purpose of this type of attack is to retrieve the encryption key which is being used for breaking the mechanism of encryption in IoT system [55]. Cryptanalysis Attacks let the possession of plaintext. Chosen-plaintext attack, Known-plaintext attack, Ciphertext-only attack and Chosen Ciphertext attack are some examples of cryptanalysis attack.

2) *Side channel Attack*: In this type of attack, the attacker can find the encryption key which is used for the purpose of decrypting and encrypting data [56]. By this way the adversary can get access to hacked data by using some particular techniques such as Electromagnetic analysis and power..

3) *Man In the Middle Attack*: During a mechanism of challenge-response when two authorized users in an IoT network establishing a secure communication [57], then this time an attacker position himself and intercepting the signals. The adversary can also interfere in the communication between the users by exchanging the keys and then the attacker will able to perform encryption or decryption.

### III. COUNTERMEASURES AT DIFFERENT LAYERS

In this section countermeasure of the above mention attacks are discussed.

#### A. Physical Layer Security

Physical Layer is the bottom most layer of IoT network which provides different features of security to the hardware. Security at physical layer is discussed in four various types as discussed below:

1) *Secure Physical Design*: In Physical Layer most of the threats are resolved by designing the devices which are physically secure. Designing of such component [58] like acquisition unit, radio frequency circuits etc should not be changeable and not be of high quality. In WSN the design of antenna is physically secure and has ability to communicate over long distance.

2) *Device authentication*: When a new physical device enters in to the IoT network, then before sending and receiving of data the device should authenticate itself [59]. When the device has accurately identified then the system always keeps the malicious devices out of the network.

3) *Secure Booting*: Authenticity and originality of the software can be checked by applying cryptographic hash algorithm. This algorithm verifies the software on the devices by digital signature [60]. Many cryptographic hash algorithms cannot be implemented because of low processing capability on many devices. Some cryptographic hash algorithm such as NH and WH cryptographic algorithm are suitable for some devices which has low utilization of power.

4) *Data Confidentiality*: In data confidentiality all tags and data of each physical device should be encrypted before sending the data to provide confidentiality [61]. Strong technique of cryptographic encryption such as AES cannot be applied because power consumption is low. So Blowfish or RSA can be applied on these devices because these techniques have low processing power.

5) *Data integrity*: To avoid the tempering of sensitive data, the technique of error detection [62] is provided at each physical device. Better error detection techniques can be applied such as WH cryptographic hash method but it refers to that type of mechanism which have ability to utilize low power such as Cyclic Redundancy Checks (CRC) and parity bit.

6) *Data Privacy*: Symmetric and asymmetric encryption function like DSA, RSA, BLOWFISH and DES etc guaranteed data privacy by preventing the attacker to unauthorized access of essential data when data is sending to the destination. These encryption algorithms can be easily applied because of their less consumption of power.

7) *Risk Assessment*: Dynamical Risk Assessment technique provides confidentiality of data and avoiding security breaches in an IoT network [63]. It is essential for security perspective of IoT by discovering different types of threats to the network. When an error is discovered with such security strategies than RFID runs an automatic kill command of tags of RFID which stops unauthorized access to data.

8) *Privacy of sensitive information*: Privacy of sensitive information is the most crucial concept for providing security to the data onto the system. With the help K-anonymity [64] technique it provides mechanism to hide the sensitive information on the system hence anonymity of identity is achieved by providing protection for the information such as location and identity etc.

9) *Anonymity*: Identification of nodes and hiding of private information like data address and location are very important for confidentiality. Zero-Knowledge technique [65] would be the best solution for anonymity but it has a drawback that having a large processing power because of strong algorithm it cannot be implemented on the devices which have less consumption power. So K-anonymity is a best approach for less power physical devices in IoT network [66].

10) *IPSec Security channel*: IPSec Security channel has two types of secure functionalities, encryption and authentication which provides security [67]. Node tempering and eavesdropping can be stopped by encryption which ensures confidentiality of data. The receiver can identify that the sender of the data onto IP is fake or real.

#### B. Network Layer Security

The network layer is threatening by many types of attacks. Due to the observance of the many wireless channels, attacker can easily control the communication between devices. The security of network layer is splits in four types which are described below.

1) *Data privacy*: The safety control procedure control the network of any type of error occurs and hence integrity of data has applied to justified that data received to the user is similar to the original [68] like encryption of point to point. Authentication mechanism is used to avoid illegal access to data onto sensor node.

2) *Security aware ad hoc routing*: Security aware ad hoc routing (SAR) protocol prevents from inside attacks of the network [69] of IoT. Some security measurements are added to the packets and the adversary is dropped from the network after the analysis of received data.

3) *Authentication*: Illegal access of the nodes can be avoided with the help of proper authentication technique and encryption process [70]. In network layer the most common type of attack is DoS attack which can affect the network by spreading useless information.

4) *Routing security*: In many applications secure routing is essential for the sensor network. Due to the insecure routing protocols, different routing algorithms are applied to secure the confidentiality of data transferring towards various sensor nodes in IoT system [71]. However, multiple paths provide secure routing which fixed errors in the network and increase

Attack Name	References	Effects	Countermeasures	Countermeasure Reference	Countermeasure Description
Virtualization threats	[57]	Damaging the resource	Hyper Safe	[70]	Hyper Safe used for protection of the memory pages from being altered
Shared Resources	[58]	Unauthorized user can control the resources	Homomorphic encryption	[68]	cipher text is allowed to compute immediately without decryption
Application security	[51]	Data theft	Web Application Scanners	[50]	Discovery of various threats which is present on the front end of web
Data Security	[54]	Leakage of confidential data because data on Cloud	Fragmentation redundancy scattering	[67]	data on cloud is splits and allocates in to various fragments for storage in servers
Underlying infrastructure security	[55]	lower layer remains unprotected	Fragmentation redundancy scattering	[67]	Data divides and allocate to different fragments for storage
Third-party relationships	[56]	Data leakage	Encryption	[69]	In Encryption, Data is firstly encrypted and then sent to the cloud

Table 4. Application Layer Analysis

performance of the system. For routing purpose source routing is a technique in which transmitted data is stored in packets after the analysis of data it then sent to the processing.

5) *GPS location system*: GPS system encountered the spoofing attack from network layer of the IoT system [72]. S. Daneshmand et al. describe and implemented the GPS location technique which is the best solution proposed yet.

6) *Routing protocol*: Ad hoc On demand Multipath Distance Vector (AOMDV) is a routing protocol which encountered the wormhole attack [73]. Amish et al. propose this technique by producing multiple paths between the sender and receiver in every discovery of rout. In this technique route table is checked by the sender that for two nodes communication, route is available or not. If the rout is available then it provides information about routing rather it transmits the packet.

7) *Hello flood Detection cum Prevention*: Virendra et al. propose a technique to prevent hello flood attack in IoT. In this technique a node sends hello message to check the strength of signal if strength is similar as in radio range then receiver accepts the message and information about routing is sent to the rout [74].

8) *Data Integrity*: A cryptographic hash mechanism is used to for the integrity of data [75]. This function is used to check the transmission of data onto the other node. When tempering of data is proved error correction process can also be used.

### C. Processing Layer Security

There are some concepts of security measures in processing layer which is discussed below:

1) *Web application scanners*: This application is using for identification of different threats [76] which is present in the front end of web. Other web firewall applications are also detecting the attacks of potential attacker.

2) *Fragmentation redundancy scattering (FRS)*: In FRS the essential data onto cloud [77] is splits and allocates in to various fragments of storage in servers. The fragment has not any useful information about the data so risk of data theft is minimized in this scenario.

3) *Homomorphic encryption*: This technique is based on entire mechanism of homomorphic encryption. In this technique [78] cipher text is allowed to compute immediately without decryption. High computation requires for data security in this method.

4) *Encryption*: Encryption technique is used to ensure the data confidentiality in IoT. Data is firstly encrypted and then sent to the cloud. Encryption helps to overcome against side channel attacks [79]. There are various kinds of encryption such as Advanced Encryption Standard etc.

5) *Hyper Safe*: Hyper safe provides protection for the memory pages from being altered and also allows restriction of pointing index that changes monitored data onto the pointer indexes [80].

#### D. Application Layer Security

The categorization of security mechanism in application layer is discussed below:

1) *Data security*: For securing the confidentiality of data and privacy of entire IoT system Encryption, Authentication and Integrity are the most essential procedure at this level. It avoids any unauthorized access to the data and protecting data to be hacked or theft.

2) *Access Control Lists (ACLs)*: Setting up the rules and allows request for the access and monitoring of the network is the important part which ensures the confidentiality of the system and data privacy [81]. ACL can manage by stopping or allowing incoming or outgoing traffic and monitors access requests from many users in the IoT system.

3) *Intrusion Detection*: Intrusion Detection process [82] provides security solutions to many threats by producing an alarm when any uncertain action is performed in the system because of continuous controlling a log of intruder's activity. Intrusion detection can be done by various detection techniques such as anomaly detection in data mining [83].

4) *Risk Assessment*: The risk assessment produces effective security approaches and gives enhancement of [84] already existing architectures and planning of security.

5) *Firewalls*: When encryption, authentication and ACLs process failed to block the unauthorized user then firewall comes in process [85] for the blockage. When weak password was chosen then encryption and authentication process can be failed. In firewall, filtration of packets is done hence unwanted packets are blocked by this process.

6) *Anti-virus, Anti-spyware and Anti-adware*: Software which provides security such as anti-virus, anti-spyware and anti-adware is essential for the confidentiality, reliability and integrity of the IoT network. IV. PERFORMANCE EVALUATION Evaluation of security threats on the network of IoT are done in this section and discuss their countermeasures. Furthermore, this paper mentions the effects of these attacks on IoT network and also presents separate countermeasures for which it reduces the damage of the IoT and prevention of vulnerability. Detail of our performance analysis is discussed in table.1 in which detailed analysis has been done on the basis of each attack.

#### IV. PERFORMANCE EVALUATION

Evaluation of security threats and their countermeasures to the network of IoT is discussed in this section. Furthermore, this paper mentions the effects of these attacks of IoT network and also presents separate countermeasures for which it reduces the damage to the IoT and prevention of vulnerability. Detail of our performance analysis is discussed in table.1 in which detailed analysis has been done on the basis of each attack.

#### V. CONCLUSION

IoT has been considered as an important research topic for the last few years where physical objects would communicate by using various network technologies. The vast advancement of the services of IoT requires the authentic and factual security

mechanism. This paper gives a broad overview of IoT by describing the working of layers and then discusses different security loopholes on different layers of IoT (Physical Layer, Network Layer, Processing Layer and Application Layer). Furthermore it presents the countermeasures against security threats from the prevention of any damage to IoT network. As IoT is going to be an essential part of our life, steps should be taken to ensure security and privacy of the user.

#### REFERENCES

- [1] Daniele Miorandi, Sabrina Sicarib, Francesco De Pellegrini, "Internet of things: Vision, applications and research challenges" survey paper September 2012, pp 1497–1516.
- [2] Misra, Gourav, et al. "Internet of things (iot)—a technological analysis and survey on vision, concepts, challenges, innovation directions, technologies, and applications (an upcoming or future generation computer communication system technology)." American Journal of Electrical and Electronic Engineering 4.1 (2016): 23-32.
- [3] Torğul, Belkız, Lütfü Şağbaşıua, and Figen Balo. "Internet of Things: A Survey." (2016): 104-110.
- [4] Krajjak, Surapon, and Panwit Tuwanut. "A survey on internet of things architecture, protocols, possible applications, security, privacy, real-world implementation and future trends." Communication Technology (ICCT), 2015 IEEE 16th International Conference on. IEEE, 2015.
- [5] Sicari, Sabrina, et al. "Security, privacy and trust in Internet of Things: The road ahead." Computer Networks 76 (2015): 146-164.
- [6] Alaba, Fadele Ayotunde, et al. "Internet of things Security: A Survey." Journal of Network and Computer Applications (2017).
- [7] Geng Yang, Jian Xu, etc.: Security Characteristic and Technology in the Internet of Things. J. Journal of Nanjing University of Posts and Telecommunications (Natural Science). 30(4) (2010) (in Chinese)
- [8] Andrea, Ioannis, Chrysostomos Chrysostomou, and George Hadjichristofi. "Internet of Things: Security vulnerabilities and challenges." Computers and Communication (ISCC), 2015 IEEE Symposium on. IEEE, 2015.
- [9] A. Mukherjee, "Physical-Layer Security in the Internet of Things: Sensing and Communication Confidentiality Under Resource Constraints," Proc. IEEE, vol. 103, no. 10, pp. 1747–1761, 2015.
- [10] H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the internet of things: A review," Proc. - 2012 Int. Conf. Comput. Sci. Electron. Eng. ICCSEE 2012, vol. 3, pp. 648–651, 2012.
- [11] Pan, Yao, et al. "Taxonomies for Reasoning About Cyber-physical Attacks in IoT-based Manufacturing Systems." International Journal of Interactive Multimedia & Artificial Intelligence 4.3 (2017).
- [12] Kaushal, Kanchan, and Varsha Sahni. "DoS Attacks on different Layers of WSN: A Review." International Journal of Computer Applications 130.17 (2015).
- [13] Wahid, Abdul, et al. "A Survey on Attacks, Challenges and Security Mechanisms in Wireless Sensor Network." International Journal For Innovative Research in Science & Technology 1 (2015).
- [14] Sonar, Krushang, and Hardik Upadhyay. "A Survey: DDOS Attack on Internet of Things." International Journal of Engineering Research and Development 10.11 (2014): 58-63.
- [15] Peris-Lopez, Pedro, et al. "attacking RFID systems." Security in RFID and Sensor Networks (2016): 29.
- [16] Illiano, Vittorio P., and Emil C. Lupu. "Detecting malicious data injections in wireless sensor networks: A survey." ACM Computing Surveys (CSUR) 48.2 (2015): 24.
- [17] Jacobson, Michael. "Vulnerable Progress: The Internet of Things, the Department of Defense and the Dangers of Networked Warfare." (2015).
- [18] Ghafir, Ibrahim, et al. "Social engineering attack strategies and defence approaches." Future Internet of Things and Cloud (FiCloud), 2016 IEEE 4th International Conference on. IEEE, 2016.
- [19] Kumar, Sathish Alampalayam, Tyler Vealey, and Harshit Srivastava. "Security in internet of things: Challenges, solutions and future directions." System Sciences (HICSS), 2016 49th Hawaii International Conference on. IEEE, 2016.

- [20] Farooq, M. U., et al. "A critical analysis on the security concerns of internet of things (IoT)." *International Journal of Computer Applications* 111.7 (2015).
- [21] Da Xu, Li, Wu He, and Shancang Li. "Internet of things in industries: A survey." *IEEE Transactions on industrial informatics* 10.4 (2014): 2233-2243.
- [22] Doinea, Mihai, et al. "Internet of Things Based Systems for Food Safety Management." *Informatica Economica* 19.1 (2015): 87.
- [23] Rahman, Abdul Fuad Abdul, Maslina Daud, and Madihah Zulfa Mohamad. "Securing Sensor to Cloud Ecosystem using Internet of Things (IoT) Security Framework." *Proceedings of the International Conference on Internet of things and Cloud Computing*. ACM, 2016.
- [24] Jeyanthi, N., Shreyansh Banthia, and Akhil Sharma. "Security in IoT Devices." *Security Breaches and Threat Prevention in the Internet of Things*. IGI Global, 2017. 96-116.
- [25] Sadeghi, Ahmad-Reza, Christian Wachsmann, and Michael Waidner. "Security and privacy challenges in industrial internet of things." *Design Automation Conference (DAC), 2015 52nd ACM/EDAC/IEEE*. IEEE, 2015.
- [26] Babar, Sachin, et al. "Proposed embedded security framework for internet of things (iot)." *Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE), 2011 2nd International Conference on*. IEEE, 2011.
- [27] Roman, Rodrigo, Jianying Zhou, and Javier Lopez. "On the features and challenges of security and privacy in distributed internet of things." *Computer Networks* 57.10 (2013): 2266-2279.
- [28] Jan, Mian Ahmad, and Muhammad Khan. "Denial of Service Attacks and Their Countermeasures in WSN." *IRACST-International Journal of Computer Networks and Wireless Communications (IJCNWC)* 3 (2013).
- [29] Puthal, Deepak, et al. "Threats to Networking Cloud and Edge Data-centers in the Internet of Things." *IEEE Cloud Computing* 3.3 (2016): 64-71.
- [30] Hossain, Md Mahmud, Maziar Fotouhi, and Ragib Hasan. "Towards an analysis of security issues, challenges, and open problems in the internet of things." *Services (SERVICES), 2015 IEEE World Congress on*. IEEE, 2015.
- [31] Sari, Arif. "Security issues in RFID Middleware Systems: Proposed EPC implementation for network layer attacks." *Transactions on Networks and Communications* 2.5 (2014): 01-06.
- [32] Nia, Arsalan Mohsen, and Niraj K. Jha. "A comprehensive study of security of internet-of-things." *IEEE Transactions on Emerging Topics in Computing* (2016).
- [33] Borgohain, Tuhin, Uday Kumar, and Sugata Sanyal. "Survey of security and privacy issues of Internet of Things." *arXiv preprint arXiv:1501.02211* (2015).
- [34] Cvitić, Ivan, Miroslav Vujić, and Siniša Husnjak. "Classification of security risks in the IoT environment." *26th International DAAAM Symposium on Intelligent Manufacturing and Automation*. 2016.
- [35] Jing, Qi, et al. "Security of the Internet of Things: perspectives and challenges." *Wireless Networks* 20.8 (2014): 2481-2501.
- [36] Conti, Mauro, Nicola Dragoni, and Viktor Lesyk. "A Survey of Man In The Middle Attacks." *IEEE Communications Surveys & Tutorials* 18.3 (2016): 2027-2051.
- [37] Zhang, Congyingzi, and Robert Green. "Communication security in internet of thing: preventive measure and avoid DDoS attack over IoT network." *Proceedings of the 18th Symposium on Communications & Networking*. Society for Computer Simulation International, 2015.
- [38] Jing, Qi, et al. "Security of the Internet of Things: perspectives and challenges." *Wireless Networks* 20.8 (2014): 2481-2501.
- [39] Pongle, Pavan, and Gurunath Chavan. "Real time intrusion and worm-hole attack detection in internet of things." *International Journal of Computer Applications* 121.9 (2015).
- [40] Yassen, Muneer Bani, Shadi Aljawaerneh, and Reema Abdulraziq. "Secure low energy adaptive clustering hierarchal based on internet of things for wireless sensor network (WSN): Survey." *Engineering & MIS (ICEMIS), International Conference on*. IEEE, 2016.
- [41] Airehrour, David, Jairo Gutierrez, and Sayan Kumar Ray. "A trust-aware RPL routing protocol to detect blackhole and selective forwarding attacks." *Australian Journal of Telecommunications and the Digital Economy* 5.1 (2017): 50.
- [42] Ye, Ning, et al. "An efficient authentication and access control scheme for perception layer of internet of things." (2014).
- [43] Chen, Long. *Security Management for The Internet of Things*. Diss. University of Windsor (Canada), 2017.
- [44] A. Razzaq, K. Latif, H. F. Ahmad, A. Hur, Z. Anwar, and P. C. Bloodsworth, "Semantic security against web application attacks," *Inf. Sci. (Ny)*., vol. 254, pp. 19–38, 2014.
- [45] D. H. Patil, "Data Security over Cloud," *Int. J. Comput. Appl.*, pp. 11–14, 2012.
- [46] B. R. Chandramouli and P. Mell, "State of Security Readiness," *Crossroads*, vol. 16, no. 3, pp. 23–25, 2010.
- [47] K. Hashizume, D. G. Rosado, E. Fernández-medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," pp. 1–13, 2013.
- [48] N. Kilari and C. Applications, "A Survey on Security Threats for Cloud Computing," *Int. J. Eng. Res. Technol.*, vol. 1, no. 7, pp. 1–10, 2012.
- [49] K. Dahbur, "A Survey of Risks , Threats and Vulnerabilities in Cloud Computing," in *International Conference on Intelligent Semantic Web-Services and Applications*, 2011.
- [50] Zhang, Zhi-Kai, et al. "IoT security: ongoing challenges and research opportunities." *Service-Oriented Computing and Applications (SOCA), 2014 IEEE 7th International Conference on*. IEEE, 2014.
- [51] Thakur, Himani, and Supreet Kaur. "A Survey Paper On Phishing Detection." *International Journal of Advanced Research in Computer Science* 7.4 (2016).
- [52] Akbari Roumani, M., et al. "Value analysis of cyber security based on attack types." *ITMSOC: Transactions on Innovation and Business Engineering* 1 (2016): 34-39.
- [53] Dorsemaine, Bruno, et al. "A new approach to investigate IoT threats based on a four layer model." *New Technologies for Distributed Systems (NOTERE), 2016 13th International Conference on*. IEEE, 2016.
- [54] A. Viejo, "Systems and methods for reducing unauthorized data recovery from solid-state storage devices," Merry, Jr. al, vol. 2, no. 12, p. Merry, Jr. et al, 2011.
- [55] Ndibanje, Bruce, Hoon-Jae Lee, and Sang-Gon Lee. "Security analysis and improvements of authentication and access control in the internet of things." *Sensors* 14.8 (2014): 14786-14805.
- [56] Choi, Jaehak, and Youngseop Kim. "An improved LEA block encryption algorithm to prevent side-channel attack in the IoT system." *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016 Asia-Pacific*. IEEE, 2016.
- [57] Conti, Mauro, Nicola Dragoni, and Viktor Lesyk. "A Survey of Man In The Middle Attacks." *IEEE Communications Surveys & Tutorials* 18.3 (2016): 2027-2051.
- [58] B. Y. Mo, T. H. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber – Physical Security of a Smart Grid Infrastructure," *Proc. IEEE*, vol. 100, no. 1, pp. 195–209, 2012.
- [59] Trappe, Wade. "The challenges facing physical layer security." *IEEE Communications Magazine* 53.6 (2015): 16-20.
- [60] Lake, David, et al. "Internet of things: Architectural framework for ehealth security." *Journal of ICT Standardization*, River Publishing 1 (2014).
- [61] J. P. Kaps, "Cryptography for ultra-low power devices." PhD diss., WORCESTER POLYTECHNIC INSTITUTE, 2006.
- [62] Rault, Tifenn, Abdelmadjid Bouabdallah, and Yacine Challal. "Energy efficiency in wireless sensor networks: A top-down survey." *Computer Networks* 67 (2014): 104-122.
- [63] C. Liu, Y. Zhang, J. Zeng, L. Peng, R. Chen, Research on Dynamical Security Risk Assessment for the Internet of Things inspired by immunology, in *Eighth International Conference on Natural Computation (ICNC)*, 2012
- [64] K.E. Emam, F.K. Dankar, Protecting Privacy Using kAnonymity, in *Journal of the American Medical Informatics Association*, Volume 15, Number 5, 2008

