



International Journal of Advanced Computer Science and Applications

Volume 4 Issue 9

September 2013



ISSN 2156-5570(Online)

ISSN 2158-107X(Print)



www.ijacsa.thesai.org



W H E R E W I S D O M S H A R E S

INTERNATIONAL JOURNAL OF
ADVANCED COMPUTER SCIENCE AND APPLICATIONS



THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org



Editorial Preface

From the Desk of Managing Editor...

It is our pleasure to present to you the September 2013 Issue of International Journal of Advanced Computer Science and Applications.

Today, it is incredible to consider that in 1969 men landed on the moon using a computer with a 32-kilobyte memory that was only programmable by the use of punch cards. In 1973, Astronaut Alan Shepherd participated in the first computer "hack" while orbiting the moon in his landing vehicle, as two programmers back on Earth attempted to "hack" into the duplicate computer, to find a way for Shepherd to convince his computer that a catastrophe requiring a mission abort was not happening; the successful hack took 45 minutes to accomplish, and Shepherd went on to hit his golf ball on the moon. Today, the average computer sitting on the desk of a suburban home office has more computing power than the entire U.S. space program that put humans on another world!!

Computer science has affected the human condition in many radical ways. Throughout its history, its developers have striven to make calculation and computation easier, as well as to offer new means by which the other sciences can be advanced. Modern massively-paralleled super-computers help scientists with previously unfeasible problems such as fluid dynamics, complex function convergence, finite element analysis and real-time weather dynamics.

At IJACSA we believe in spreading the subject knowledge with effectiveness in all classes of audience. Nevertheless, the promise of increased engagement requires that we consider how this might be accomplished, delivering up-to-date and authoritative coverage of advanced computer science and applications.

Throughout our archives, new ideas and technologies have been welcomed, carefully critiqued, and discarded or accepted by qualified reviewers and associate editors. Our efforts to improve the quality of the articles published and expand their reach to the interested audience will continue, and these efforts will require critical minds and careful consideration to assess the quality, relevance, and readability of individual articles.

To summarise, the journal has offered its readership thought provoking theoretical, philosophical, and empirical ideas from some of the finest minds worldwide. We thank all our readers for their continued support and goodwill for IJACSA. We will keep you posted on updates about the new programmes launched in collaboration.

Lastly, 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.

We hope that materials contained in this volume will satisfy your expectations and entice you to submit your own contributions in upcoming issues of IJACSA

Thank you for Sharing Wisdom!

Managing Editor
IJACSA
Volume 4 Issue 9 September 2013
ISSN 2156-5570 (Online)
ISSN 2158-107X (Print)
©2013 The Science and Information (SAI) Organization

Associate Editors

Dr. Kohei Arai – Editor-in-Chief

Saga University

Domains of Research: Human-Computer Interaction, Networking, Information Retrievals, Optimization Theory, Modeling and Simulation, Satellite Remote Sensing, Computer Vision, Decision Making Methodology

Dr. Ka Lok Man

Xi'an Jiaotong-Liverpool University (XJTLU)

Domain of Research: Computer Science and Microelectronics

Dr. Sasan Adibi

Research In Motion (RIM)

Domain of Research: Security of wireless systems, Quality of Service

Dr. Zuqing Zuh

University of Science and Technology of China

Domains of Research : Optical Communication Systems, Optical network architecture and design, Next generation Internet, Signal processing, Broadband access network, such as cable access (DOCSIS) networks, passive optical networks (PON), fiber to the home (FTTH), Energy-efficient network and green technologies

Dr. Sikha Bagui

University of West Florida

Domain of Research: Database, database modeling, ER diagrams, XML data, web databases, data mining, association rule mining, data preprocessing

Dr. T. V. Prasad

Lingaya's University

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

Dr. Mohd Helmy Abd Wahab

Universiti Tun Hussein Onn Malaysia

Domain of Research: Data Mining, Database, Web-based Application, Mobile Computing

Reviewer Board Members

- **A Kathirvel**
Karpaga Vinayaka College of Engineering and Technology
- **A.V. Senthil Kumar**
Hindusthan College of Arts and Science
- **Abbas Karimi**
Islamic Azad University Arak Branch
- **Abdel-Hameed Badawy**
Arkansas Tech University
- **Abdelghni Lakehal**
Fsdm Sidi Mohammed Ben Abdellah University
- **Abdul Wahid**
Gautam Buddha University
- **Abdul Hannan**
Vivekanand College
- **Abdul Khader Jilani Saudagar**
Al-Imam Muhammad Ibn Saud Islamic University
- **Abdur Rashid Khan**
Gomal Unversity
- **Abeer Elkorny**
Faculty of computers and information, Cairo University
- **Ahmed Boutejdar**
- **Dr. Ahmed Nabih Zaki Rashed**
Menoufia University
- **Ajantha Herath**
University of Fiji
- **Aderemi A. Atayero**
Covenant University
- **Ahmed Mahmood**
- **Akbar Hossin**
- **Akram Belghith**
University Of California, San Diego
- **Albert Alexander**
Kongu Engineering College
- **Alcinia Zita Sampaio**
Technical University of Lisbon
- **Ali Ismail Awad**
Luleå University of Technology
- **Amit Verma**
Department in Rayat & Bahra Engineering College
- **Amitava Biswas**
Cisco Systems
- **Anand Nayyar**
KCL Institute of Management and Technology, Jalandhar
- **Andi Wahyu Rahardjo Emanuel**
Maranatha Christian University, INDONESIA
- **Anirban Sarkar**
National Institute of Technology, Durgapur, India
- **Andrews Samraj**
Mahendra Engineering College
- **Arash Habibi Lashakri**
University Technology Malaysia (UTM)
- **Aris Skander**
Constantine University
- **Ashraf Mohammed Iqbal**
Dalhousie University and Capital Health
- **Ashok Matani**
- **Ashraf Owis**
Cairo University
- **Asoke Nath**
St. Xaviers College
- **B R SARATH KUMAR**
Lenora College of Engineering
- **Babatunde Opeoluwa Akinkunmi**
University of Ibadan
- **Badre Bossoufi**
University of Liege
- **Balakrushna Tripathy**
VIT University
- **Basil Hamed**
Islamic University of Gaza
- **Bharat Bhushan Agarwal**
I.F.T.M.UNIVERSITY
- **Bharti Waman Gawali**
Department of Computer Science & information
- **Bhanu Prasad Pinnamaneni**
Rajalakshmi Engineering College; Matrix Vision GmbH
- **Bilian Song**
LinkedIn
- **Brahim Raouyane**
INPT
- **Brij Gupta**
University of New Brunswick