- [65] Flood, Pdraig, and Michael Schukat. "Peer to peer authentication for small embedded systems: A zero-knowledge-based approach to security for the Internet of Things." *Digital Technologies (DT)*, 2014 10th International Conference on. IEEE, 2014.
- [66] Sun, Gang, et al. "Efficient location privacy algorithm for Internet of Things (IoT) services and applications." *Journal of Network and Computer Applications* (2016).
- [67] D. Migault, D. Palomares, E. Herbert, W. You, G. Ganne, G. Arfaoui, and M. Laurent, "E2E: An Optimized IPsec Architecture for Secure And Fast Offload," in *Seventh International Conference on Availability, Reliability and Security E2E*, 2012.
- [68] Abomhara, Mohamed, and Geir M. Kjøien. "Security and privacy in the Internet of Things: Current status and open issues." *Privacy and Security in Mobile Systems (PRISMS)*, 2014 International Conference on. IEEE, 2014.
- [69] S. Sharmila, "Detection of sinkhole Attack in Wireless Sensor Networks using Message Digest Algorithms," IEEE, pp. 0–5, 2011.
- [70] Zhao, Kai, and Lina Ge. "A survey on the internet of things security." *Computational Intelligence and Security (CIS)*, 2013 9th International Conference on. IEEE, 2013.
- [71] Granjal, Jorge, Edmundo Monteiro, and Jorge Sá Silva. "Security for the internet of things: a survey of existing protocols and open research issues." *IEEE Communications Surveys & Tutorials* 17.3 (2015): 1294-1312.
- [72] S. Daneshmand, A. Jafarnia-jahromi, A. Broumandan, and G. Lachapelle, "A Low-Complexity GPS Anti-Spoofing Method Using a Multi-Antenna Array," *ION GNSS12 Conf.*, pp. 1–11, 2012.
- [73] Amish, Parmar, and V. B. Vaghela. "Detection and prevention of wormhole attack in wireless sensor network using AOMDV protocol." *Procedia computer science* 79 (2016): 700-707.
- [74] Singh, Virendra Pal, Aishwarya S. Anand Ukey, and Sweta Jain. "Signal strength based hello flood attack detection and prevention in wireless sensor networks." *International Journal of Computer Applications* 62.15 (2013).
- [75] Singla, Aashima, and Ratika Sachdeva. "Review on security issues and attacks in wireless sensor networks." *International Journal of Advanced Research in Computer Science and Software Engineering* 3.4 (2013).
- [76] B. L. Suto, "Analyzing the Accuracy and Time Costs of Web Application Security Scanners," *San Fr.*, no. October 2007, 2010.
- [77] Y. Singh, F. Kandah, and W. Zhang, "A Secured Cost-effective Multi-Cloud Storage in Cloud Computing," *IEEE INFOCOM*, pp. 619–624, 2011.
- [78] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE," *SIAM J. Comput.*, vol. 43.2, pp. 831–871, 2014.
- [79] D. Koo, J. Hur, and H. Yoon, "Secure and efficient data retrieval over encrypted data using attribute-based encryption in cloud storage q," *Comput. Electr. Eng.*, vol. 39, no. 1, pp. 34–46, 2013.
- [80] S. Kumar, S. Pal, A. Kumar, and J. Ali, "Virtualization , The Great Thing and Issues in Cloud Computing," *Int. J. Curr. Eng. Technol.*, pp. 338–341, 2013.
- [81] M. Ongtang, S. McLaughlin, W. Enck, and P. McDaniel, "Semantically rich application-centric security in Android," *Secur. Commun. Networks*, no. August 2011, pp. 658–673, 2012.
- [82] Animesh Patcha, Jung-Min Park, An overview of anomaly detection techniques: Existing solutions and latest technological trends, in *Computer Networks*, Volume 51, Issue 2, 2007
- [83] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic and Marimuthu Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions."
- [84] C. Liu and Y. Zhang, "Research on Dynamical Security Risk Assessment for the Internet of Things Inspired by Immunology," in *8th International Conference on Natural Computation*, 2012, no. *Icnc*, pp. 874–878.
- [85] S. L. Wiley, O. Park, and U. S. C, "Pin-hole firewall for communicating data packets on a packet network," 2011.
- [86] M. M. Ahmed, M. A. Shah and A. Wahid, "IoT Security: A Layered Approach for Attacks & Defenses," in *International conference on communication technologies*, 2017,

# Ladder Networks: Learning under Massive Label Deficit

Behroz Mirza and Tahir Syed  
National University of Computer and Emerging Sciences,  
Karachi, Pakistan

Jamshed Memon  
Barrett Hodgson University,  
Karachi, Pakistan

Yameen Malik  
Smartlytics  
Karachi, Pakistan

**Abstract**—Advancement in deep unsupervised learning are finally bringing machine learning close to natural learning, which happens with as few as one labeled instance. Ladder Networks are the newest deep learning architecture that proposes semi-supervised learning at scale. This work discusses how the ladder network model successfully combines supervised and unsupervised learning taking it beyond the pre-training realm. The model learns from the structure, rather than the labels alone transforming it from a label learner to a structural observer. We extend the previously-reported results by lowering the number of labels, and report an error of 1.27 on 40 labels only, on the MNIST dataset that in a fully supervised setting, uses 60000 labeled training instances.

**Keywords**—Ladder networks; semi-supervised learning; deep learning; structure observer

## I. INTRODUCTION

Over the past decades, there has been an effort in machine learning theoretical research to move from supervised to unsupervised methods, for the reasons of 1. arduous effort in labeling data, and 2. the inherent aptitude of unsupervised approaches to discover the latent structure of data without the guiding (or misguiding) external influence of labels.

The paper discusses the opportunities and strengths in deep unsupervised learning and its implications towards unsupervised and weekly supervised learning in general. The model selected for this purpose is the recently-introduced *ladder network* designed by Valpola [1]. This work modifies the model configuration and reports an error on the extremely popular MNIST benchmark of 1.27 using 40 labels only. This is 10 labels fewer than previously reported results.

Ladder networks successfully combine supervised learning with unsupervised learning in deep neural networks models. Prior to this unsupervised learning was used for specialized pre-training task, followed by supervised learning. However ladder networks, are trained to simultaneously minimize the sum of supervised and unsupervised cost functions using backpropagation, thus eliminating the need for layer-wise pre-training.

The model has the distinctive feature of learning from the structure in the data instead of solely from the labels alone. This novelty results in minimizing the amount of labelled data required for training the network. As most of the data are unlabeled, the model learns principal features from the small set of labelled data and correlated features from the large set

of unlabeled data concurrently [2]. This makes the machine learning process narrowly closer to natural learning.

The rest of the paper is organized as follows. Section II - ‘Deep Unsupervised Learning’ discusses the models namely RBM and Auto encoders. Section III - ‘Semi Supervised Learning’ discusses the Ladder Networks model followed by the experiments and results. Section IV - ‘Conclusion’ concludes the paper and discusses future research directions.

## II. DEEP UNSUPERVISED LEARNING

### A. Relaxing Supervision

Unsupervised learning forms a class of machine learning techniques of deducing a function to disentangle hidden structure from unlabeled data. What clearly distinguishes unsupervised learning from supervised learning is unlabeled samples are used during training so there is no error or reward signal to evaluate a potential solution. As unsupervised learning attempts to draw inferences from datasets consisting of input data without labeled responses it is closely related to the problem of density estimation in statistics [3].

Hinton and Salakudinov [4] proposed the idea of the stochastic RBM; symmetrical arrangement of binary stochastic neurons in a Boltzmann Machine where the two layers of the model for a bipartite graph. Later works by [5] suggested Auto Encoders for pre training; pre train each successive layers using unsupervised measure thus producing an enriched useful higher-level representation from the lower-level representation output. State of the art generalization can later be achieved by running Gradient descent on supervised format. Transitioning probability to unsupervised learning looks promising based on the fact - natural learning is unsupervised; we learn the structure around us by observing not by the names of the associated objects.

### B. Greedy Unsupervised Pre-training

The year 2006 marks the breakthrough in training deep architectures as RBM were proposed followed by stacked autoencoders (SAEs) (Fig. 1). Both approaches used the notion of Greedy layer-wise unsupervised pre-training followed by supervised fine-tuning. The concepts Greedy layer wise pre training and Supervised fine tuning have profound impacts on Unsupervised Learning. Unsupervised pre-training leads to

- Pre-conditioning the model, whereby arranging the parameter values in suitable ranges later to be used in

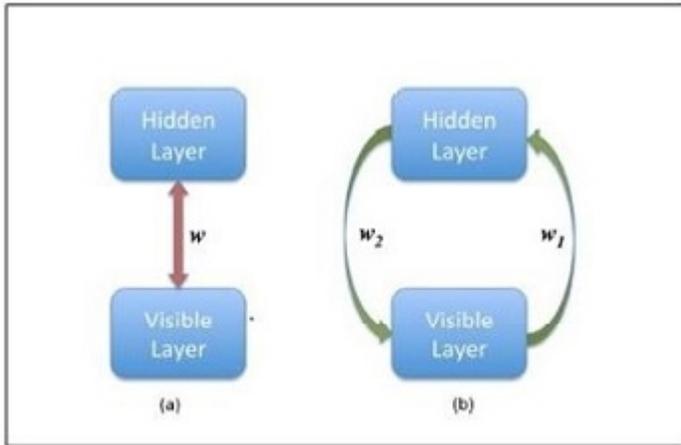


Fig. 1. RBM and Autoencoder architectures [6].

supervised training. This is completely different than a generic random initialization.

- Model initialization to an already close to optimal configuration in parameter space leads to an optimization of the optimization process. The appropriate range increases backpropagation efficiency and error minimization.

Greedy layer-wise unsupervised pre-training introduces a useful prior to the supervised fine-tuning training [7].

- Unsupervised learning is used to draw inferences from datasets consisting of input data without labeled responses, this makes it different from the supervised counterpart.
- Only use the inputs  $x(t)$  for learning.
- Auto extracting meaningful features from data. This distinctive feature extraction capability without utilization of labels, makes unsupervised learning a prime candidate simulating near to human learning behavior.
- Leverage the availability of unlabeled data. Most data being unlabeled and natural intelligence deals with analyzing the bulk of objects via structures rather than labels, therefore unsupervised learning can utilize this massive data in training deep models exhibiting the distinction of learning via structures and not simply labels.

Two popular neural models for unsupervised learning are:

- 1) Restricted Boltzmann Machines
- 2) Autoencoders

### C. Restricted Boltzmann Machines

Restricted Boltzmann Machines are autoencoder models that have the distinctive capability of transforming and reducing high dimensional data to low dimension. These use an effective way to initialize the weights of the model with calibrated values followed by gradient descent for fine tuning.

This model uses the sigmoid non-linearity, thus be called nonlinear generalization of the PCA. The model has outperformed the Linear PCA producing outstanding results, [8].

Trials confirm that weight optimization is challenging in Deep Non Linear Auto Encoders. Initialization with Large weight values leads to poor local minima problem, while initializing with small weights leads to very small gradients. However if an Auto Encoder initialized is with good calibrated weights, Gradient Descent performs well, but this initialization requires a novel algorithm which learns one layer of feature at a time. Each layer captures strong, high-order correlations b/w activities of units in layer below it.

As examples of unsupervised learning, RBM are used for pre-training phase extensively. The phase consists of learning a stacked RBM each having only one layer of feature detectors. The learned feature activations of one RBM are used as data for training the next RBM in the stack. After pre training these RBM are unfolded to create a deep auto encoder, which is later fine-tuned using back propagation.

### D. Free Energy

The RBM model defines a distribution over  $x$  with latent variables via an energy function  $E$ . The function gives the probability distribution  $P(v, h)$  where:

$$E(v, h) = - \sum_{i \in pixels} b_i v_i - \sum_{j \in features} b_j v_j - \sum_{i, j} v_i h_j w_{ij} \quad (1)$$

If  $w$  is negative it leads to high energy and the probability decreases, if  $w$  is positive it leads to low energy and the probability increases. The challenge here: the function is divided by the partition function  $Z$ , the sum over all values of  $v$  and  $h$ . As  $v, h$  are binary, so  $Z$  can take many values leading to an exponential sum over the numerator, thus making computing it intractable. To overcome this challenge Hinton et al. proposed *contrastive divergence*.

### E. Contrastive Divergence: The Negative Sample

Contrastive divergence [9] was proposed by Hinton. It uses Gibbs sampling to approximate joint distribution when direct sampling is difficult. Alternating between layers, given one unit in visible layer, all units are independent in hidden layer, values in one layer be sampled given a value in another layer (Fig. 2). Using contrastive Divergence we get the following interesting relations between Energy and Probability.

- Increase probability of observing  $x_t$  at hidden layer, decrease the energy
- Decrease probability of observing  $x_t$  at hidden layer increase the energy.
- Increase probability of observing digits from training set.
- Decrease probability of observing noise.
- Decreasing energy of things that looks like what is in training set.

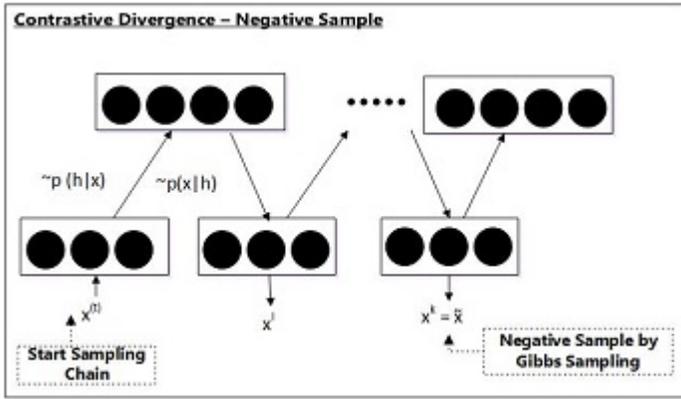


Fig. 2. Negative sample.

- Increase energy of things that are hallucinated or sampled by the model
- Ultimately model spits out things similar to the what in the model i.e. after a series of cycles  $x$  becomes closer to  $x^t$

F. Auto Encoders

An autoencoder model [10] attempts to regenerate its input. It has a hidden layer that describes a code used to represent the input. The model network comprises of an encoder function  $h = f(x)$  and a decoder that produces a reconstruction  $r = g(h)$ . Auto encoders are designed to be unable to learn to copy perfectly. As the model is forced to prioritize which aspects of the input should be copied, it often learns useful properties of the data [11].

As discussed above the model has two parts namely the encoder and the decoder. Encoder takes an input and encodes it in a linear representation. Encoder is sigmoid of a linear transformation. Decoder takes latent representation  $h(x)$ , passes it to non-linearity, generating the output. As the Auto encoder attempts to minimize the reconstruction error between actual and generated value during training, a typical Loss function is:

$$L(x, x') = \|x - x'\|^2 = \|x - \sigma_2(W'(\sigma_1(Wx + b) + b'))\|^2 \tag{2}$$

For deep autoencoders, their representational power, layer size and depth can be elaborated as:

- Universal Approximation Theorem [12] says feed forward network with a linear output layer and one hidden layer with non-linearity based squashing function can approximates any function. However the hidden layer may be large leading to generalization failure.
- Reduction in number of units can happen in deep models to represent the desired function get a better and generalization error.
- Exponential Reduction in Computational cost of representing a function and training data needed to learn can be done by increasing the depth of the model.

TABLE I. RESTRICTED BOLTZMANN MACHINE AND AUTO ENCODERS

RBM	Auto encoder
Stochastic model setting with symmetric connectivity b/w the visible and hidden layers.	Deterministic model with two weight matrices $w_1$ and $w_2$ representing the flow of data from the visible-to-hidden and hidden-to-visible layers.
Energy based models.	Achieved considerable success via de-noising the input.
Trained using contrastive divergence that performs Gibbs sampling and is used inside a gradient descent procedure.	Trained to perform optimal reconstruction of the visible layer by minimizing the mean-squared error in a reconstruction task.
Preferred in High Noise Scenarios & Speech Recognition.	Preferred in Low Noise Scenarios.

- Deep auto encoders yield better compression than shallow or linear auto encoders [13].

G. RBM and Auto Encoders

Restricted Boltzmann Machines and Auto encoder models; both are used for training deep architectures using an unsupervised greedy layer wise pre training followed by supervised fine tuning. Table I compares these two models.

III. SEMI-SUPERVISED LEARNING

Semi-supervised learning a class of machine learning techniques that has the capability of utilizing small volume of labeled data with a large volume of unlabeled data; leading to substantial improvement in learning precision. Thus it therefore falls between unsupervised learning and supervised learning realm. As the acquisition and labelling cost of labeled data for a specific learning problem is high making the dataset set infeasible, whereas acquisition of unlabeled data is relatively inexpensive. These settings make semi-supervised learning of great technical and practical value.

It is worth noting that supervised learning has achieved good results as opposed to unsupervised learning. The reason being that implementations of unsupervised learning are not are incompatible with supervised learning. Supervised Learning processes filter out non relevant information preserving only the important features where as unsupervised learning methods try retain and represent as much information about the original data as possible. This is where semi-supervised learning comes into play, integrating both supervised and unsupervised learning together in a novel architecture - The Ladder Networks.

A. Ladder Network An Autoencoder with shortcut connections

The Ladder Network model is an autoencoder with adjacent shortcut connections from the encoder to decoder at each layer. These connections let the higher layers to focus on abstract invariant and consistent features. Comparatively standard auto encoders are equivalent to latent variable models with a single layer of stochastic variables only; however the ladder network is equivalent in strength to hierarchically ranked latent variables models [14].

Ladder networks combine supervised learning with unsupervised learning in deep neural networks. As stated before, in

classical setting, unsupervised learning was used only for pre-training the network this was followed by supervised learning. However ladder networks integrates the two together. Similar to feedforward networks, learning occurs via minimizing the relevant cost function. Another important aspect is higher layers can focus on consistent features only leaving the details for the lower layers to represent. Classically unsupervised learning has been used for pre-training prior to supervised learning. However here, it continues to work after Supervised Learning has commenced. Relevant features are selected via supervised learning using labelled samples while unsupervised learning selects correlated features using bulk of unlabeled data. This improves generalization to new samples.

It is important to note that in this integrated model, once supervised learning starts selecting significant features, unsupervised learning only focuses on and selects co-related features which are useful for supervised learning. This characteristic is completely in contrast with the classic pre-training approach where unsupervised learning selected all the features via which input could be re generated.

*B. Stochastic Latent Variable Models and Deterministic Autoencoders*

Unsupervised learning methods where latent variables generate observed data are called Latent variable models. These have the capability of forecasting the mean of the observations. Minimizing the mismatch between the observed data and its reconstruction results in inference of the unknown latent variables.

Single layer latent models do not discard information and represent everything about the data; discarding information in a single layer model would increase the reconstruction error. For better reconstruction error management, subsidiary information along with the abstract consistent features is also required. The solution being using latent variables in a multi-layer hierarchy. This leads to higher layer levels to focus on abstract consistent features and leaving the details to lower levels. Finally higher-level latent variable models can easily represent the mean of the lower-level variables. This hierarchical arrangement provides stochasticity to latent variables but with a forewarning. Computing the posterior probability of the latent variables and their parameters is mathematically inflexible. Another problem with these models is the extensive usage of variational Bayesian probabilistic methods or training leading to a significant compromise in their performance.

Autoencoders are equivalent in representational power to single-layer latent variable models. Learning is based on minimizing the difference between the observation and its reconstruction. Similar to latent variable models, auto encoders have the capability of being stacked together. Training follows a layer wise approach as new layers are added to the previously trained network. After this, training later continues in a supervised manner, [15]. An important point to be noted is a hierarchical/stacked auto encoder is not equivalent in representational power to the hierarchical latent variable model counterpart. The difference being that the middle layers in an auto encoder are strongly deterministic while the hierarchical latent variable model has complete stochasticity. To elaborate more, regardless of whatever the priors are, stochastic variables

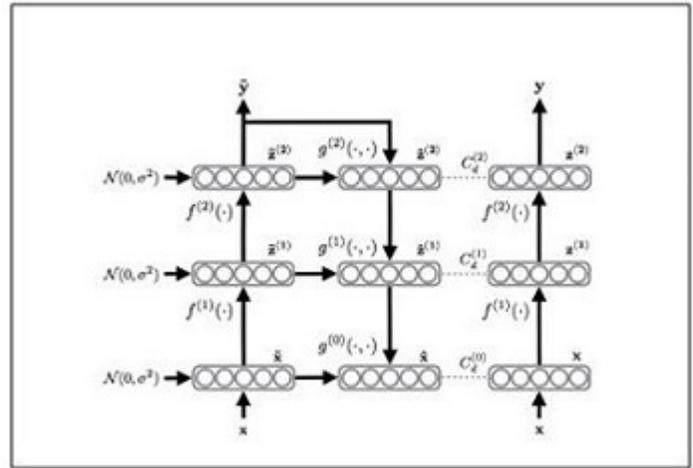


Fig. 3. A two layer ladder network [15].

have independent representational aptitude and thus can add information at the time of reconstruction. In stark contrast, deterministic variables cannot add any information at reconstruction time [16].

This clearly shows that both autoencoders and latent variable models have their pros and cons. This paves way for a more robust model that has the capability of combining the strengths of the two models making it an ideal candidate for deep semi-supervised learning.

*C. The Ladder model*

The ladder network attempts to combine best of the breed from both Latent Variable Models and Stacked Auto encoder models. As is evident from the equations,  $h(t)$  in an auto encoder depends only on top-down information and cannot add any new information to the representation because it does not receive information from the bottom-up path. On the contrary inference of  $s(l)(t)$  in the latent model combines information from top-down priors and bottom-up likelihood.

The Hierarchical Latent Variable Model can therefore be expressed as:

$$p(s^l(t)|s^{l+1}(t), \xi^l) \tag{3}$$

The Stacked Auto encoder can be expressed as:

$$h^{l-1}(t) = g^l(h^l(t)) \tag{4}$$

Likewise the Ladder Network can be expressed as:

$$h^{l-1}(t) = g^l(h^l(t), h^{(l-1)}(t)) \tag{5}$$

Transformation of an Autoencoder into a Latent Model when shortcut adjacent connections are added from the bottom-up encoder path to the modified top-down decoder path. This empowers the hidden layer to recover information which is missing from the higher layers, thus relieving higher layers to represent all the details. The mapping function  $g$  in the Ladder Model combines abstract information from higher layers with detailed information from lower layers (Fig. 3).

#### D. Distributed Layer-wise Learning

Deep networks have error functions only at the output layer. Keeping this in mind if similar training procedure is adopted for Ladder Networks as with Auto encoders; this leads to a considerable increase in the reconstruction error as the lateral connections also contributes to the reconstruction, leaving minimum error contribution from the higher layers which are the custodians of most consistent information. Auto encoder model has the inherent capability of routing all the signals through all layers and making learning difficult and slow. However latent variable models have cost functions for all stochastic variables. As the ladder network combines its structure from both these models, it introduces training signals at each level of the hierarchy exhibiting the novel feature of distributed learning.

The cost Function of the ladder network can be expressed as:

$$C = \frac{1}{T} \sum_{t=1}^T \|s(t) - s'(t)\|^2 = \frac{1}{T} \sum_{t=1}^T \|f^1 x(t) - s'(t)\|^2 \quad (6)$$

$$C^l = \frac{1}{T} \sum_{t=1}^T \|h^l(t) - h^l(t)\|^2 \quad (7)$$

The first cost function of the ladder network expressed as (6) does not directly refer to the input but only to the latent variable and its corrupted counterpart. This leads to the fact that each layer in the hierarchical model contributes to the cost function, bringing the basis of training signals close to the parameters on the relevant layer. The 2nd function expressed as (7) shows that error term can be used at each layer, thus leading to Layered Learning novelty. Layer wise denoising is also used in the Model. The idea taken from [13] corrupting the input of auto encoders with noise and let the network attempt to reconstruct the original uncorrupted inputs. This forces the auto encoder to learn how to denoise the corrupted inputs. Bengio stated denoising not only the inputs but on all levels of the encoder path in a hierarchical model results in efficient sampling; and called such networks generative stochastic networks (GSN) [17]. Another important aspect is that a model with reconstruction capability of missing data can be turned into a probability density estimator.

#### E. Implementation

A typical implementation of the ladder network include:

- 1) A feedforward model which serves supervised learning with 2 encoder - clean and corrupted. The corrupted encoder adds Gaussian noise at all layers.
- 2) A decoder which has the capability to invert the mappings on each encoder layer and provisions unsupervised learning. Decoder uses a denoising function to reconstruct the activations of each layer. The target at each layer is the clean version of the activation where as the difference between the reconstruction and the clean version serves as the denoising cost of that layer.
- 3) The supervised cost is calculated from the output of the corrupted encoder and the output target. The

unsupervised cost is the sum of denoising cost of all layers scaled by a hyper parameter. The final cost is the sum of supervised and unsupervised cost.

- 4) Train the whole network in a semi-supervised arrangement using standard optimization techniques.

#### F. Experiments and Results

TABLE II. ERROR REPORTED WITH 40 LABELS ON (SET 1)

Initial Learning rate = 0.002 Annealed linearly to zero Decay Rate = 0.67 Layers 784-1000-500-250-250-10 Epochs = 150 / Dataset = MNIST				
Models/ No of Labels	40	100	1000	All
Full Model	1.27	1.07	0.71	0.6
Sigma (Top) Model	3.9	3.4	1.9	0.9
Bottom Model	1.30	1.09	1.0	0.71

TABLE III. DENOISING VALUES FOR 6 LAYERS (SET 1)

Denoising Cost configured on Layer wise basis			
Models/ No of Labels	100	1000	All
Full Model	1000,10,0.1, 0.1,0.1,0.1, 0.1	2000,20,0.1, 0.1,0.1,0.1, 0.1	1000,1,0.01, 0.01,0.01,0.01, 0.01
Sigma (Top) Model	0,0,0,0,0,0.5	0,0,0,0,0,10	0,0,0,0,0,2
Bottom Model	5000,0,0,0,0, 0,0	2000,0,0,0,0, 0,0	2000,0,0,0,0, 0,0

TABLE IV. ERROR REPORTED WITH 100 LABELS (SET 2)

Initial Learning rate = 0.002, Annealed linearly to zero Decay Rate = 0.67 Layers 784-1000-500-250-10 Epochs = 150 / Dataset = MNIST			
Models/ No of Labels	100	1000	All
Full Model	1.22	0.852	0.72
Sigma (Top) Model	4.08	1.28	1.08
Bottom Model	1.308	1.23	0.852

The Tables II, III, IV summarize the results of the experiments conducted in this research. Table II shows the results obtained using a 6 layers model while Table IV shows the results obtained using a 4 layers model. The Table III shows the layer wise denoising values used. These results are further elaborated in the Inferences section below.

#### G. Inferences

- 1) Table II shows, with 60000 labels the error rate is around 0.6, while with 1000 labels its 0.71, with 100 labels it is 1.06 and with merely 40 labels its 1.27.
- 2) The Top Model in Table II has also impressive results of 3.4 with 100 labels. It has cost function at only top layer, thus most of its denoising part can be bypassed. It has the capability of being plugged in any neural network.
- 3) Reducing the number of layers from 6 (Table II) to 4 (Table IV), increases the error by around 21
- 4) Top and bottom models have denoising cost at the top and bottom layer only as shown in Table III.
- 5) Applying noise to each layer and specifically to the first layer leads to regularization thus minimizing the generalization error as shown in Table III.

#### IV. CONCLUSION

This exploratory research highlights the impact ladder networks has brought upon deep unsupervised learning by alleviating pre-training and successfully synthesizing the two forms of learning which are usually considered under exclusivity. As for semi-supervised learning, it can be concluded that the most important contribution is made by the lateral adjacent connections. These connections being a vital component to the level that removing them declines the performance for all of the semi supervised tasks. The second important contribution is introduction of noise at each layer. As the number of labeled examples increases, the adjacent connections and the reconstruction criterion become less significant and the generalization enhancement coming from the induction of noise in each layer and vice versa [18]. Ladder networks have transformed the neural network model from a label learner to a structural observer, thus narrowing the gap between machine learning and machine intelligence.

In our opinion, the ladder network could be applied in these future directions: data generated in high velocity environments are unlabeled, which makes training models difficult; we envisage to use the Ladder model exhibiting semi supervised learning in these settings. Semi-supervised models require large training time, optimizing it is a potential challenge. Ladder Networks have shown promising results with low dimensional data sets, so their use on high dimensional data sets is a possibility for exploration.

#### REFERENCES

- [1] Harri Valpola. From neural PCA to deep unsupervised learning. In *Adv. in Independent Component Analysis and Learning Machines*, pages 143171, Elsevier. arXiv:1411.7783, 2015
- [2] Antti Rasmus, Tapani Raiko, and Harri Valpola. Denoising auto encoder with modulated lateral connections learns invariant representations of natural images. arXiv:1412.7210, 2015
- [3] Krizhevsky, A., I Sutskever, and G. E Hinton. ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, pp. 11061114 , NIPS, 2012
- [4] Hinton, G. and R. Salakhutdinov. Reducing the dimensionality of data with neural networks. In: *Science* 313.5786, pp. 504507, 2006
- [5] Pascal Vincent, Hugo Larochelle, Yoshua Bengio. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. In *Journal of Machine Learning Research*, 2010
- [6] Temporal Auto encoding Restricted Boltzmann Machine. Chris Hausler and Alex Susemihl. arXiv:1210.8353v1, 2012
- [7] Ian Goodfellow, Yoshua Bengio and Aaron Courville. The Deep Learning textbook, An MIT Press book. Under Publication, 2016
- [8] Salakhutdinov, R. , Mnih, A. , Hinton, G. Restricted Boltzmann machines for collaborative filtering. *Proceedings of the 24th international conference on Machine learning - ICML 07.* p. 791. ISBN 9781595937933,doi:10.1145/1273496.1273596, 2007
- [9] Miguel A. Carreira-Perpinan and Geoffrey Hinton. On contrastive divergence learning. *Artificial Intelligence and Statistics*, 2005
- [10] Bengio, Y. Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning*. 2. doi:10.1561/22000000006, 2009
- [11] Yoshua Bengio, Li Yao, Guillaume Alain, and Pascal Vincent. Generalized denoising auto-encoders as generative models. In *Advances in Neural Information Processing Systems 26*, pages 899 907, 2013
- [12] Cybenko., G. Approximations by superpositions of sigmoidal functions, *Mathematics of Control, Signals, and Systems*, 2 (4), 303-314, 1989
- [13] Salah Rifai, Pascal Vincent, Xavier Muller. Contractive Auto-Encoders: Explicit Invariance During Feature Extraction. In *Proceedings of the Twenty-eight International Conference on Machine Learning*, 2011
- [14] Oja, E. The nonlinear PCA learning rule in independent component analysis. In: *Neurocomputing 17.1*, pp. 2546, 1997
- [15] Antti Rasmus, Harri Valpola, Mikko Honkala, Mathias Berglund, and Tapani Raiko. Semi-supervised learning with ladder networks. arXiv preprint arXiv:1507.02672, 2015
- [16] Ilin, A. and H. Valpola. On the effect of the form of the posterior approximation in variational learning of ICA models. In: *Proc. of the 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (ICA2003)*. Nara, Japan, pp. 915920, 2003
- [17] Yoshua Bengio, Eric Thibodeau-Laufer, Guillaume Alain, Jason Yosinski. Deep Generative Stochastic Networks Trainable by Backprop. arXiv:1306.1091, 2013
- [18] Mohammad Pezeshki, Philemon Brakel, Aaron Courville, Yoshua Bengio. Deconstructing The Ladder Network Architecture. Under review as a conference paper at ICLR, 2016

# Privacy-preserving Twitter-based Solution for Visually Impaired People

Dina Ahmed Abdraboo

Faculty of Computers and Informatics  
Suez Canal University  
Ismailia, Egypt

Tarek Gaber

Faculty of Computers and Informatics  
Suez Canal University  
Ismailia, Egypt

Mohamed El Sayed Wahed

Faculty of Computers and Informatics  
Suez Canal University  
Ismailia, Egypt

**Abstract**—Visually impaired people is a big community all over the world. They usually seek help to perform their daily activities such as reading the expired date of food cans or medicine, reading out PIN of a certain ATM Visa, identifying the color of clothes or differentiate between the money notes and other objects with the same shape. A number of IT-based solutions have been proposed to help and assist blind and/or visually impaired people. Generally speaking, these solutions, however, do not support Arabic languages nor protect blind users' privacy. In this paper, Trusted Blind Society (TBS) mobile application is proposed. It is an android application which allows blind users to recognize their unknown surroundings by utilizing two concepts: social networks sites and friendsourcing. These two concepts were employed by allowing family members and the trusted friends, who are registered on Twitter, to answer blind users' questions on a real time. The solution is also bilingual, supports (Arabic/English) and allows screen reader using Android talk-back service. The performance of the TBS system was evaluated using loader.io to check its stability under the heavy load and it was tested by a number of blind volunteers and the results showed good performance comparing to most related work.

**Keywords**—Human powered technology; blind people; visually impaired people; user's privacy; IT-based solution; social networks; friend-sourcing; crowd-sourcing; accessibility; low vision; bilingual; screen reader

## I. INTRODUCTION

The World Health Organization (WHO) reported that 285 million people around the world are visually impaired (39 million are blind and 246 have low vision) [1]. The huge number of the blind people and the evolution of technology encourage the researchers to introduce solutions to assist blind/visually impaired people to normally practice their daily activities. In the other hand, the social network sites such as Facebook, Twitter, and Instagram, are widely used nowadays and the users of them are growing rapidly. Social networking sites (SNSs) are sites that asset people to communicate together in many ways via the Internet. The target of these websites are not only to interact with each others and making friends but also to gather the information and getting help from known people and share the experiences with them [2].

Arab Social Media Report (ASMR) reported that there are 3.7 million Arab Twitter users [2]. As there are huge number of people in using SNSs and the increase of the Arabic visual impaired, these two important facts motivate us to design and introduce an Arabic application supported

by human powered technology while assisting the visually impaired users in identifying the unknown surrounding objects.

Recent solutions are proposed using human-powered technology for assisting the blind such as VizWiz [3]. The VizWiz App allows users to capture an unknown object only one shot to then send to crowd worker who should response with the name of this object. However, using only one shot is not enough because the blind user can't determine the dimensions of the image needed to be recognize. In addition, it is not easy for the blind to capture a clear picture of the unknown object. Other solutions such as Chorus [4], Third Eye [5], they use video to address the limitation of VizWiz but they do not support consumer's privacy as they employ crowd-sourcing concept. In other words, they enable the blind to send the video over Internet and any anonymous helper will get it and help him. Thus, the user shouldn't send any private data to save his privacy.

Human powered technology (friend-sourcing) is the major part of our solution as it considered the identification process while supporting privacy protection to the blind users as well as providing accurate responses to the blind people. Indeed, they can ask anyone for the unknown object but there were some confidential data and situations that he can't share with any strangers except close friend or a family's member [6].

In this paper, Trusted Blind Society (TBS) application is proposed. The TBS is Android-based mobile application enabling blind users to recognize their unknown surroundings by utilizing two concepts: online social networks and friend-sourcing. These two concepts were employed by allowing the family members and the trusted friends who are registered on Twitter to answer the blind's private questions on a real-time while other Twitter friends could answer any other questions. The solution also supports two languages (Arabic/English).

This paper is organized as follows. Section II presents the related works while Section III provides the details about the proposed solution. Section IV reports the evaluation of the proposed solution, and Section V presents the discussion. Finally, Section VI gives the conclusion and future work.

## II. RELATED WORK

Technology trend aims to found the full automated solutions. However the advances in the technology can't stand alone in facing blind needs such as identification of unknown

objects supporting their daily activities. For example, OCR failed to read the text on a road sign that captured by mobile phone<sup>1</sup>, object recognition technique worked well with sighted people. However it doesn't work effectively with the blind users. This is because the captured area often isn't clear for the camera [3, 7, 8]. Moreover, screen reader software use some techniques to deliver the content of the screen to the blind, the content is always confusing or unreliable for the blind [9, 10].

Instead of automated solutions, disabled people depend on a wide network of friends, volunteers or strangers. An example of that, the volunteer can perform a given task (reading a book for blind) for a certain time of the day but not all the time. Technology evolution engaged the human intelligence and introduces new solutions to assist blind in their daily situations instead of the physically proximate for the helper with the blind. Human-powered access technology [11] has introduced as technology solutions that facilitate the human assistant to help disabled people remotely with flexible support rather than the fully automated solutions that almost not complete or inadequate in daily real situations [12]. Expansion of the human powered solution makes it widely used, therefore researchers deeply study it and classified it into two major concepts: crowd-sourcing and friend-sourcing. The crowd-sourcing enables users to get help from anonymous helpers while friend-sourcing allows the users to get help from known friends.

Indeed, many solutions are proposed to help blind. These include ThirdEye [5], VIZWIZ [3] and LendAnEye [13], and others [6, 13–15]. These solutions make use of the integration between human resource and the information technology. An overview of these solutions is given below.

- **VizWiz**  
Bigham et al. [3] introduced an application to address the challenges that faced blinds in their everyday activities (the ability to see food products expired date, recognize ATM ID, etc.). Therefore, they presented an iPhone application called VizWiz that allows blind users to send a picture along with a recorded question about their the unknown object to workers then receive quick answers which are read to the blind users by converting the text to speech. However, VizWiz is not a time efficient solution as sometimes the picture is not clear, thus the helper asks the blind to recapture the image, and repeat the processing again. Moreover, a helper sometimes explains the image content without listening to the question carefully, so the blind would need to repeat the questions again. In addition to the lack of workers available during some time a day. Also, the malicious workers are considered one of the most limitation in this approach [15].
- **Third Eye**  
Lamya et al. [5] discussed the problem for the visually impaired as they faced challenges when practicing their daily tasks. Then, they proposed an mobile application which provides descriptions of all the pictures or videos captured by visually impaired people. The blind can capture an image or record a video

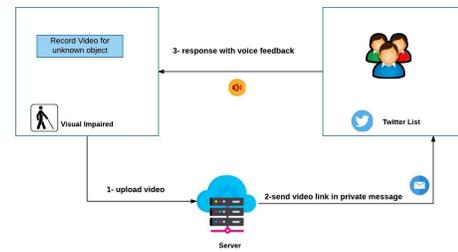


Fig. 1. Data flow of the proposed solution.

about the unknown object then sends it through the server and the volunteer will view and then send his/her feedback as speech to the blind user again. This proposed application presents good performance in recognition as it depends on crowd-sourcing for serving visual impaired requests. However, the privacy factor is not covered as the blind user doesn't know who is going to answer his/her request as it could be stranger volunteers. If the object contains confidential data, such as credit card PIN, it will be known to this strangers.

- **LendAnEye**  
LendAnEye [5] is proposed to introduce a solution for blinds in their daily activates to identify environment surround them. It's a mobile application that has two interfaces one for volunteers and the other one for low-vision user. Un sighted users can contact helpers by double click the screen then a live video call is started, and helper can communicate with the blind via a continuous video call that enhances reality and highly response time. However, the privacy of blind is not guaranteed because helpers are anonymous strangers people the privacy of his request is not achieve. Also this require involving video call service which could be expensive to most of the blind users in the developing countries.

### III. PROPOSED SOLUTION

Mobile Application called: Trust Blind Society (TBS) is proposed which is an android mobile application that assists the blind in identifying the confidential data. As shown in Fig. ??, the idea of this application is based on the friend-sourcing concept while providing help. The user can record a video about the unknown object that he needs to recognize and then sends it over the social network (Twitter) to his selected friend's list. Once the helper identifies the unknown object, he will reply by a message through Twitter, then the TBS will read out the content of the message to the user (using talk back). TBS will access the message of the user Twitter's account through (Twitter API). TBS offers video recording because it's more effective and accurate than the image as the video recording guarantee that the whole range of the unknown object is recorded besides the image may be blurred so the identification process will consume time [3].

TBS is based on friend-sourcing (family members and close friends) which provides the protection of the blind users' privacy as only family members or close friends could see

<sup>1</sup>Twitter statistics. 2014, <http://www.statisticbrain.com/twitter-statistics/> (accessed May 2015)

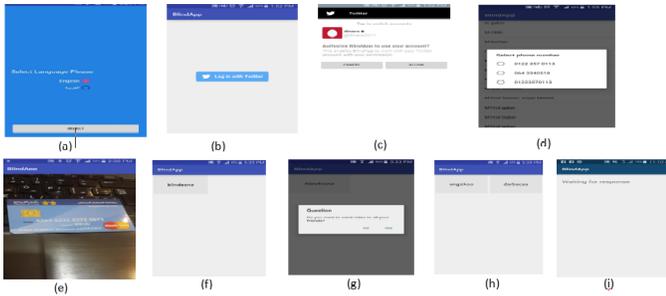


Fig. 2. TBS screen shots.

blind users' private data such as credit card numbers, emails passwords, accounts, private situations that may embarrass him and any private data in his surrounding environment. In other words, TBS integrates the features of (a) Android OS (i.e. Talkback that considered the screen reader), (b) Twitter (i.e. while broadcasting the video to helpers), and (c) friend-sourcing (which is considered the helper side). The implementation of the TBS is done using two main APIs:

- PHP developed APIs: we developed our own APIs for uploading the recorded video to a server.
- Twitter APIs: we embedded twitter API into our application for the social communication while it passes the requested video from the blind user to the helper then get the response back through private message.

These are the main steps, also depicted in Fig. ??, showing how TBS is working:

- 1 Login to the user Twitter account (through twitter API), to enable the user to access his account data.
- 2 Select his preferred language (Arabic - English)
- 3 Select his prefer Twitter's List/ Friends or family members (helper side)
- 4 Select his featured friend from his phonebook that is embedded in our solution so that when the user has no response up to 30 seconds, a SMS will be sent to his selected friend.
- 5 Record his video that includes the unknown object and then send it out to the Twitter's list.
- 6 Read messages for certain conversation as long as the user is log into his account.

#### IV. EVALUATION

The evaluation process for the proposed solution is held to determine if TBS is useful and easy to use or not. The evaluation is done in two ways: target users satisfaction and performance test. Users satisfaction is done to test the functionalities of the solution, the ease of use and accuracy. The performance test is done by using a stress test called, loader.io, to check the response time, evaluate the robustness and availability under the heavy loads.

For the user's satisfaction, 35 blind participants, 13 girls and 22 male, are asked to join our evaluation questionnaire while using our proposed solution. Blinds will access the application trying to identify the unknown object then they

will give their feedback about the proposed solution through our questionnaire which consists of the following questions:

- Questionnaire
  - 1 Name, Age
  - 2 Education level
  - 3 Experience with android
  - 4 Is the solution easy to use?
  - 5 Do you think the Application is useful in your daily activities?
  - 6 What do you think is more helpful photo or video?
  - 7 Give some examples in which situations you could use our solution.
  - 8 Rate the solution.
  - 9 Addition ideas to enhance any solution needs more ideas and maintained of it , so this question will assist us in the future work to enhance and present well functional solution.

- Responses

The results of the survey are shown in Fig. ?? . Eleven blind users reported that they strongly agree that the application is easy to use and user friendly, while Seven of them agree with this. Only two users are disagree with it because they aren't connected to twitter in addition they suggested to supply them with image recognition technique which would be better for them as they don't want to depend on human factor. In another hand, all of the blind testers agreed that TBS is very secure and save their privacy. So, it could be said that almost 90% find that TBS is easy to use. Moreover, all of them agree that this application will assist them in their daily activities as they have tested our applications with different examples such as:

  - Money identification.
  - Password identification.
  - Is the makeup matched with the girl?
  - Differentiating between colors and paper money.
  - If the blind's hit himself he can ask his close friend if it's ok or not).
  - Differentiating between boxes that have similar shapes and sizes, like soft drink cans and juice boxes.
  - Differentiating between medicine containers and eye drops.
  - Checking the expiration dates of different products.

The following is screen shots from videos that recorded via the blind testers (Fig. 3). In addition, all of the respondents reported that recording a video is useful for them as the video give wide range of the target and they can say all their detailed question while recording (Fig. 5).

- Performance evaluation

Stress test is used to check the performance of our solution under the heavy load of users. We applied it using loader.io software. Stress testing aims to achieve the best performance for websites such as set-up static caching servers such as Varnish that enables the software to serve much more users with the same

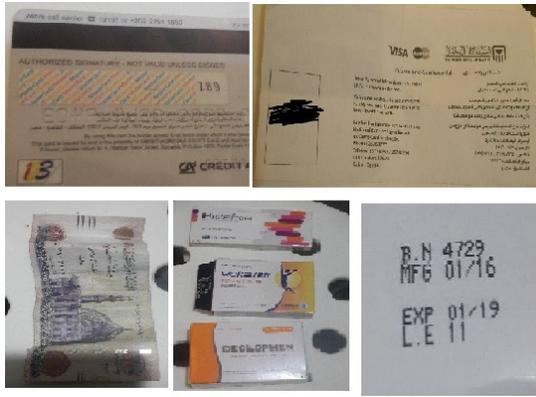


Fig. 3. All screen shots.

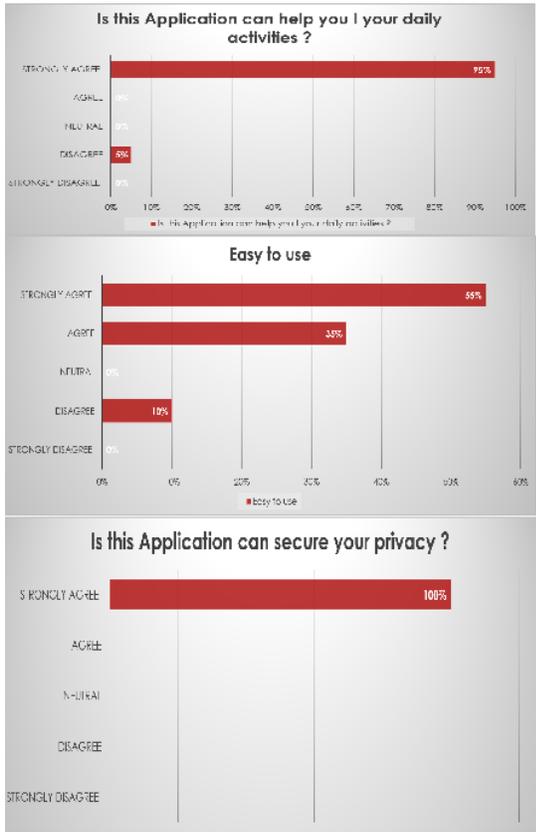


Fig. 4. Evaluation: All evaluation screen shots.

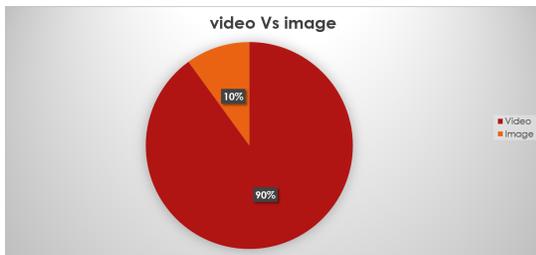


Fig. 5. Evaluation: Video Vs Image.

hardware. Thus the optimization of the website can be held via load testing. Stress testing can be held through



Fig. 6. Stress test simulation about TBS solution by Loader.io.

a new free software tool called Loader.io which is a FREE load testing service that allows you to perform stress test for your web-apps and APIs with thousands of concurrent connections. It simulates connections to our application for certain duration to Monitor and test if the application will perform sufficiently under the load on the application so that the response time can be examined and the availability under the heavy load of the connected people to the application [16]. In Addition, it can measure the response time of our application. We examined if the application can be maintained in a very good way to respond and work well under all circumstances such as (the heavy load of connected users) [17]. The result of applying the stress test to our application is showing in Fig. ??:

## V. DISCUSSION

The successful results of using the TBS application by blind users means that our TBS app is blind user friendly. This achieved by using auto-focus camera allowing the blind users to feel the vibrating of the mobile when the object is in the camera focus. It was noticed that 30 persons were successfully recorded a video stream about the desired object from the *first trial*, and 3 of them were successfully completed this cycle from the *second trial* and two blind users successfully record a video after three trials. it could be said that that the TBS application is easy to use and has friendly interface. Choosing the android as mobile OS was to support the fact that the majority of smart-phones is android-based and they are cheap too. This fact was confirmed by the the volunteers who appreciated the fact that the TBS is an android based smart-phone as it is now available with low prices which support the economic case of most blind peoples in Egypt and the Middle East. Moreover, the proposed TBS application is supported with Arabic. So, the TBS application will be much affordable than related applications which are only working at iOS-based smart-phones and only supporting English language.

TABLE I. COMPARISON BETWEEN FEATURES OF RELATED SOLUTIONS

Feature	VIZWIZ	Chrous-View	ScanSearch	ThirdEye	Legion	SmartEye	EasySnap	LendAnEye	TBS
Add Video	X	✓	✓	✓	X	X	X	✓	✓
Based on friend sourcing	X	X	X	X	X	X	X	X	✓
Bilingual	X	X	X	✓	X	X	X	X	✓
Pre-defined known helpers	X	X	X	X	X	X	X	X	✓
Support Android	X	✓	X	X	X	X	✓	X	✓
Free service	X	✓	✓	✓	✓	✓	✓	✓	✓
Send SMS to pre-defined number	X	X	X	X	X	X	X	X	✓

A comparison with most related work is conducted and its results are shown in Table ??. From these results, the following

remarks can be drawn.

- Privacy Privacy factor is not covered in all the previous work. All these work are based on crowdsourcing (anonymous helpers who are not trusted by the blind users). So, they could not be comfortable and feel non-secured in case of identifying confidential things.
- Supporting offline mode All previous related work are based on the availability of the Internet. However, in the developing countries, the availability of the internet isn't guaranteed. In our solution, this problem has been addressed through sending the video to a close friend's mobile phone using video message.

## VI. CONCLUSION

IT-based solutions supporting the daily life activities of visually impaired people are the most important due to a large number of blind/visually impaired people all over the world. However, most of the current solutions do not support the user's privacy. In this paper, we proposed an assistive Mobile solution called: Trust Blind Society (TBS), which is an Android application that allows blind users to recognize their unknown surroundings while protecting their privacy. This solution utilized two concepts: online social networks (i.e., Twitter) and friendsourcing. In this solution, the close friend can answer the confidential questions (e.g., bank card security question) while the public friend answers the other questions (e.g., identifying currency notes or food cans). Thus, by utilizing the friendsourcing (friends on Twitter), the privacy of the blind user will be protected. This TBS solution has to be implemented using built-in Android functionalities such as the auto-focus camera for video recording and talk-back accessibility service to read out what is on the screen reader. It also made use of Twitter APIs to handle the connection between the blind users and his/her helpers who have accounts on Twitter. The solution is evaluated in terms of its usability by 35 blind testers. The results were very promising and acknowledged by the testers. Also, the app performance was tested using stress test (loader.io) and the results proved that TBS is stable under the heavy loads and the average response time was 320 ms. In the future, different features can be added to the TBS App and these include: adding automated identification algorithm for basic images, colors and money, supporting iOS version so that the application will be compatible with the iPhone.

## REFERENCES

- [1] M. El-Gayyar, H. F. ElYamany, T. Gaber, and A. E. Hassanien, "Social network framework for deaf and blind people based on cloud computing," in *Computer Science and Information Systems (FedCSIS), 2013 Federated Conference on*. IEEE, 2013, pp. 1313–1319.
- [2] F. Salem, R. Mourtada, and S. Alshaer, "Transforming education in the arab world: Breaking barriers in the age of social learning," *Arab Social Media Report, Dubai School of Government–DSG*, 2013.
- [3] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White et al., "Vizwiz: nearly real-time answers to visual questions," in *Proceedings of the 23rd annual ACM symposium on User interface software and technology*. ACM, 2010, pp. 333–342.
- [4] W. S. Lasecki, P. Thiha, Y. Zhong, E. Brady, and J. P. Bigham, "Answering visual questions with conversational crowd assistants," in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 2013, p. 18.
- [5] L. Albraheem, R. AlDosari, S. AlKathiri, H. AlMotiry, H. Abahussain, L. AlHammad, and M. Alshehri, "Third eye: An eye for the blind to identify objects using human-powered technology," in *Cloud Computing (ICCC), 2015 International Conference on*. IEEE, 2015, pp. 1–6.
- [6] Y. Zhong, P. J. Garrigues, and J. P. Bigham, "Real time object scanning using a mobile phone and cloud-based visual search engine," in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 2013, p. 20.
- [7] R. Manduchi, J. Coughlan, and V. Ivanchenko, "Search strategies of visually impaired persons using a camera phone wayfinding system," *Computers Helping People with Special Needs*, pp. 1135–1140, 2008.
- [8] H. Samuel White and J. P. Bigham, "Easysnap: Enabling blind people to take photographs," *UIST 2010-Demos*, 2010.
- [9] Y. Borodin, J. P. Bigham, G. Dausch, and I. Ramakrishnan, "More than meets the eye: a survey of screen-reader browsing strategies," in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*. ACM, 2010, p. 13.
- [10] Y. Borodin, J. P. Bigham, R. Raman, and I. Ramakrishnan, "What's new?: making web page updates accessible," in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2008, pp. 145–152.
- [11] S. Kawanaka, Y. Borodin, J. P. Bigham, D. Lunn, H. Takagi, and C. Asakawa, "Accessibility commons: a metadata infrastructure for web accessibility," in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2008, pp. 153–160.
- [12] J. P. Bigham, R. E. Ladner, and Y. Borodin, "The design of human-powered access technology," in *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2011, pp. 3–10.
- [13] W. S. Lasecki, K. I. Murray, S. White, R. C. Miller, and J. P. Bigham, "Real-time crowd control of existing interfaces," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 2011, pp. 23–32.
- [14] S. White, H. Ji, and J. P. Bigham, "Easysnap: real-time audio feedback for blind photography," in *Adjunct proceedings of the 23rd annual ACM symposium on User interface software and technology*. ACM, 2010, pp. 409–410.
- [15] W. S. Lasecki, R. Wesley, J. Nichols, A. Kulkarni, J. F. Allen, and J. P. Bigham, "Chorus: a crowd-powered conversational assistant," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*. ACM, 2013, pp. 151–162.
- [16] K. Keeker, "Improving web site usability and appeal," *Guidelines Complied by MSN Usability Research*. Available at <http://msdn.microsoft.com/library/default.asp>, 1997.
- [17] P. Khanna, S. Jain, and B. Babu, "Distributed cloud brokerage: solution to real world service provisioning problems," *ARPN Journal of Engineering and Applied Sciences*, vol. 10, no. 5, pp. 2011–2016, 2015.

# A Text based Authentication Scheme for Improving Security of Textual Passwords

Shah Zaman Nizamani

Department of Information Technology  
Quaid-e-Awam University of Engineering, Science &  
Technology, Pakistan

Tariq Jamil Khanzada

Department of Computer Systems Engineering  
Mehran University of Engineering & Technology, Pakistan

Syed Raheel Hassan

Department of Computer Systems Engineering  
Quaid-e-Awam University of Engineering, Science &  
Technology, Pakistan

Mohd Zalisham Jali

Faculty of Science and Technology  
Universiti Sains Islam (USIM), Malaysia

**Abstract**—User authentication through textual passwords is very common in computer systems due to its ease of use. However textual passwords are vulnerable to different kinds of security attacks, such as spyware and dictionary attacks. In order to overcome the deficiencies of textual password scheme, many graphical password schemes have been proposed. The proposed schemes could not fully replace textual passwords, due to usability and security issues. In this paper a text based user authentication scheme is proposed which improves the security of textual password scheme by modifying the password input method and adding a password transformation layer. In the proposed scheme alphanumeric password characters are represented by random decimal numbers which resist online security attacks such as shoulder surfing and key logger attacks. In the registration process password string is converted into a completely new string of symbols or characters before encryption. This strategy improves password security against offline attacks such as brute-force and dictionary attacks. In the proposed scheme passwords consist of alphanumeric characters therefore users are not required to remember any new kind of passwords such as used in graphical authentication. Hence password memorability burden has been minimized. However mean authentication time of the proposed scheme is higher than the textual password scheme due to the security measures taken for the online attacks.

**Keywords**—Password security; security; usability; alphanumeric passwords; authentication

## I. INTRODUCTION

Despite of many weaknesses user authentication through textual passwords is widely used since long time. In textual password scheme credentials are directly inserted into login fields, which results in easy capture of password through spyware attack, and shoulder Surfing attack [1]. Other problem with textual password scheme is that users tend to set short and easy to remember passwords, such passwords are easy to break through brute force or dictionary attack [2]. Therefore users are restricted to add numbers or special characters in their passwords but such policies make the passwords hard to remember.

By recognizing the memorability and security issues in textual passwords, researchers proposed different graphical password techniques. In this category of authentication passwords are consist of some pictures, lines or x, y coordinates inside a picture. Generally graphical passwords have memorability advantage over textual passwords because visual information is easy to remember and recall than alphanumeric characters [3] [4]. While security and usability of graphical password techniques varies from one scheme to another.

Graphical password technique was first proposed by Blonder [5] in 1996, since then many graphical password techniques are proposed but none has replaced textual password scheme. Shoulder surfing and spyware attacks are common threat to different graphical password schemes. Android unlock scheme [6], is the only graphical password scheme being largely used in smart phones because the scheme is easy to use. Although this technique has many security weaknesses such as shoulder surfing attack but due to nature of the device, attackers have very little access to launch security attacks. Due to security weaknesses Android unlock scheme is not used in online systems for authentication. Secure graphical password schemes have timing and adoptability issues. Such schemes require large amount of physical and mental work to do for authentication and users have to remember different kinds of passwords that is why many usability issues arises.

User authentication can be made secure by biometric or token based authentication techniques but they require special hardware for processing. The other easy to use authentication option remains the knowledge based technique. Authentication through this technique is improved by two approaches. In first, different graphical password schemes have been proposed, while in second approach schemes are suggested by enhancing or mixing text based and graphical password techniques. In this paper second approach has been taken for improving the security of traditional textual passwords. Proposed scheme provides enhancements in the login screen and the way passwords are stored into the database. In the login screen every time user inserts a new set of numbers which

represent the password, therefore proposed scheme provides resistance from spyware and man-in-the-middle attacks. In the proposed password storage technique, alphanumeric characters of a user's password are transformed to different alphanumeric characters and symbols and then stored into database. This password transformation makes harder to apply dictionary and brute force attacks.

The remaining paper is divided into six sections. In section 2 literature review is given regarding the field of user authentication. Proposed authentication scheme along with technique to store passwords are explained in section 3. In section 4 analysis of the proposed scheme is given with respect to security, usability and memorability. Proposed scheme is compared with famous authentication techniques in section 5. Finally conclusion is given in section 6.

## II. LITERATURE REVIEW

User authentication works on the basis of something user knows (Knowledge-based), something user has (token-based) or something user are (Biometric). Focus of this research is to design an efficient user authentication scheme under the category of knowledge-based authentication. Therefore literature review targets knowledge-based authentication. This section has been divided into two parts in first, different user authentication schemes are discussed which are related with the research work. In the second part, problems in user authentication schemes are briefly discussed.

### A. Related work

Zhao and Li [7] proposed some changes in textual password scheme for adding resistance to shoulder surfing attack and called the scheme as S3PAS. In this scheme registration process is same as textual password scheme but the login process is different. In the login screen alphanumeric characters are randomly shown in the image format and a user has to click on the logical triangles formed by the password elements or type characters which belong to each password triangle. S3PAS scheme provides resistance from shoulder surfing, keystroke logger and mouse logger attacks. Searching password triangles is time consuming task, therefore this scheme is very difficult to use. The scheme is also vulnerable to dictionary and brute force attacks.

Ziran et.al [8] proposed a text-based password scheme. In this scheme a user set password by drawing a shape inside a registration screen. In the login screen a grid filled with 0s and 1s are randomly shown, a user is required to insert a list of 0s and 1s, such that they form the shape of password. Proposed scheme provides resistance from spyware attacks but the scheme is vulnerable to brute force, dictionary and shoulder surfing attacks.

Chen et al. [9] proposed a mixed textual and graphical password scheme for resisting shoulder surfing attack. In this scheme passwords consist of some characters and numbers along with a colour. In the login screen characters and numbers are shown in circular format. Password is entered by rotating password characters in front of the colour chosen during registration. Proposed scheme does not contain symbols therefore it has small password space and password entry process requires physically efforts.

Rao and Yalamanchili [10] proposed an authentication scheme known as Pair Pass Char (PPC), in this scheme registration process is same as ordinary textual password scheme. In the login screen all alphanumeric characters are shown in  $10 * 10$  grid. For password entry a user has to search logical rectangles, formed by different pairs of password characters and then click on the corner characters of the rectangles. The scheme contain different rules for rectangle searching therefore the scheme is difficult to learn. Average authentication time for 6 characters password is 47.4 seconds which is quite high.

First graphical password scheme was proposed by Blonder [5]. He proposed a scheme where a password consists of certain points inside a password picture. Blonder's scheme has many security issues such as shoulder surfing attack and mouse logger attack. Wiedenbeck [11] proposed "PassPoint" scheme based upon Blonder's scheme. In PassPoint scheme users have freedom to click on any point inside the password picture, this freedom was not available in Blonders scheme. Passpoint scheme is better than Blonders scheme with respect to brute force and dictionary attacks but it is not resilient to shoulder surfing and spyware attacks.

Wiedenbeck et al. [12] proposed a shoulder surfing resilient graphical password scheme known as CHC (Convex Hull Click). In this scheme users are given multiple challenges for authentication. In each challenge users have to find out three password images and then need to click inside an invisible triangle formed by the password images. This scheme provides resistant from shoulder surfing attack but authentication time is 71.66 seconds which is quite high.

Lopez et al. [13] suggested a challenge response based shoulder surfing and spyware attack resilient graphical password scheme. In this scheme three images per row are shown in the login screen. A user has to identify whether number of password images are even or odd in different rows. The scheme is weak with respect to brute force attack because small number of images are used in this scheme. Combined screen scrapper and key logger attack become successful after multiple rounds of recordings.

Weinshall [14] proposed a recognition based graphical password scheme known as cognitive authentication scheme. It provides resistance form key logger and mouse logger spyware attacks. In the the scheme, 80 icons are presented into  $8 * 10$  grid based login screen. Password icons are selected by computing a path generated by the icons. Learnability and high authentication time are the issues with this scheme.

Google introduced android unlock scheme, in which nine points are given into a  $3 * 3$  grid based login screen. Password of the scheme consists of some lines inside the grid. This scheme is very easy to use but the passwords can be captured by shoulder surfing attack and the scheme also provides low password space [6]. Microsoft introduced a graphical password scheme in windows 8, in which passwords consist of some points, lines or circles inside a picture. This scheme is also very easy to use but it has Hot-Spot and shoulder surfing issues [15].

Akputat et al. [16] proposed a hybrid graphical password scheme known as T&C. In hybrid schemes multiple user authentication schemes are combined into single scheme. In this scheme passwords are consist of alphanumeric characters

and a location inside a picture. Users enter alphanumeric part of a password in text field through keyboard while location is identified through mouse. Usability is not a big issues in T&C scheme but passwords can be captured by online attacks, because they are directly inserted into the login screen. Another hybrid graphical password scheme was proposed by Alsaieri et al. [17], the scheme is known as Gotpass. GOTPass scheme is designed by combining properties of Android unlock, Deja Vu and textual password schemes. For authentication a user has to draw password lines and insert some codes which represent different password images. The scheme provide resistance from key logger, mouse logger and dictionary attacks but combined screen scrapper and key-mouse logger attacks can reveal passwords. This scheme has many usability issues such as high error rate and authentication time. This scheme also requires large amount of information to memorize.

### III. PROPOSED AUTHENTICATION SCHEME

In this research a user authentication scheme is proposed which reduces the security weaknesses of textual password scheme. The proposed scheme has two common authentication phases which are registration and login. Registration phase is same as ordinary textual password scheme but passwords are saved with different methodology. In the login phase changes are made in password entry screen and password verification process. Both phases are explained here.

#### A. Registration Phase

In this phase authentication information of a new user is registered. In the proposed scheme registration information is taken in same way as in ordinary textual password scheme. Therefore registration process is required to be executed in a secure machine and environment, where no one should be able to monitor the process. A secure channel should be used during registration time such as SSL/TLS [18] [19] for collecting password from a user. Generally registration phase is consist of three layers, which are password collection, password encryption and password storage into the database. In order to improve the password security from offline guessing attacks, transformation layer is added into the registration phase. The transformation layer is described here.

#### B. Password Transformation

In this layer alphanumeric characters of a password are converted into different alphanumeric character or symbols. Password transformation helps in resisting from brute force and dictionary attacks. For resisting brute force attack, theoretical password space and effective password space need to be high. Theoretical password space is the total number of passwords available in an authentication scheme, while effective password space is the total number of passwords being used by the users inside a scheme. Theoretical Password space and effective password space are increased by adding password transformation layer into the proposed scheme. Standard keyboard contains 94 alphanumeric characters excluding space key, therefore theoretical password space of textual password can be described with equation 1.

$$\sum_{i=1}^{94} 94^i \quad (1)$$

Majority of the users create password from less than 13 alphanumeric characters [20], therefore effective password space can be described with equation 2.

$$\sum_{i=1}^{12} 94^i \quad (2)$$

In order to decrypt a password, attackers need to check all the passwords belong to effective password space or in special case theoretical password space. Password transformation layer helps in increasing the size of theoretical and effective password space by adding symbols along with 94 alphanumeric characters.

Password transformation can be static or dynamic. In static transformation, same password of different users generate same transformed string. While in dynamic transformation, different transformed strings are generated from same password of different users. Password transformation can be carried out with many techniques, for example one strategy for static transformation is described using the following steps.

- (i) Create a list of alphanumeric characters as shown in Table I. The table contains all 94 alphanumeric characters.
- (ii) Create a combined alphanumeric characters and symbols list as shown in Table II. The list may be consist of more than two hundred elements.
- (iii) Find out the index number in Table I, which belongs to first character of a password.
- (iv) Get an element from Table II, which has same index number generated from previous step. The element would be transformed character or symbol.
- (v) Fetch index number of next character of the password from Table I.
- (vi) Sum previous index and current index of the elements, generated from Table I.
- (vii) Fetch an element from Table II, which has the index number generated after summation in step vi.
- (viii) Step v to vii will continue until all password characters are transformed.

TABLE I. LIST OF ALPHANUMERIC CHARACTERS

index	character
1	a
2	b
3	c
4	d
5	e
6	f
7	g
...	...
94	9

With the above transformation method the password “bdg” will be transformed to “βYσ” through the following steps, if the alphanumeric characters are stored in the form of Table I and symbols are is stored in the form of Table II.

- (i) System picks the index of first password character ‘b’ from Table I. The index of ‘b’ is ‘2’.
- (ii) System gets an element from Table II which has index ‘2’ . In this case the element is ‘β’.
- (iii) System fetches index of second password character ‘d’ from Table I. The index of ‘d’ is ‘4’.

TABLE II. LIST OF SYMBOLS AND ALPHANUMERIC CHARACTERS

index	character
1	$\alpha$
2	$\beta$
3	X
4	$\delta$
5	b
6	Y
7	$\lambda$
8	$\Theta$
9	$\chi$
10	g
11	$\sigma$
...	...

- (iv) System generates new index ‘6’ by adding current index ‘4’ with previous index ‘2’.
- (v) System fetches an element from Table II which has the index ‘6’. In this case the element is ‘Y’.
- (vi) System picks the index of last password character ‘g’ from Table I. Here the index of ‘g’ is ‘7’.
- (vii) System generates index “11” by adding current index ‘7’ with previous index ‘4’.
- (viii) System fetches an element which has index “11” in Table II. In this case the element is ‘ $\sigma$ ’.

Dynamic password transformation is also achieved by different methods, one of the method is password concatenation. In this method before applying password transformation steps, some characters are added into the password of a user. For example first three characters of user’s email address can be concatenated with the password. Every user has different email address, therefore same password of two users will have different transformed string.

### C. Login Phase

Authentication process of the proposed scheme is different from ordinary textual password scheme. In the password field users need to enter decimal numbers which represent the alphanumeric character of their password. For authentication, decimal numbers entered by a user are mapped into alphanumeric characters and then the characters are matched against stored password. Login phase is further divided into three parts, which are login screen generation, password entry and password matching.

1) *Login Screen Generation*: Login screen is a medium through which authentication information is collected and sent to a server. Login screen of the proposed scheme contains all alphanumeric characters along with some numbers as shown in Figure 1.

The alphanumeric characters are represented by decimal numbers from 0 to 9 (total 10 numbers). Each decimal number is assigned to 9 or 10 alphanumeric characters, because 94 alphanumeric characters are shown in the login screen and they are represented by 10 decimal numbers.

Each time a user opens the login page, the decimal numbers are randomly assigned to the alphanumeric characters. For example characters (g m x F G P X ) > ) are represented by decimal number ‘4’ in the login screen as shown in Figure 1. While in another session the alphanumeric characters (f h r y O X = [ ]) are assigned to the same decimal number ‘4’ as shown in Figure 2.

Through algorithm 1, every alphanumeric character is assigned a random decimal number within the range of 0 to 9. All alphanumeric characters and their corresponding decimal numbers are saved into session variable for password matching.

---

#### Algorithm 1 Numbers to characters mapping

---

```

1: alphaNum = List of alphanumeric characters
2: counter = 0
3: comment: Each decimal number is stored 10 times
4: for i = 0 to 9 do
5:   for j = 0 to 9 do
6:     numbers[counter]=i
7:     numbers++
8:   end for
9: end for
10: CLength = 100
11: comment: Rearranges elements of numbers array
12: for i = 0 to 100 do
13:   tempElement = Null
14: comment: A random index is generated
15:   ind = Random(0,cLength)
16:   tempElement = numbers[ind]
17:   numbers[ind] = numbers[cLength]
18:   numbers[cLength] = tempElement
19:   cLength = cLength - 1
20: end for
21: for i = 0 to 94 do
22:   alphaNum[i][0] = numbers[i]
23: end for

```

---

2) *Password Entry*: Passwords of the proposed scheme consists of alphanumeric characters but in the password field some decimal numbers are entered, which represent the password characters. For example if a user’s password is “bdg”, then the user has to enter “724” in the password field, if the login screen is same as shown In Figure 1. Here first digit ‘7’ represents the password character ‘b’ next digit ‘2’ represents the character ‘d’ and last digit ‘4’ represents the password character ‘g’.

3) *Password Matching*: In the password matching phase, authentication information provided by a user is compared against the stored authentication information. Password matching is further divided into two steps, password re-transformation and password numbers matching. Whole process of password matching is given in the flowchart as shown in Figure 3.

a) *Password re-transformation*: In this step a password is restored into its original form based upon the username provided by the user. A password is first decrypted if encrypted value is stored into a database and then the password is re-transformed by the following steps.

- (i) Index of first symbol or element is fetched from Table II.
- (ii) A password character is fetched from Table I, which contains the index number generated from step i.
- (iii) System fetches index of next element from Table II.
- (iv) An index number is generated by subtracting the index of Table I (belongs to previous password character) from the index generated in step iii.

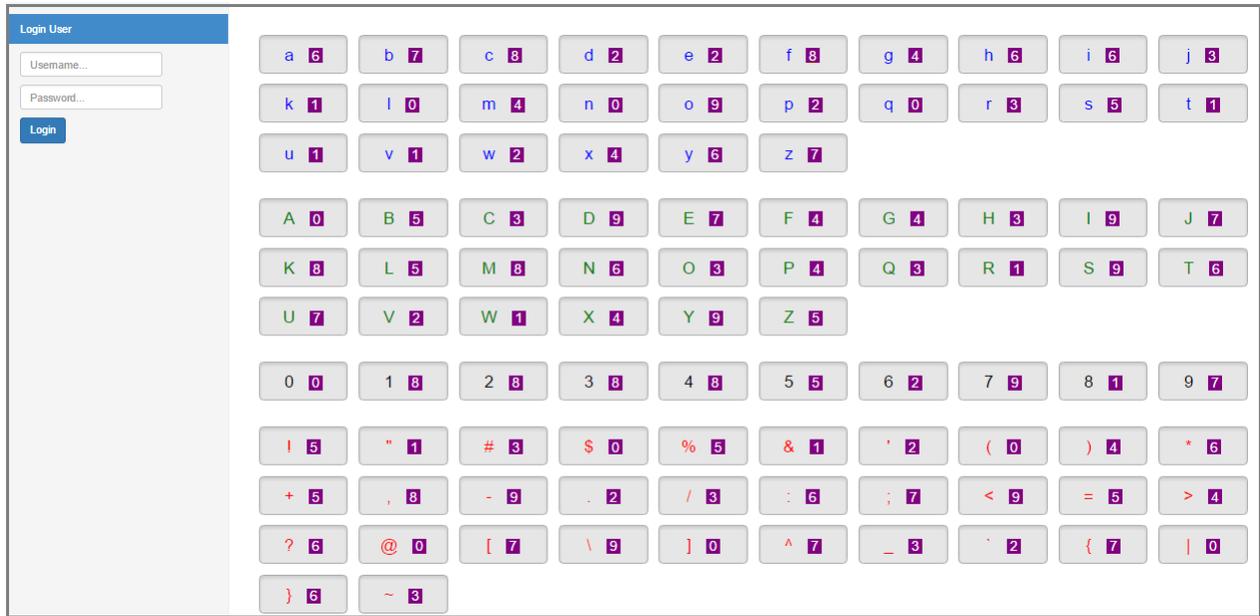


Fig. 1. Login screen



Fig. 2. Login screen 2

- (v) A password character is fetched from Table I which belongs to the index number generated in step iv.
- (vi) Steps iii to v will continue until all the password elements are fetched.

In the current scenario saved symbols “ $\beta Y \sigma$ ”, are converted to “bdg” through the following steps.

- (i) System fetches the index of first symbol ' $\beta$ ' from Table II, in this case the index is '2'.
- (ii) System fetches a character from Table I, which has same index '2'. In this case the system fetches character 'b', which is the first element of the password.
- (iii) System then fetches the index of next symbol which is 'Y' from Table II. The index of 'Y' is '6' in the table.
- (iv) System subtracts the index of 'b' which is '2' from the index of 'Y' which is '6'. The new index becomes '4'.
- (v) System fetches a character from Table I which has index position '4'. In this case the element is 'd' which is the second character of the password.
- (vi) System then fetches the index of ' $\sigma$ ' from Table II, the index is "11".
- (vii) An index number is generated by subtracting the index of 'd' from Table I, with the symbol ' $\sigma$ ' from Table II. In this case the index of 'd' is '4' and index of ' $\sigma$ ' is '11', therefore a new index '7' is generated.

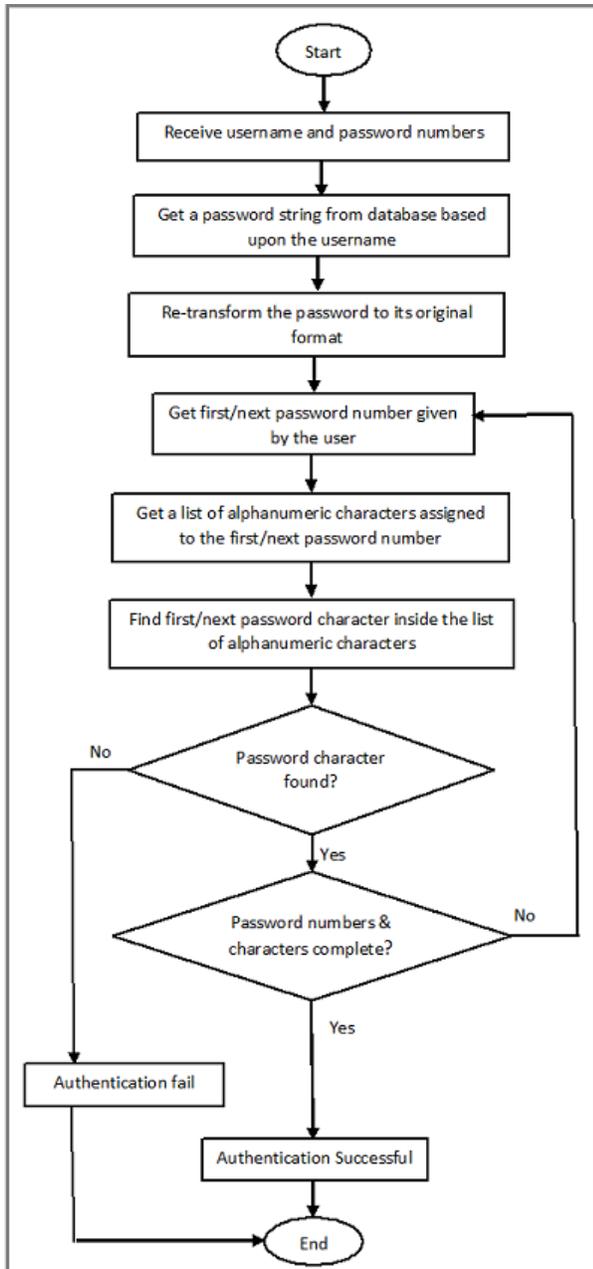


Fig. 3. Flowchart of password matching

(viii) System fetches a character which has index number ‘7’ in the Table I. In this case the character is ‘g’, which is the last character of the password.

*b) Password numbers matching:* In this step password numbers given by the users and re-transformed password are matched for authentication. Password numbers are matched with the following steps.

- (i) Server receive username and decimal numbers for authentication.
- (ii) Based upon username, password is re-transformed as described in password re-transformation process.
- (iii) System fetches a list of 9 or 10 alphanumeric characters from the session variable, which belong to first password

- number provided by the user.
- (iv) System searches first character of re-transformed or original password within the list of 9 or 10 alphanumeric characters.
- (v) If system successfully finds the password element in the list of alphanumeric characters, then system will repeat step iii & iv for all password numbers.
- (vi) If all re-transformed password characters successfully matches with the corresponding list of password characters, then system will allow login, otherwise user will not be authenticated.

*4) Illustration of password matching:* If password of a user is “bdg” and the password numbers entered by the user are “724”. Based upon the login screen shown in Figure 1 the system matches the password with the following steps.

- (i) System will fetch first decimal number, which is ‘7’.
- (ii) Corresponding alphanumeric characters are fetched from the session variable. In this case the characters belong to decimal number ‘7’ are (b z E G U 9 ; [ ^ {) see login screen as shown in Figure 1.
- (iii) System successfully finds first character of the password, which is ‘b’ within the set of ten elements (b z E G U 9 ; [ ^ {).
- (iv) System will fetch next decimal number, which is ‘2’ and its corresponding alphanumeric characters, which are (d e p w V ’ 6 . ^).
- (v) System successfully finds second password character ‘d’ within the set (d e p w V ’ 6 . ^).
- (vi) Finally the system will fetch last decimal number, which is ‘4’ and its corresponding alphanumeric characters, which are (g m x F G P X ) >).
- (vii) Password character ‘g’ is also present in the set (g m x F G P X ) >), therefore user will be authenticated.

#### IV. ANALYSIS OF THE PROPOSED SCHEME

This section consists of two parts, in first proposed scheme is analyzed with respect to security. In second part, usability analysis of the scheme is given.

##### A. Security Analysis

Proposed scheme improves the security of textual password scheme, by taking measures in password entry and database storage. Many security attacks are applied for password theft, in this section different security attacks are discussed with respect to the proposed scheme.

*1) Brute Force Attack:* In this attack all possible passwords exist in an authentication scheme are matched with the encrypted password value, in order to get the original password. Although proposed scheme uses same alphanumeric characters for password creation, but the theoretical and effective password space of the scheme is increased by using the transformation layer. Therefore brute force attack becomes difficult to apply.

*2) Dictionary Attack:* This is the efficient form of brute force attack, in which small number of passwords are tried to guess a password. In this attack a list of passwords are generated which is called password dictionary and the passwords are cracked by comparing every password of the dictionary

with a user's password. Password dictionary creation is very difficult in the proposed scheme, because the sequence of symbols generated after password transformation may not be same within different implementations of the scheme.

3) *Shoulder Surfing Attack*: In shoulder surfing attack, authentication information is revealed through camera recording or by directly observing login activity. This attack is not possible in the proposed scheme, because user does not directly enter the password characters but they enter some decimal numbers which represent multiple alphanumeric characters.

4) *Random Guessing Attack*: In this attack an attacker blindly try some passwords for authentication. The probability of successful random guessing attack in the proposed scheme is given by the equation 3

$$P(S) = (1/10)^N \quad (3)$$

Here n is the length of a password and the number 10 shows the range of decimal numbers used inside the scheme. If the length of a password is 8 alphanumeric characters then the probability of successful random guessing attack is 0.00000001. This probability is very low that is why attacker will not rely on this attack to break the password.

5) *keystroke/mouse logger attack*: keystroke/mouse logger programs send information to attackers without the consent of users. In textual password scheme keystroke logger can easily send the password of a user. While in the proposed scheme keystroke logger are not helpful for attacker, because they only get some random numbers instead of a password.

6) *Man-in-the-middle Attack*: There are many form of main-in-the-middle attack such as phishing and replay attacks. In phishing attack a legitimate user is redirected to a fake website, where the user enters the authentication information. From the fake website attacker gets the authentication information. In replay attack, password is recorded from a communication medium and then the recorded information is used later for authentication. Both attacks are not possible in the proposed scheme, because an attacker does not get exact password characters.

7) *Multiple Recording Attack*: Passwords in knowledge based authentication schemes can be captured by recording information of multiple login sessions. The information may be in the form of camera recordings or spyware data. Proposed scheme also suffer from this attack but it requires recordings of multiple login sessions. Equation 4 shows the condition in which password of the scheme is captured by multiple recording attack.

$$X \cap Y = Z / |Z| = 1 \quad (4)$$

Here X and Y are the set of alphanumeric characters of two login sessions belongs to a password character and Z is the intersection result. A password character is captured when cardinality of set Z comes to 1 or single alphanumeric character is generated after intersection. If cardinality of Z is not equals to 1 then recursively intersection work has to be performed until single element remains in Z. The recursion process is given by the equation 5.

$$Z_{i-1} \cap Y_i = Z_i \quad (5)$$

Here i is the count of login sessions,  $Z_{i-1}$  shows the intersection result of previous two sets and  $Y_i$  is the set of alphanumeric characters related with current login session.

In the current scenario the password "bdg" is captured by the following steps.

- (i) Attacker gets the information of first login session, which consists of password numbers "724" given by the user and screen-shot of the login screen appeared in-front of the user as shown in Figure 1.
- (ii) Attacker gets the information of second login session, which consists of password numbers "894" and screen-shot of the login screen as shown in Figure 2.
- (iii) Attacker gets the alphanumeric characters of first login session, which are related with first password digit. The password digit is '7' and its corresponding characters are (b z E G U 9 ; [ ^ { } ), see Figure 1.
- (iv) Attacker gets the alphanumeric characters of second login session, which are related with first password digit. The password digit is '8' and its corresponding characters are (b m p D H L M @ \_ ), see Figure 2.
- (v) After intersection of set (b z E G U 9 ; [ ^ { } ) and (b m p D H L M @ \_ ), the attacker captures the character 'b' of the password.
- (vi) Same process will be repeated for password characters 'd' and 'g' by utilizing second and third password digits given by the user.

It is not necessary that a password character is cracked after recording information of two login sessions. Multiple login sessions may be required for capturing a password character, because password numbers are randomly assigned to nine or ten password characters.

## B. Usability Analysis

Usability study was conducted in order to analyze password entry time and input accuracy of the proposed scheme. A web based application was developed in order to perform the usability tests. The testing application was created through PHP programming language and MySQL database. In the application users performed registration and login activities. Different processes were created inside the application for storing password entropy and password entry time inside the database. Information of successful and failure login attempts were also stored inside the database for analysing the input accuracy of the proposed scheme.

1) *Procedure*: In the experiment users were asked to register once and make three successful login attempts. Maximum three attempts were allowed for a successful login. Each user performed registration and login activities in single session because memorability evaluation was not the objective of the experiment. Users performed registration activity on a traditional registration page, because password input method in the registration screen of the proposed scheme is same as textual password scheme. While graphical interface of the login page was similar to Figure 1. Before starting the test, the purpose of testing and how to perform registration and login activities were explained to the users.

2) *Participant*: Application was tested by 30 volunteer users belong to Quaid-e-Awam University. From volunteers, 8 users were faculty members while remaining 22 users were students of different departments. Both male and female users participated in the experiment.

3) *Experiment Results*: Experiment data was collected from the database of the application in order to analyse the authentication time and input accuracy. Total 30 users made 123 login attempts, which were both fail and successful login attempts. Average password length of the thirty users was 7.83 alphanumeric characters and average password entropy was 38.28 bits.

a) *Password entry time*: Table III shows the password entry timings of the proposed scheme. Mean authentication time of the proposed scheme is 16.73 seconds which is larger than textual password scheme because users require extra time for searching the alphanumeric password characters in the login screen.

TABLE III. PASSWORD ENTRY TIMINGS OF THE PROPOSED SCHEME IN SECONDS

Mean	Lowest time	Largest Time	Standard Deviation
16.73	9	34.79	7.54

b) *Input accuracy*: Input accuracy of the proposed scheme with respect to first login attempt and within three login attempts is given in the table IV.

TABLE IV. INPUT ACCURACY

Accuracy	Percentage
First Attempts	73.17 %
Within three attempts	100%

Results show that the number of input errors are high in first login attempt, which may be due to a new method of password insertion. The users made less number of input errors when they become familiar with the graphical interface of the login screen.

## V. SECURITY COMPARISON OF THE PROPOSED SCHEME

Many authentication schemes are proposed which have different advantages and disadvantages. In this paper proposed scheme is compared with three commercially used authentication schemes along with traditional textual password scheme. Commercially used schemes are Android unlock pattern scheme, Passface scheme and Picture Gesture Authentication (PGA) scheme used by Microsoft in windows 8 operating system. Security comparison with respect to different attacks is given in table V. For comparison three values are used depending upon the level of resistance provide by the schemes against a particular security attack. The values are "Strong", "Moderate" and "Weak". The value "Strong" shows that the scheme provides high level resistance to the particular attack while the value "Moderate" shows that mid level resistance and the values "Weak" shows the scheme is weak or not resilient to the particular attack.

Table V shows that proposed scheme improves the security of traditional textual passwords against different type of attacks, only multiple recording attacks is a threat with the scheme. Proposed scheme is resilient to brute force and dictionary attacks due to password conversion process presented

in the paper. Proposed scheme also provides better security in comparison to different commercial authentication schemes.

## VI. CONCLUSION

The idea of proposed scheme is not to replace textual password scheme, but to enhance the scheme for improving the security aspect. Users can easily shift towards the proposed scheme from textual password scheme, because old textual passwords can be used inside the scheme. The proposed scheme is also easy to learn because a very simple approach is used for login process.

Proposed scheme uses alphanumeric character based passwords, therefore memorability results would be same as textual password scheme. The problem with alphanumeric passwords is that easy to remember passwords are easy to guess through dictionary attack [21]. However in the proposed scheme easy to remember passwords are not easy to guess due to password transformation layer. Proposed password transformation layer can also be used for other knowledge based authentication schemes for improving the security against dictionary and brute force attacks.

Many security threats can be resisted when passwords are indirectly inserted into an authentication scheme, but this approach has usability cost [1]. Users may require more mental or physical work to do in order to indirectly insert the passwords. That is why login time of the proposed scheme is higher than traditional textual password scheme, but this usability disadvantage is less than security advantages provides by the proposed scheme.

## REFERENCES

- [1] Q. Yan, J. Han, Y. Li, H. DENG *et al.*, "On limitations of designing usable leakage-resilient password systems: Attacks, principles and usability," 2012.
- [2] A. Adams and M. A. Sasse, "Users are not the enemy," *Communications of the ACM*, vol. 42, no. 12, pp. 40–46, 1999.
- [3] D. Davis, F. Monrose, and M. K. Reiter, "On user choice in graphical password schemes." in *USENIX Security Symposium*, vol. 13, 2004, pp. 11–11.
- [4] I. Jermyn, A. J. Mayer, F. Monrose, M. K. Reiter, A. D. Rubin *et al.*, "The design and analysis of graphical passwords." in *Usenix Security*, 1999, pp. 1–14.
- [5] G. E. Blonder, "Graphical password," Sep. 24 1996, uS Patent 5,559,961.
- [6] S. Uellenbeck, M. Dürmuth, C. Wolf, and T. Holz, "Quantifying the security of graphical passwords: the case of android unlock patterns," in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM, 2013, pp. 161–172.
- [7] H. Zhao and X. Li, "S3pas: A scalable shoulder-surfing resistant textual-graphical password authentication scheme," in *Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on*, vol. 2. IEEE, 2007, pp. 467–472.
- [8] Z. Zheng, X. Liu, L. Yin, and Z. Liu, "A stroke-based textual password authentication scheme," in *Education Technology and Computer Science, 2009. ETCS'09. First International Workshop on*, vol. 3. IEEE, 2009, pp. 90–95.
- [9] Y.-L. Chen, W.-C. Ku, Y.-C. Yeh, and D.-M. Liao, "A simple text-based shoulder surfing resistant graphical password scheme," in *Next-Generation Electronics (ISNE), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 161–164.
- [10] K. Rao and S. Yalamanchili, "Novel shoulder-surfing resistant authentication schemes using text-graphical passwords," *International Journal of Information and Network Security*, vol. 1, no. 3, p. 163, 2012.

TABLE V. SECURITY COMPARISON

Scheme	Brute Force attack	Dictionary Attack	Shoulder Surfing	Random guessing attack	Keystroke/Mouse logger Attack	Man-in-the-Middle Attack	Multiple recording attack
Android unlock	Weak	Weak	Weak	Moderate	Moderate	Weak	Weak
PassFaces	Weak	Weak	Weak	Moderate	Weak	Strong	Weak
PGA	Strong	Strong	Weak	Moderate	Moderate	Moderate	Weak
Textual passwords scheme	Moderate	Moderate	Moderate	Strong	Weak	Weak	Weak
Proposed scheme	Strong	Strong	Strong	Strong	Strong	Strong	Moderate

- [11] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon, "Passpoints: Design and longitudinal evaluation of a graphical password system," *International journal of human-computer studies*, vol. 63, no. 1, pp. 102–127, 2005.
- [12] S. Wiedenbeck, J. Waters, L. Sobrado, and J.-C. Birget, "Design and evaluation of a shoulder-surfing resistant graphical password scheme," in *Proceedings of the working conference on Advanced visual interfaces*. ACM, 2006, pp. 177–184.
- [13] N. Lopez, M. Rodriguez, C. Fellegi, D. Long, and T. Schwarz, "Even or odd: A simple graphical authentication system," *IEEE Latin America Transactions*, vol. 13, no. 3, pp. 804–809, 2015.
- [14] D. Weinshall, "Cognitive authentication schemes safe against spyware," in *Security and Privacy, 2006 IEEE Symposium on*. IEEE, 2006, pp. 6–pp.
- [15] H. Gao, W. Jia, N. Liu, and K. Li, "The hot-spots problem in windows 8 graphical password scheme," in *Cyberspace Safety and Security*. Springer, 2013, pp. 349–362.
- [16] M. Akpulat, K. Bicakci, and U. Cil, "Revisiting graphical passwords for augmenting, not replacing, text passwords," in *Proceedings of the 29th Annual Computer Security Applications Conference*. ACM, 2013, pp. 119–128.
- [17] H. Alsaiani, M. Papadaki, P. Dowland, and S. Furnell, "Graphical one-time password (gotpass): a usability evaluation," *Information Security Journal: A Global Perspective*, vol. 25, no. 1-3, pp. 94–108, 2016.
- [18] T. Dierks, "The transport layer security (tls) protocol version 1.2," 2008.
- [19] A. Freier, P. Karlton, and P. Kocher, "The secure sockets layer (ssl) protocol version 3.0," 2011.
- [20] Z. Liu, Y. Hong, and D. Pi, "A large-scale study of web password habits of chinese network users," *JSW*, vol. 9, no. 2, pp. 293–297, 2014.
- [21] X. Suo, Y. Zhu, and G. S. Owen, "Graphical passwords: A survey," in *Computer security applications conference, 21st annual*. IEEE, 2005, pp. 10–pp.

# Impact of Pulse Voltage as Desulfator to Improve Automotive Lead Acid Battery Capacity

EL MEHDI LAADISSI  
Laboratory LM2PI, ENSET,  
Mohamed V University  
Rabat, Morocco

ANAS EL FILALI  
Laboratory LM2PI, ENSET,  
Mohamed V University  
Rabat, Morocco

MALIKA ZAZI  
Laboratory LM2PI, ENSET,  
Mohamed V University  
Rabat, Morocco

**Abstract**—This paper studies the impact of Pulse Voltage as Desulfator to recover weak automotive Lead Acid Battery capacity which is caused by Sulfation. This technique is used to overcome the premature loss of battery capacity and speed up the process of charging and extend the lead acid battery life cycle 3 to 4 times compared with traditional charging methods using constant current. Sulfation represents the accumulation of lead sulfate on the electrodes (lead plates). This phenomenon appears naturally at each discharge of the battery, and disappears during a recharge. This is common with starter batteries in cars driven in the city with a load-hungry accessory. A motor in idle or at low speed cannot charge the battery sufficiently. Voltage pulse decompose the sulfate (PbSO<sub>4</sub>) attached to the electrode which is the main cause of the loss of capacity. In this paper, we study the effects of the recovery capacity of a Lead Acid Battery. Voltage pulses will be applied on a commercial automotive battery to collect data, using a charger/Desulfator prototype based on a PCUINO. The experiment results show that there is improvement of Cold Cranking Amps level, and charge time duration of the Lead Acid Battery after using our prototype.

**Keywords**—Lead acid battery; desulfator; pulse charging; cold cranking; sulfation

## I. INTRODUCTION

In the last few years there has been a growing interest in batteries, they have revolutionized the way of storing electrical energy. Its use is widespread and growing, it helps to have a reserve of electrical energy autonomous and mobile cell phones, photovoltaic systems, space equipment, laptops and other devices to public or industrial use. Especially since the battery is a power source that can partially replace the use of internal combustion systems used in the new generation of hybrid electric cars and this to reduce the emission of greenhouse gases which is now the major concern of humanity. To use a battery effectively, it is necessary to understand its dynamics and discover the parameters which may affect its performance. The problem with these types of batteries is that they need to maintain their useful life capacity as long as possible and to optimize the use of their energy [1], [2].

## Nomenclatures

$PbSO_4$	Lead(II) Sulfate
$H_2SO_4$	Sulfuric Acid
$Pb$	Lead
$PbO_2$	Lead dioxide

## Abbreviations

PWM	Pulse-Width Modulation
LAB	Lead Acid Battery
SLI	Start, Lighting, Ignition
CCA	Cold Cranking APMS
DOD	Depth of Discharge

### A. Research Purposes

The purpose of this study is:

- To examine the effect of a recovery made from lead acid battery capacity using a pulse voltage charging method on a battery that has a low storage capacity.
- Verify and validate the circuit prototype producing voltage pulses that have been made previously.

### B. Problem Formulation

How the method of “charge pulse” helps lead acid batteries to recover their capacity, what is the effect of these “impulsions” on batteries. To do so, we will take an interest in the value of power starter/cold cranking amps and the duration of charging a lead acid battery.

### C. Scope of the Experiment

Study was conducted using commercial automotive batteries on the market. The charging process is done by using our prototype charger based on PCUINO shown in Fig. 1 and an algorithm programmed with python language.

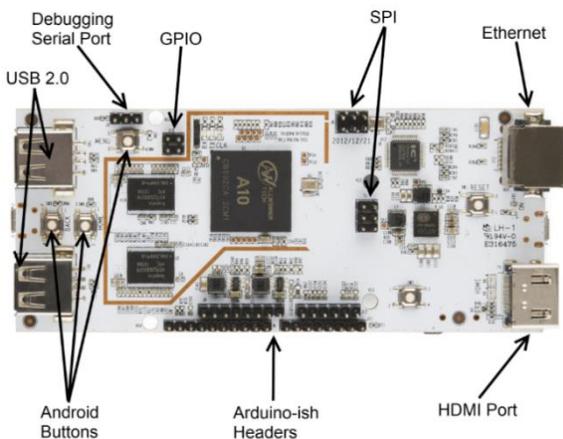
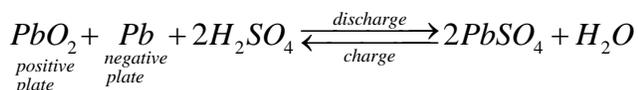


Fig. 1. Desulfator programmable main board.

## II. THEORETICAL FRAMEWORK

### A. Charging Process of a Lead Acid Battery

Lead acid battery have anode made of lead (Pb) and the cathode made from lead dioxide (PbO<sub>2</sub>), H<sub>2</sub>SO<sub>4</sub>, and a separator between the two electrodes. The chemical reaction that occurs at the positive electrode and negative electrode of the battery are as follows [3]:



### B. Sulfation

Sulfation represents the accumulation of lead sulfate on the electrodes. This phenomenon arises naturally in every discharge of the battery [4]-[6], and disappears when a refill. However, under certain conditions (too deep or prolonged discharge, large temperature, gasification of the electrolyte), plates of stable lead sulfate appear and are not dissolved during charging. Lead sulfate thus generated decreases battery capacity by preventing the reactions on the electrode and its low electrical conductivity.

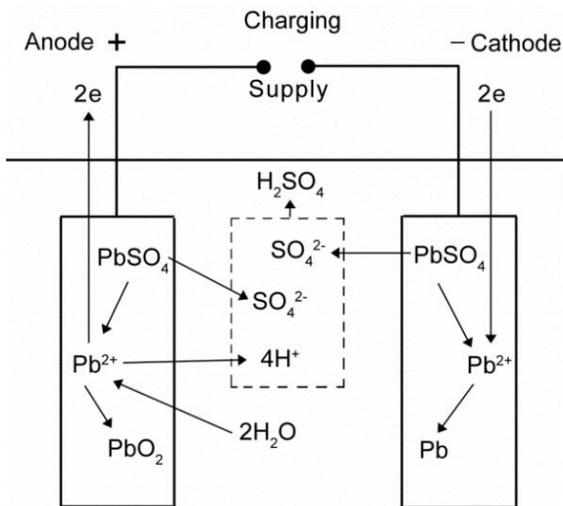


Fig. 2. Construction of a lead acid battery.

Based on the above reaction PbSO<sub>4</sub> will be formed at the positive electrode and negative electrode during charging or using the battery while in the process of charging the electrode PbSO<sub>4</sub> will decompose and become Pb, PbO<sub>2</sub> and H<sub>2</sub>SO<sub>4</sub> [7]. Lead acid battery construction can be seen in Fig. 2.

### C. Desulfation

There is a way to reverse the sulfation process of a battery. This consists of sending electrical pulses to the battery with a resonance frequency (between 2 and 6 MHz). During this process, the sulfur ions collide with the plates, which have the effect of dissolving lead sulfate covering them. The shape of the voltage pulse can be seen in Fig. 3.

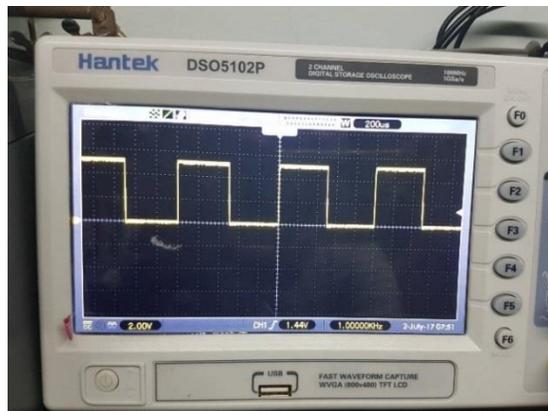


Fig. 3. Output pulse shape.

### D. Lead Acid Battery Capacity and Peukert Law

Understanding battery capacity refers to the amount of energy that can be stored. Lead acid battery manufacturers often use the specification known as Amp Hour to provide an indication of the stored battery capacity. But lead acid battery capacity will be determined by the amount of load current. This means that greater the load on the battery, the less realized capacity you will have. This phenomenon is called Peukert law [8].

### E. Lead Acid Battery Types

Based on the application, the LAB can be divided into several types [9]:

- SLI (Start, Lighting, Ignition) which is used in automotive.
- Stationary to support the power supply to generate and store electrical energy used in telecommunications systems, electric utilities centre, and computer systems.
- Traction, to power transportation equipment such as forklifts, electric cars, and mining equipment.
- Battery specially meant for the use of aircraft, submarines and military equipment.

### F. Starter Power

World power industry uses standard automotive starter with a size of Cold Cranking APMS (CCA). The number indicates the current in ampere that the battery can deliver at -18°C (0°F) for 30 seconds, while being able to provide a

voltage of 1.2 volts per cell or higher. American and European norms differ slightly.

### G. Depth of Discharge

Depth of Discharge (DOD) is used to describe how deeply the battery is discharged. If we say a battery is 100% fully charged, it means the DOD of this battery is 0%, If we say the battery have delivered 40% of its energy, here are 60% energy reserved, we say the DOD of this battery is 40%. And if a battery is 100% empty, the DOD of this battery is 100%. DOD always can be treated as how much energy that the battery delivered.

## III. RESEARCH METHODS

### A. Stages of Research

Flowchart of this research can be seen in Fig. 4. The research began with a literature study, followed by the preparation of the batteries to be tested. Battery preparation is intended to obtain a battery which has a weak storage capacity, so that we can see the effects of the voltage pulse charging technique on the capacity of the lead acid battery. Battery preparation is done by charging with a constant current for 30 cycles. After the battery is ready, treatment is initiated by the load voltage impulsion, respecting the manufacturer's specifications. Throughout the charge cycle the battery data are collected and stored to analyse the effects of this voltage pulses on the capacity of the lead acid battery.

### B. Data and Sources

Data:

- 1) Results of voltage measurement.
- 2) The capacity and time.
- 3) Battery power starter.

Sources:

- 1) Instruments measurement of capacity.
- 2) Battery manufacturer specifications.

## IV. RESULTS AND DISCUSSION

### A. Battery Preparation

The lead acid battery used were NS60, that has been used for 4-6 months on the vehicle, with the condition Voltage 12,29 Volt, 195 CCA starter power, according to standards Issued by Yuasa, NS60 starter has a standard power of 325 CCA.<sup>1</sup>

Aside from the value of the starter power of the lead acid battery, capacity can also be seen through the value of the voltage, a battery which has a value voltage of 12.65 V is said to have a capacity of 100%, 12.40 V is said to have a capacity of 75% [10]-[13].

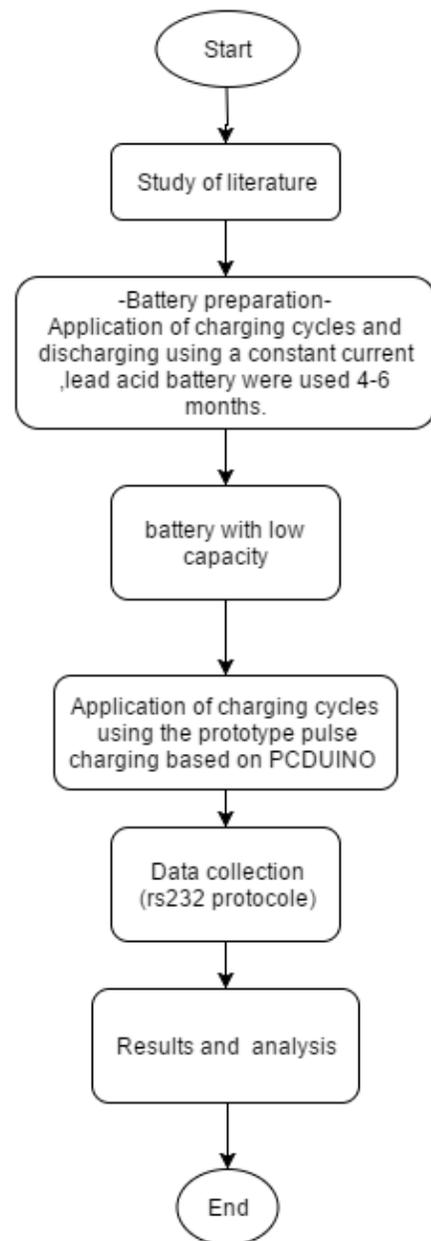


Fig. 4. Stages of research.

### B. Experiment and Data Collection

The lead acid battery that has been used is measured by Battery Analyzer BT 747 DHC, to know the initial conditions of power starter / battery capacity. Furthermore, we charge the measured battery with shock pulses with a frequency between 2 kHz and 6 MHz, the frequency is selected by the user through a menu programmed in the PCDUINO until the voltage reaches 13.5 V.

<sup>1</sup>[http://www.yuasa.com.tw/\\_english/02\\_automobile/01\\_list.php?CID=1](http://www.yuasa.com.tw/_english/02_automobile/01_list.php?CID=1)

After that, the battery rested for 60 minutes to obtain a stable voltage to subsequently remeasure Battery Analyzer to determine the value of starter of the battery, then discharged for 60 minutes. The cycle is carried out repeatedly for 20 cycles so that the change in the value of Starter CCA each cycle of charging and discharging can be known. Circuit producing a voltage pulse can be seen in Fig. 5 and 6 while the shape of the pulses generated can be seen in Fig. 3.

C. Pulse Voltage Effect on the Battery Capacity

A major increase can clearly be seen in the value of the “CCA” during the first cycles. While in the subsequent cycles we see an increase that was not significant Fig. 7. While Fig. 8 shows an increase in voltage on the lead acid battery.

D. Influence of the Charging Cycles on the Charging Time

For the first cycles, we needed a large amount of time to obtain the required voltage 13.5V but faster one reaches this value for subsequent cycles. This is due to the value of the internal resistance of the battery which is caused by sulfation.

With the passage of the charging cycle, lead sulfate decreases so that the charging process becomes faster. Thus, the battery regains the ability to store an electric charge. In addition, its time to full charge is visually improved.

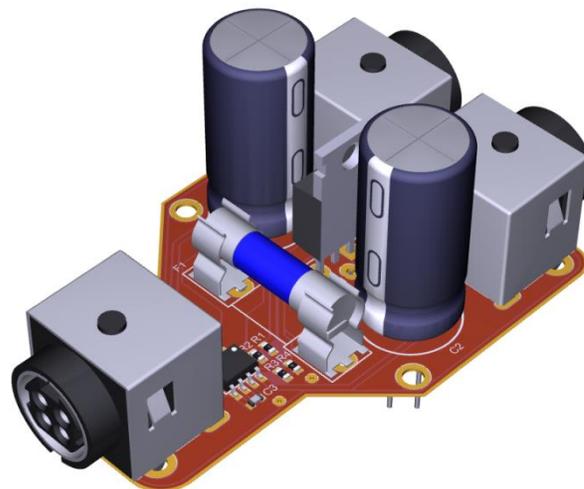


Fig. 6. 3D view of the pulse voltage generator.

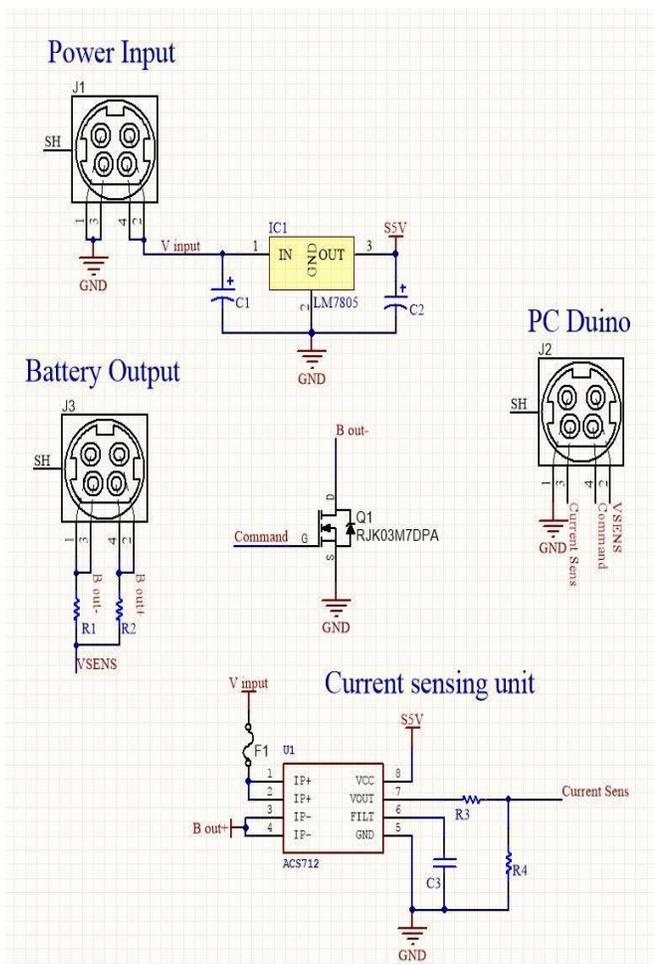


Fig. 5. Circuit voltage Pulse generator.

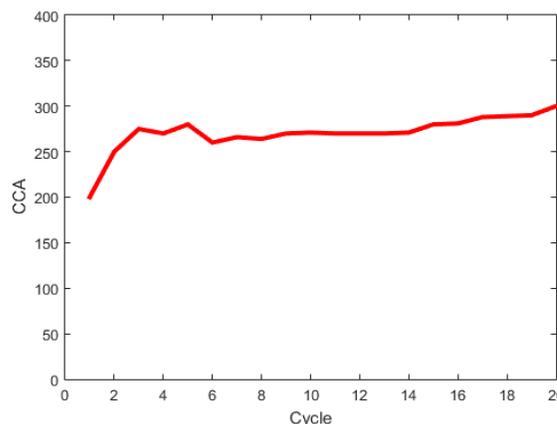


Fig. 7. Evolution of the CCA value.

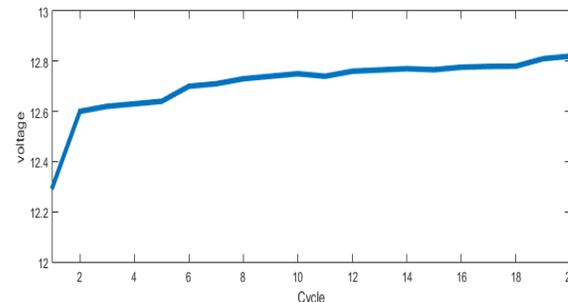


Fig. 8. Voltage value after each charge-discharge cycle.

V. CONCLUSIONS

Based on the above experimental results, charging the lead acid battery with the generated voltage pulses will increase storage capacity, the battery will be getting better after charging and discharging cycles.

At the beginning, the cycle of the charging process will take longer because of the high value of internal resistance which is due to sulfation, but with the declining value of the

internal resistance the lead acid battery storage capacity will be recovered and the charging time will be shorter. The prototype generates a voltage pulse that can increase battery capacity.

The use of voltage pulse charging technology is a highly promising method to be applied to batteries made from lead sulfate to extend the service life of the lead acid battery, other than that, it would be good to reduce the environmental pollution caused by the lead acid battery waste.

#### REFERENCES

- [1] D. Berndt, Maintenance-Free Batteries, second ed., Wiley, New York, 1997, p. 453.
- [2] Spanos, C., Berlinger, S. A., Jayan, A., & West, A. C. (2016). Inverse charging techniques for sulfation reversal in flooded lead-acid batteries. *Journal of The Electrochemical Society*, 163(8), A1612-A1618
- [3] Kujundžić, G., Ileš, Š., Matuško, J., & Vašak, M. (2017). Optimal charging of valve-regulated lead-acid batteries based on model predictive control. *Applied Energy*, 187, 189-202.
- [4] P. Ruetschi, *J. Power Sources*, 127 p. 33–44, 2004. G. Eason, B. Noble,
- [5] H. Ikeda, S. Minami, Song JieHou, Y. Onishi, A Kozawa. *J. Asian Electric Vehicle*, 3, p. 683, 2005.
- [6] B. Culpin, A.F. Hollenkamp, D.A. J. Rand, *J. Power Sources*, 38, p.63-74, 1992.
- [7] C.V. D'Alkaine, L.M.M. De Souza, P.R. Impinnisi. *J. Power Sources*, 158, p. 997, 2006.
- [8] Cugnet, M.G.; Dubarry, M.; Liaw, B.Y. Peukert's law of a lead-acid battery simulated by a mathematical model. *ECS Trans.* 2010, 25, 223–233.
- [9] E. Rocca, J. Steinmetz, S. Weber, Mechanism of formation of dense anodic films of PbO on lead and lead alloys in sulfuric acid, *J. Electrochem. Soc.* 146 (1999) 54–58. M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.
- [10] P. Rüetschi, B.D. Cahan, Anodic corrosion and hydrogen and oxygen overvoltage on lead and lead–antimony alloys, *J. Electrochem. Soc.* 104 (1957) 406–413.
- [11] S. Atlung, B. Zachau-Christiansen, Degradation of the positive plate of the lead–acid battery during cycling, *J. Power Sources* 30 (1990) 131–141.
- [12] J. Alzieu, N. Koechlin, J. Robert, Internal stress variations in lead–acid batteries during cycling, *J. Electrochem. Soc.* 134 (1987) 1881–1884.
- [13] Laadissi El, El Filali A., & Zazi, M (2016). Nonlinear Black Box Modeling Of A Lead Acid Battery Using Hammerstein-Wiener Model. *Journal of Theoretical and Applied Information Technology*, 89(2), 476.

# The Dynamics of IT Workaround Practices

## A Theoretical Concept and an Empirical Assessment

Ahmed Alojairi

King Fahd University of Petroleum and Minerals (KFUPM)  
Dhahran, Saudi Arabia

**Abstract**—An interesting phenomenon that has received limited attention in the extant literature is that of IT workaround practices. Based on Ashby’s Law of Requisite Variety, workarounds were found to be used to accomplish the basic task of matching unmatched variety in the system. The Interaction Effectiveness (IE) ratio of 1.4 was used as a baseline to uncover potential sources of workarounds. The Echo method was used to collect data from 42 users in a high-technology company (HTC). Enablers of and barriers to workaround practices were divided into four main categories: flexibility, reliability, ease of use, and coordination whereas workarounds were divided into three categories: using other tools, seeking help, and accepting. The results of the case study indicate that “reliability” is the dominant category for both helpful and non-helpful incidents, whereas “coordination” was the least significant. Of the workaround mechanisms, “using other tools” was the most significant category for all users. The findings suggest cycles of continuous improvement to the IE ratio to alleviate the need for workarounds, but a more fundamental issue concerning the source of workaround behaviors is a function of misfits between input variety by users and variety handling capabilities of the system.

**Keywords**—IT effectiveness; workarounds; cybernetics

### I. INTRODUCTION

Despite broad recognition of ERP systems for their seamless integration of all information flows within organizations and their ability to standardize the processes of different departments [1], [2], substantial research has shown that many ERP systems are unexpectedly complex to adopt, and their ultimate benefits are uncertain, resulting in sub-optimal operating practices [3], [4]. The challenges to ERP system implementation and success include, but are not limited to, inflexibility [5]; disturbance to organizational culture [6]; the requirement of significant investments of money, time, and expertise [7]; strain on the organization [4]; and inadequate training and support for end users [8].

Even with the implementation of ERP systems, which were introduced as a means to enforce standardization and control, the probability of external activity beyond the system seems almost impossible to control. As [9] indicates, enterprise or computing systems often replace other legacy systems and processes, establishing “open systems” that are incapable of including all contingencies. When IT systems are perceived as a barrier, two distinct but related phenomena may be observed: resistance to change and workaround practices. Resistance to change is usually perceived as a negative behavior in which users oppose the disturbance of a perceived

flaw in a system [10]. On the other hand, workaround practices are seen as a positive behavior in which users adapt in order to overcome the shortcomings of a system [10].

All workaround practices share common attributes. Whenever a user attempts to bypass formal system processes to overcome a barrier imposed by the system in order to complete a task, he or she is engaging in a workaround practice. Similarly, workaround practices occur at the post-implementation phase of any system, which may extend beyond the formal systems [11], [12]. In [13], author offers a broader and more inclusive definition of workaround as follows:

A workaround is a goal-driven adaptation, improvisation, or other change to one or more aspects of an existing work system in order to overcome, bypass, or minimize the impact of obstacles, exceptions, anomalies, mishaps, established practices, management expectations, or structural constraints that are perceived as preventing that work system or its participants from achieving a desired level of efficiency, effectiveness, or other organizational or personal goals (p. 1044).

Although workaround as an activity is well recognized in many fields such as nursery, project management, the military, and budgeting, some researchers have asserted that workaround theories remain somewhat understudied and underdeveloped [13]. Building a theory of workarounds will help to improve organizations, management, work practices, standards, and technology design and adoption and will also provide a unified view of workarounds and augment existing definitions.

To date, very little research has examined IT workaround practices from a qualitative perspective. The basic objective of this study, therefore, is to develop an in-depth understanding of workaround-related interactions and what governs such interactions. To this end, this study encompasses two primary objectives. First, this study presents a dynamic theory of workaround-related interactions within complex social networks. This study’s theoretical development draws from Ashby’s Law of Requisite Variety – to formulate the social and technical components of workaround practices as a network of task-related social interactions within an organizational context. Second, this study presents relevant quantitative and qualitative results from a field study in which the proposed theoretical framework is used. The data for this study is extracted from a high-technology company (hereafter referred to as HTC). Specifically, this study will uncover the motivations and preconditions for the occurrence of a

workaround, as well as the main types of workaround strategies that users employ to overcome barriers to accomplishing their goals.

The paper is structured as follows. The next section reviews relevant literature on the subject of workaround practices in various fields. The section that follows discusses the study's methodology, which adopts a qualitative interview method known as the Echo method [14] to allow the study's participants to reflect on their unique interactions and experiences with these systems. The remaining sections of the paper present and discuss the findings of a case study for which data was obtained from a company that has implemented an ERP solution.

## II. LITERATURE REVIEW

A workaround is a strategy of using a computer system in a manner that it was not designed to be used for or using alternative methods to accomplish a work task. A workaround is a useful strategy to solve an immediate and pressing problem [9]. In [15], author explained workaround by noting that "when a path to goal is blocked, people use their knowledge to create and execute an alternate path to that goal" (p. 71). In addition, they defined such a path, or workaround, as "a temporary fix that implies a genuine solution to the problem is needed" [15]. When users are unable to collect data they need from existing IT systems, they employ their own unique methods to collect such data [16]. In [17], author has defined a 'workaround' as a practice that "involves (1) a specific policy procedure or rule enforceable by bureaucratic superiors (2) that constrains or impedes local implementation and goal attainment and (3) prompts a local response that is counter to the procedure or rule but responsive to the underlying policy intent." Similarly, [18] defines a workaround as "the substitutive method that is used to overcome a constraint in information interaction in CIS with a specific motive to complete a work task" (p.381). They analyze workarounds in terms of the process involving the antecedent conditions, the actual workaround, and the consequences. They concluded that a workaround is a technical trick to interact with information systems, representing an everyday strategy used by workers to obtain better information in order to do their job [19], [20].

The workaround phenomenon has been discussed in many areas of organizational studies, such as public administration, healthcare, technology, finance, and accounting [21]. Workarounds have three forms: data adjustment, procedural adjustment, and backup systems [9]. In [21], author considers workarounds to be an essential expectation for those who rely upon open systems. In order for workarounds to be effective, some parameters should be established: educate employees about workarounds and the possibility of positive or negative impacts on the organization, observe the existing process of work, and encourage discussion among employees to identify the challenges [19], [21].

In [9], author identified three tactics users can adopt to overcome barriers imposed by any system. The first is fitting, which refers to any activity that attempts to modify the computing structure of the work by adapting to the computing error. Examples include adjusting work schedules and

commitments due to the backlog of information requests demanded by users of the system. The second is augmenting, which basically entails supplementing or expanding the current system to cover for system discrepancies, such as consolidating data resources from multiple sources that are currently fragmented. Last are workarounds, which involve the deliberate use of the system in ways not initially intended to achieve a required target by utilizing other methods to accomplish tasks. Common examples of workarounds are data adjustment or data manipulation to arrive at the desired result. One the other hand, [10] classified workarounds into three types based upon their consequences to organizations. The first is a hindrance workaround, which is the bypassing of the formal system due to perceived time-consuming, tedious or problematic procedures and/or processes. The second is a harmless workaround, which basically does not significantly interrupt the workflow or the accuracy of data. The last is an essential workaround, which involves necessary actions to complete the task at hand.

The reasons that lead to workaround practices include barriers in workflow, additional demands for work, rigid organizational rules, and poorly designed systems, which in turn lead users to employ workaround tactics to overcome such barriers imposed by the system's rigidity, which inhibits users from fulfilling their work requirements [21], [22]. In addition, workarounds are used when the information needed to meet an external demand is limited or lacking [18]. Reasons that lead to workarounds include a block in workflow, additional demands for work, organizational rules, and poorly designed systems; in general, users employ workarounds to accomplish tasks when they perceive the system to be inflexible and incompatible with their workflow [21]. Workaround practices may be inspired by an expert opinion or website help page. Most often, however, they are discovered by trial and error [15]. Interestingly, when a workaround is not available, users may change their goal to conform to a work process they already know the system is capable of executing. For example, if sending a file through an email is not working, a possible workaround is to put it on a web server. Workarounds can occur not only because of software defects, but also when software and environment requirements are inadequate [15]. Technologies that are related to physical structure, such as the lack of wireless connectivity, may also lead to workarounds [21]. In addition, a lack of expert and well-trained users leads to an increase in the existence of poor conformance (i.e. workarounds) within the system [23].

Some researchers argue that workarounds have both positive and negative consequences. Positive consequences include increased awareness, the availability of better information, and saved time [18]. Additionally, workarounds help to identify defects that need to be addressed, and they can be employed to speed up the processes of an organization and help to avoid unnecessary barriers to quality service and care [23]. In addition, a workaround is essential to the integration of IS; otherwise, service and performance will decrease dramatically. Thus, workaround practices are dependent upon the relationships between users, specialists, and key actors [9]. In [18], author considers workaround practices to be an innovative means of customizing IS in ways that will not

affect the accuracy of data. Reference [13] asserts that workaround practices are essential for performing everyday work. Additionally, IT workarounds may provide benefits to the individual and the organization as a whole, such as through the identification of existing gaps in IS between what centralized information systems offer and what users would like to have [11], [16].

However, workarounds may also represent an obstacle to improvements and result in decreased effectiveness [21]. Reference [13] argues that workaround practices might be viewed as undesirable and unethical or illegal violations of procedures. Workarounds may include some sophisticated technological solutions, such as designing local databases that are able to provide needed information. However, such practices may result in significant costs related to the use of workarounds [16], [22]. Along the same line, [12] has found that some applications originally designed to reduce process variations in healthcare settings (e.g., medication ordering and dispensing systems) actually resulted in increased process variations. This may obstruct the objectivity and reliability of the system, generating inaccurate or inconsistent reports when required by users.

Overall, review of the extant literature on workaround practices indicates that the workaround phenomenon thus far lacks theoretical elucidation necessary to provide a deeper understanding and predictive explanations. In support of this view, [13] asserts that there is no published comprehensive theory of workarounds.

### III. THEORETICAL BACKGROUND

The cybernetics theory, specifically, Ashby's Law of Requisite Variety, is used to develop a theoretical approach to modeling workaround practices. This model views the social and technical component of workaround practices as a task-related social network within an organizational context. Such a model helps in understanding Human-Computer interactions, as well as what governs such interactions [26].

Ashby's Law is a part of the cybernetics theory that advocates for the ability of any system to achieve its goals while maintaining viability [24]. In other words, any open system must adapt to its environment in order to survive. Ashby's Law of Requisite Variety states, "Only variety can destroy variety" [25], (p. 207). Reference [26] explained Ashby's Law of Requisite Variety as follows:

A system survives to the extent that the range of responses it is able to marshal – as it attempts to adapt to imposing tensions – successfully matches the range of situations – threats and opportunities – confronting it (p. 282).

Variety is defined as the number of different possible states the system can assume and their relative probability of occurring, which is indicative of the level of complexity a system can handle [25]. Thus, the internal complexity of a system must match the external complexity it provokes to remain stable [25]. Input variety can be seen as a disturbance to the system's stability. There are two sources of input variety: external variety and internal variety. External variety refers to variety generated by the organization's environment. An example is variability in a supplier's behavior, such as late

deliveries and inconsistent quality. Internal variety, on the other hand, refers to variety generated by the system itself that affects its own performance, such as poorly designed systems, machinery breakdowns, and human error.

In general, [27] suggested two approaches to maintain the stability of any system: reducing variety at the source and/or increasing the variety handling capability of the system. For example, a manufacturing organization can cope with a certain amount of raw material variability from suppliers by either implementing policies to force suppliers to ship parts on time (i.e. stimulus simplification) or increasing inventory levels as a buffer against any future variability in shipment arrivals (i.e. response complexification). Reference [26] summarized how a system is capable of responding to its environment in adaptive ways as follows:

1) Simplify the complexity of incoming stimuli so as to economize on the resources that need to be expended in responding.

2) Invest more resources in the response than they judge to be strictly necessary so as to ensure some degree of adaptation (p. 279).

On an abstract level, ERP systems may be viewed as a set of predefined assumptions and preconditions about what organizations are and how they should function. Such rule-based systems can become static and difficult to change, much like organizational bureaucracy. In order to overcome such rigidity in the system, users may engage in machine-like behaviors and permit their behaviors to be formalized, or users may engage in informal workaround practices to realign the system with organizational requirements. In a way, workaround practices enable ERP systems to be flexible enough to adjust to dynamic changes in organizational requirements and the environment.

In this context, any ERP system is originally designed to effectively conform to organizational requirements by handling input variety. That is, ERP systems are variety regulating systems to the extent that input variety is continuously absorbed (destroyed) in order to regulate the output. However, while ERP systems produce this requisite variety, it inevitably creates some unintended and undesirable variety due to changes in organizational requirements or environmental demands. Therefore, the effectiveness of the system is dependent upon the fit between the input variety and the variety handling capabilities of the system [28]. That is, as the organization adapts to internal (e.g., business requirements) and external (e.g., environmental demands) changes, ERP systems are employed as a variety handling mechanism to match new variety. However, amplified misfits between variety and variety handling will result in the generation of more unpredictable outputs and more unstable systems.

IT misalignment, or IT misfit, within the organization will result in excess variety, which must be addressed either by reduced variety at the source (e.g., customization) or increased variety handling (e.g., workarounds) to reduce such input variety. On the one hand, IT fit needs to be managed on a continuous basis. On the other hand, ERP systems are complex system that are difficult to modify, even if needed,

because of costs associated with the changes, lack of accurate documentation, lack of highly skilled programmers, and the risk of possible dysfunctional effects [29]. Thus, organizations may instead employ micro-level informal mechanisms (i.e. workarounds) to maintain IT's alignment with changing organizational requirements. This is, in principle, matching variety. From this viewpoint, workarounds address excesses generated by the ERP system in order to maintain the stability of the organization.

IV. METHODOLOGY

This study was conducted at a leading high-technology company (HTC) in the Middle East. The HTC's primary business includes selling computer hardware, software, electronics, semiconductors, and computer services. The company has implemented an ERP solution in an attempt to streamline current processes and improve business operations.

Semi- structured interviews based on the Echo method originally developed by [14] were conducted. The Echo method is designed to investigate users' task-related interactions with an ERP system. Reference [30] defined the Echo method as follows:

A way of observing, quantifying, and describing what people value and believe is a way to describe the patterns of value and influence that are felt, verbally expressed, and often acted upon in groups or organizations (p. 4).

This method is a qualitative socio-technical interview method that identifies positive and negative aspects of any system by eliciting concrete examples of others' "helpful" and "non-helpful" task-related interactions [31]. The use of "helpful" and "non-helpful" incidents is a practical way of representing the concept of "variety" to participants, since helpful incidents from one node reduce variety on the part of the recipient node, and non-helpful incidents increase variety to the recipient node [32], [33]. That is, helpful and non-helpful incidents are situational factors that can be associated with workarounds performed within the enterprise system environment. In addition, participants were asked about the manner in which they handle each type of non-helpful incident. Corrective actions available at the recipient node are used to indicate variety handling mechanisms employed (i.e., workarounds). The purpose of this question is to find any patterns of workarounds that are specific to ERP systems.

Forty-four employees from different departments and with various hierarchal ranks in the HTC constituted the sample frame of this study (see Table 1).

TABLE I. PARTICIPANTS BY DEPARTMENT AND MANAGERIAL POSITION

Department	Number of Participants	Managers	Non-managers
Sales	6	6	-
Accounting	5	1	4
Marketing	4	2	2
Operations/ Planning	13	6	7
Production	4	-	4
IT	10	2	8
Total	42	17	25

V. RESULTS

After all of the interviews were transcribed, data were coded systematically into three main categories in accordance with the structure of the interview questions: helpful incidents, non-helpful incidents, and corrective actions to non-helpful incidents (i.e., workarounds). Summarizing data in this way is essential to preparing the data for analysis and extracting meaning. The following Table 2 summarizes the number of examples provided by participants and presents a typical example for each category.

TABLE II. FREQUENCY OF EXAMPLES PER QUALITATIVE CATEGORIES

Category	# of Examples	Typical Example
Helpful	229	User-friendly, easy to enter and modify data
Non-helpful	163	Lack of integration between different processes
Workaround	161	Export the file and manually copy and paste required records in MS Excel
Total	553	

A. Macro-level Analysis: Organizational IE Average Ratio

This study used a quantitative measure that indicates the relative effectiveness of the link between nodes in the task-related social network. Reference [34] referred to this ratio as the link's interaction effectiveness (IE) ratio. Mathematically, IE ratio is calculated by dividing the total number of helpful incidents by the total number of non-helpful incidents for all nodes. In the HTC's case, the organizational IE average ratio was estimated at 1.4. An IE ratio of 1.4 indicates that approximately three helpful incidents exist for approximately every two non-helpful incidents.

B. Micro-level Analysis: Departmental-level IE Average Ratio

At the departmental level of analysis, the range of IE ratios (from 0.86 for the marketing department to 4.5 for the production department) reflects variability in the relationships between ERP and other nodes.

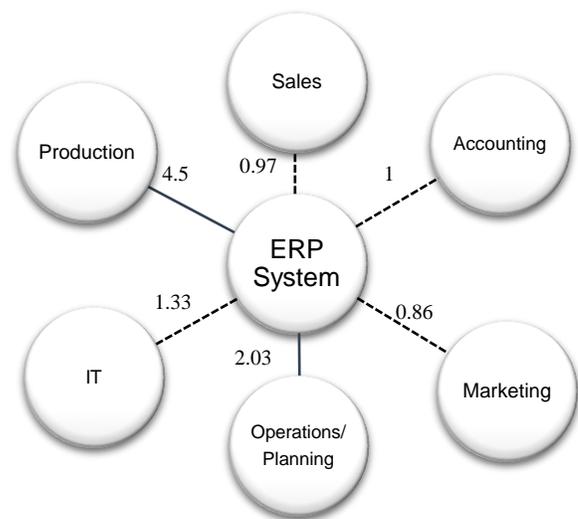


Fig. 1. IE ratios for all departments.

In Fig. 1, ratios of greater than 1.4 (indicating that effectiveness is above the organizational average) are shown as solid lines, while ratios of less than 1.4 (indicating that effectiveness is below the organizational average) are shown as dashed lines. Fig. 1 shows the IE ratio for all nodes within the HTC’s task-related social network in relation to the ERP system.

C. Categorical Analysis of the Qualitative Data

Thematic analysis was used to identify, analyze, and report patterns (themes) within data. Interview responses were classified into categories that emerged from the data, rather than relying upon predetermined categories imposed by the researcher. The collected helpful and non-helpful examples were found to cluster around four major categories: flexibility, reliability, ease of use and coordination. In addition, workaround mechanisms were found to have three common forms: using other tools, seeking help and accepting. Table 3 shows the basic properties for each category.

TABLE. III. SUBCATEGORIES FOR EACH QUALITATIVE CATEGORY AND RELATED CONTENT PROPERTY

Qualitative Category	Properties of Messages
<b>Helpful Category Labels</b>	
<b>Flexibility</b>	The ability of the system to be modified and to be responsive to different types of requests
<b>Coordination</b>	The ability to transfer information from one unit to another, and the ability of different functions/parts to work together
<b>Reliability</b>	The quality of being dependable
<b>Ease of use</b>	Requiring little effort from the user
<b>Non-helpful Category Labels</b>	
<b>Inflexibility</b>	The inability of the system to be modified or to respond to different types of requests
<b>Lack of coordination</b>	The inability to transfer information from one unit to another or for different functions/parts to work together
<b>Unreliability</b>	The quality of not being dependable
<b>Not easy to use</b>	Requiring great effort from the user
<b>Workaround Category Labels</b>	
<b>Seeking help</b>	Involving a third party to help solve the problem
<b>Using other tools</b>	The use of other means (manual or automated) to fix the problem
<b>Accepting</b>	Problems beyond the control of the recipient node

The collected data were presented under each category, and the percentages of each category were calculated to determine the category’s importance, as well as how it affects the interaction with the ERP system. Understanding these incidents may uncover flaws in ERP system design or in the implementation of ERP systems. In the following section, categories of helpful and non-helpful incidents and workaround mechanisms are discussed in more detail.

D. Enablers Alleviating Workaround Practices: Helpful Categories

As a proportion of total helpful incidents, “Reliability” (39.52%) received the highest number of favorable comments,

followed by “Ease of use” (26.66%) and “Flexibility” (24.76%). “Coordination” (9.04%) was mentioned the least. Table 4 shows the distribution of all helpful examples as perceived by all departments.

TABLE. IV. CATEGORIES OF HELPFUL EXAMPLES WITH PERCENTAGES

Helpful Categories	%	Typical Example
<b>Reliability</b>	39.52%	The data produced by the system is accurate and complete
<b>Ease of use</b>	26.66%	Easy to monitor production information
<b>Flexibility</b>	24.76%	Ability to upload all required documents directly to the system
<b>Coordination</b>	9.04%	Customers are able to report their feedback about product performance through the system

E. Barriers Elevating Workaround Practices: Non-helpful Categories

As a proportion of total non-helpful behavioral examples, “Unreliability” (42.26%) received the highest number of unfavorable comments, followed by “Inflexibility” (30.35%) and “Not easy to use” (18.45%). “Lack of Coordination” (8.92%) received the fewest examples. The following Table 5 shows the distribution of all non-helpful examples as perceived by all units.

TABLE. V. CATEGORIES OF NON-HELPFUL EXAMPLES WITH PERCENTAGES

Non-helpful Categories	%	Typical Example
<b>Unreliability</b>	42.26%	Some of the data is not accurate (e.g., equipment descriptions)
<b>Inflexibility</b>	30.35%	Unable to see the subtotal in the cost field
<b>Not easy to use</b>	18.45%	The are no clear standards for end users to follow
<b>Lack of coordination</b>	8.92%	The system does not support tracking the movement of the products between the warehouse and the printing shop in order to print the labels

F. Variety Handling Mechanisms: Types of Workaround Practices

Each participant was asked about the manner in which they address non-helpful incidents. As a proportion of total workaround mechanism examples, “Using other tools” (57.14%) received the highest number of comments, followed by “Seeking help” (23.60%). On the other hand, the lowest proportion of workaround mechanism examples fell into the category of “Accepting” (19.25%). Table 6 presents the distribution of all workaround mechanism examples as perceived by all units.

TABLE. VI. CATEGORIES OF WORKAROUND MECHANISMS WITH PERCENTAGES

Workaround Mechanisms Categories	%	Typical Example
<b>Using other tools</b>	57.14%	Using Microsoft Excel application to store and back up data
<b>Seeking help</b>	23.60%	Asking the manager to intervene
<b>Accepting</b>	19.25%	It is beyond our control

## VI. DISCUSSION

This study demonstrates that the interaction effectiveness (IE) ratio can be used in several ways to assess the interaction between users and ERP systems within organizations. One of this study's most important findings relates to the estimated organizational IE ratio (1.4), which is based on the total number of helpful and non-helpful incidents identified by respondents from all departments. The IE ratio can aid in understanding the quality of interactions with the ERP system, as well as potential sources of workarounds. This IE ratio is very important, because it is often difficult to detect and measure the quality of human-computer interactions.

In the HTC's case, the range of interaction effectiveness ratios (from 0.86 to 4.5) reflects variability in the relationships with the ERP system. Departments with an IE ratio below the organizational average of 1.4 (sales, accounting, and marketing) are considered to be ineffective, whereas nodes above the organizational average (operations/planning and production) are considered to be effective. The IT staff department falls almost on the average. Ineffective links indicate that the ERP system is perceived by users as a variety generator and is therefore sending excess input variety beyond the capacity of the users. Conceptually, the IE ratio may be used as an indicator to assess the volume of workaround practices within the organization.

One possible explanation for this wide range of IE ratios may be linked to the organizational design of tasks in terms of both personal elements (knowledge and skills the employee should possess) and structural elements (job design and requirements). For example, the skills, knowledge, and job requirements for production are different from accounting or sales job requirements. Thus, the design and the requirements of the job may affect how a user will interact with the ERP system. Moreover, the level of dependence on the technology to deliver the required work may affect the volume of interaction. For example, the production department, which yielded the highest IE ratio, may not depend on the system as much as other departments in all of its daily tasks; thus, it may see the system as effective. On the other hand, the marketing, sales, and accounting departments, whose IE ratios were less than the average, are generally more dependent upon the system to perform their daily tasks, which enables them to realize the shortcomings of the system. Also, the non-helpful incidents may have different impacts on different departments based on the routine, structure, and complexity of the work. During the interviews, one comment from production (IE 4.5) that was related to the unreliability of the system was: "It is totally dependent on the Internet; thus, if the Internet is slow or shuts down, our work performance will suffer." On the other hand, a comment from the sales department (IE 0.97) relevant to the same category was: "Delivery of items in the invoices cannot be tracked." These two comments reveal how job tasks and requirements affect different departments' interactions with the system. Moreover, the comments demonstrate the nature of the problems confronted by each department. The production problem can be easily fixed by providing a stable Internet connection, while the sales

department's problem requires more work and greater customization to fix.

In terms of the qualitative data, "Reliability" (39.52%) received the highest number of comments in the helpful category, while "Coordination" (9.04%) received the lowest number of comments. Conversely, "Unreliability" (42.26%) received the highest number of comments in the non-helpful category, whereas "Lack of coordination" (8.92%) received the lowest number of comments. The high proportion of comments pertaining to "Reliability" reflects the relative importance of this factor in ERP systems.

These results indicate that users tend to have more concerns about the reliability of the system (e.g., performance, shutdown, speed, and accuracy of results), which has a direct and high impact on the continuity of the work. Moreover, coordination received less attention, because users seem to focus more on their local goals in performing their tasks and are less concerned about relationships with or the performance of other departments. In the same way, [33] assert that coordination difficulties are common, because each department has its own goals and tends to speak its own specialized language. According to the data collected, users were less concerned about their interactions or coordination with others inside or outside the organization.

In analyzing the workaround mechanisms, "Using other tools" (57.14%) was found to be most commonly employed, followed by "Seeking help" (23.60%) and, finally, "Accepting" (19.25%). In other words, users tend to first seek to solve problems by using other tools, such as manuals or other technological solutions. If the problem is not solved by this method, users tend to seek help from others (e.g., colleagues, superiors, IT staff, etc.). If the problem is still not solved, users then tend to accept the hindrance of the system [22]. These mechanisms seem to follow a logical order depending on the cost of coordination. Seeking help involves some costs, such as time, money, effort, and delay of work, which may explain the tendency of users to employ other tools before seeking help. Seeking help as a requisite variety appears to be more difficult to execute and may create additional undesirable variety. That is, reducing the input variety is perceived as better than addressing it, and internal input variety is easier to control than external input variety. For example, undesirable variety may appear in the form of favors in which one individual expects to gain the advantage of benefiting from someone else based on a previous service. In this context, favors themselves can take the form of workaround behaviors, in which people use social mechanisms to bend rules in order to reduce input variety on the recipient end; hence resulting in reciprocal cycles of favors.

## VII. CONCLUSION AND RECOMMENDATIONS

To summarize, Ashby's Law of Requisite Variety was used to illustrate how workarounds serve to maintain stability within an organization. According to Ashby's Law of Requisite Variety, "Only variety can destroy variety" [25] (p. 207). Organizational environments are becoming more difficult to handle and predict, particularly with ongoing,

dramatic changes in technology. These ongoing changes to organizations may result in unforeseen variety that cannot be handled by a formal system. Organizations must respond to such dynamic changes with an increase in variety handling capabilities. One possible way is to engage in informal adjustments to address unmatched external variety generated by ERP systems: namely, workarounds. Workaround practices appear to be derived from misfits between input variety and variety handling capabilities. Workaround practices were found to be a useful mechanism to maintain a good fit between IT and organizational requirements, thus ensuring stability within the organization. Workarounds appear to play a significant role in adding requisite variety to the organization if implemented appropriately and communicated effectively.

The helpful and non-helpful examples provided by interviewees were divided into four main categories: flexibility, reliability, ease of use, and coordination. "Reliability" and "flexibility" are the most important categories from the users' point of view, and they received the highest number of comments from users. In light of these findings, they should be taken into account in the development of ERP systems. All findings show that workarounds are positively employed to eliminate interruptions and errors and to maintain performance throughout day-to-day tasks. The actions taken by users to solve the non-helpful behaviors were divided into three categories: using other tools, seeking help, and accepting. "Using other tools" appeared to be the best workaround mechanism for all users in all positions within the organization. Non-helpful incidents within the system (Unreliability, Inflexibility, Not easy to use, and Lack of coordination) appeared to be the main reasons that lead users to employ workarounds. In addition, users differ in the way that they interact with the ERP and in those workaround mechanisms they choose to employ. These differences stem from the nature of the tasks assigned to each department, the nature of the problems confronting users, and the power structure. These disparities lead to discrepancies in how users view the system.

Based on the analysis of the responses gathered through the interviews, three recommendations are offered to improve the usage of ERP systems. First, to increase the interaction effectiveness (IE) between ERP systems and various departments, non-helpful incidents noted in ERP systems should be reduced. The results indicate that unreliability is the major source of non-helpful incidents in terms of shutdown, poor speed, and poor performance of such systems. These problems can be solved by providing high-speed connections, scheduling preventive maintenance, and so on.

Second, the interaction between users and ERP systems should be enhanced by focusing on the helpful features and increasing their frequency of occurrence. For example, developers of ERP systems should invest more time and money into enhancing the user interface so as to increase the user-friendliness of such systems thus increasing positive interactions between the user and the system.

Third, workaround practices appear to be temporal in nature. Therefore, there is a need to provide a platform to help

transform such temporary and localized solutions into planned change. Knowledge management systems (KMS) will help to spread localized workaround experiences across the organization over longer time periods. Over time, this will help users overcome barriers imposed by the formal or centralized system, which is primarily due to the associated less flexible capabilities inherent in such organization-wide solutions (e.g., ERP systems). Establishing a knowledge management system to share information and answer users' questions will increase helpful behaviors and decrease non-helpful behaviors, and it will also offer users quick solutions to the problems they confront, thereby reducing the cost of seeking help.

#### ACKNOWLEDGMENT

The author of the paper would like to acknowledge the support brought by King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia.

#### REFERENCES

- [1] S. Dezdar, and S. Ainin, "Examining ERP implementation success from a project environment perspective," *Business Process Management Journal*, 17(6): pp. 919–39, 2011.
- [2] B. Efe, "An integrated fuzzy multi criteria group decision making approach for ERP system selection," *Applied Soft Computing*, 38: pp.106–117, 2016.
- [3] AG. Chofreh, FA. Goni, S. Ismail, and AM. Shaharoun, "A Master Plan for the Implementation of Sustainable Enterprise Resource Planning Systems (Part I): Concept and Methodology," *Journal of Cleaner Production*, 176-182, 2016.
- [4] H. Yen, and C. Sheu, "Aligning ERP implementation with competitive priorities of manufacturing firms: An exploratory study," *International Journal of Production Economics*, pp. 207-220, 2004.
- [5] P. Poba-Nzaou, L. Raymond, and B. Fabi, "Adoption and risk of ERP systems in manufacturing SMEs: a positivist case study," *Business Process Management Journal*, 1997.
- [6] W. Ke, and K. Wei, "Organizational culture and leadership in ERP implementation," *Decision Support Systems*, pp. 208-218, 2008.
- [7] T. Davenport, "Putting the Enterprise into the Enterprise System," *Harvard Business Review*, 1998.
- [8] A. Nicolaou, and S. Bhattacharya, "Organizational performance effects of ERP systems usage: The impact of post-implementation changes," *International Journal of Accounting Information Systems*, pp. 18-35, 2006.
- [9] L. Gasser, "The Integration of Computing and Routine work," *ACM Transactions on Office Information Systems*, pp. 205-225, 1986.
- [10] E. Femeley, and P. Sobreperez, "Resist, comply or workaround? An examination of different facets of user engagement with information systems," *European Journal of Information Systems*, pp. 345-356, 2006.
- [11] S. Alter, "A Workaround Design System for Anticipating, Designing, and/or Preventing Workarounds," *Advanced Information System Engineering*, June 8-12, 2015.
- [12] B. Azad, and N. King, "Enacting computer workaround practices within a medication dispensing system," *European Journal of Information Systems*, pp. 264-278, 2008.
- [13] S. Alter, "Theory of Workarounds," *Communications of the Association for Information Systems*, pp. 1041-1066, 2014.
- [14] A. Bavelas, "A Method for Investigating Individual and Group Ideology," *American Sociological Association*, 371-377, 1942.
- [15] P. Koopman, and R. Hoffman, "Work-arounds, Make-work, and Kludges," *IEEE Intelligent Systems*, pp. 70-75, 2003.
- [16] L. Petrides, S. McClelland, and T. Nodine, "Costs and benefits of the workaround: inventive solution or costly alternative," *International Journal of Educational Management*, pp. 100-108, 2004.

- [17] D. Campbell, "Public Managers in Integrated Services Collaboratives: What Works Is Workarounds," *Public Administration Review*, (72), pp. 721–730, 2012.
- [18] S. Huuskonen, and P. Vakkari, "I Did It My Way: Social workers as secondary designers of a client information system," *Information Processing and Management*, 2012.
- [19] N. Röder, M. Wiesche, M. Schermann, and H. Krcmar, "Workaround Aware Business Process Modeling," *International Conference on Wirtschaftsinformatik*, Osnabrück, Germany, 2015.
- [20] GD. Schiff, and L. Zucker, "Medical scribes: salvation for primary care or workaround for poor EMR usability? " *The Journal of General Internal Medicine*, 31(9): pp.979–981, 2016.
- [21] C. Lally, and K. Malloch, "Workarounds: The Hidden Pathway to Excellence," *Mosby Inc.*, 2010.
- [22] D. Drum, R. Standifer, and K. Bourne, "Facing the Consequences: Examining a Workaround Outcomes-Based Model," *Journal of Information Systems*, 29(2): pp. 137-159, 2015.
- [23] A. Wheeler, J. Halbesleben, and K. Harris, "How job-level HRM effectiveness influences employee intent to turnover and workarounds in hospitals," *Jornal of Business Research*, pp. 547-554, 2012.
- [24] M. Trask, *Story of cybernetics*. Studio Vista, London, 1971.
- [25] W. Ashby, *Introduction to cybernetics*, Chapman & Hall, London, 1956.
- [26] M. Boisot, and B. McKelvey, "Complexity and organization-environment relations: revisiting Ashby's Law of Requisite Variety," *The SAGE Handbook of Complexity and Management*, 279-298, 2011.
- [27] S. Beer, *Designing freedom*, Radio Canada International, Toronto, 1974.
- [28] A. Alojairi, and F. Safayeni, "The dynamics of inter-node coordination in social networks: A theoretical perspective and empirical evidence," *International Journal of Project Management*, 2011.
- [29] K. Hong , and Y. Gul Kim, "The critical success factors for ERP implementation: an organizational fit perspective," *Information & Management*, pp. 25-40, 2002.
- [30] J. Cunningham, *Researching Organizational Values and Beliefs: The Echo Approach*. Quorum Books, Westport, CT, 2001.
- [31] C. Poile, "Echo metod: Investigation socio-technical interactions," *ACM*, 2008.
- [32] P. Duimering, *Organizational impact of the just-in-time production system*, Master thesis, University of Waterloo, Waterloo, Ontario, Canada, 1991.
- [33] J. Scala, L. Purdy, and F. Safayeni, "Application of cybernetics to manufacturing flexibility: A systems perspective," *Journal of Manufacturing Technology Management*, 7(1/2): pp. 22-41, 2006.
- [34] F. Safayeni, PR. Duimering, K. Zheng, N. Derbentseva, C. Poile, and B. Ran, "Requirements engineering in new product development," *Communications of the ACM*; 51(3):pp. 77–82, 2008.