- **Constantin Filote**
Stefan cel Mare University of Suceava
- **Constantin Popescu**
Department of Mathematics and Computer Science, University of Oradea
- **Chandrashekhara Meshram**
Chhattisgarh Swami Vivekananda Technical University
- **Chao Wang**
- **Chi-Hua Chen**
National Chiao-Tung University
- **Ciprian Dobre**
University Politehnica of Bucharest
- **Chien-Pheg Ho**
Information and Communications Research Laboratories, Industrial Technology Research Institute of Taiwan
- **Prof. D. S. R. Murthy**
Sreeneedhi
- **Dana PETCU**
West University of Timisoara
- **Duck Hee Lee**
Medical Engineering R&D Center/Asan Institute for Life Sciences/Asan Medical Center
- **Deepak Garg**
Thapar University.
- **Dong-Han Ham**
Chonnam National University
- **Dr. Gunaseelan Devraj**
Jazan University, Kingdom of Saudi Arabia
- **Dr. Bright Keswani**
Associate Professor and Head, Department of Computer Applications, Suresh Gyan Vihar University, Jaipur (Rajasthan) INDIA
- **Dr. S Kumar**
Anna University
- **Dragana Becejski-Vujaklija**
University of Belgrade, Faculty of organizational sciences
- **Driss EL OUADGHIRI**
- **Dr. Omaira Al-Allaf**
Assistant Professor
- **Elena Camossi**
Joint Research Centre
- **Eui Lee**
- **Firkhan Ali Hamid Ali**
UTHM
- **Fokrul Alom Mazarbhuiya**
King Khalid University
- **Frank Ibikunle**
Covenant University
- **Fu-Chien Kao**
Da-Yeh University
- **G. Sreedhar**
Rashtriya Sanskrit University
- **Ganesh Sahoo**
RMRIMS
- **Gaurav Kumar**
Manav Bharti University, Solan Himachal Pradesh
- **Ghalem Belalem**
University of Oran (Es Senia)
- **Gufran Ahmad Ansari**
Qassim University
- **Giri Babu**
Indian Space Research Organisation
- **Giacomo Veneri**
University of Siena
- **Gerard Dumancas**
Oklahoma Medical Research Foundation
- **Georgios Galatas**
- **George Mastorakis**
Technological Educational Institute of Crete
- **Gavril Grebenisan**
University of Oradea
- **Hadj Hamma Tadjine**
IAV GmbH
- **Hanumanthappaj**
UNIVERSITY OF MYSORE
- **Hamid Alinejad-Rokny**
University of Newcastle
- **Harco Leslie Hendric Spits Warnars**
Budi Luhur University
- **Hardeep**
Ferozaepur College of Engineering & Technology, India
- **Hamez I. El Shekh Ahmed**
Pure mathematics
- **Hesham Ibrahim**
Chemical Engineering Department, Faculty of Engineering, Al-Mergheb University
- **Dr. Himanshu Aggarwal**
Punjabi University, India
- **Huda K. AL-Jobori**
Ahlia University
- **Iwan Setyawan**
Satya Wacana Christian University
- **Dr. Jamaiah Haji Yahaya**
Northern University of Malaysia (UUM), Malaysia
- **Jasvir Singh**
Communication Signal Processing Research Lab

- **James Coleman**
Edge Hill University
- **Jim Wang**
The State University of New York at Buffalo,
Buffalo, NY
- **John Salin**
George Washington University
- **Jyoti Chaudary**
high performance computing research lab
- **Jatinderkumar R. Saini**
S.P.College of Engineering, Gujarat
- **K Ramani**
K.S.Rangasamy College of Technology,
Tiruchengode
- **K V.L.N.Acharyulu**
Bapatla Engineering college
- **Kanak Saxena**
S.A.TECHNOLOGICAL INSTITUTE
- **Ka Lok Man**
Xi'an Jiaotong-Liverpool University (XJTLU)
- **Kushal Doshi**
IEEE Gujarat Section
- **Kashif Nisar**
Universiti Utara Malaysia
- **Kavya Naveen**
- **Kayhan Zrar Ghafoor**
University Technology Malaysia
- **Kiitimaporn Choochote**
Prince of Songkla University, Phuket Campus
- **Kohei Arai**
Saga University
- **Kunal Patel**
Ingenuity Systems, USA
- **Krasimir Yordzhev**
South-West University, Faculty of Mathematics
and Natural Sciences, Blagoevgrad, Bulgaria
- **Labib Francis Gergis**
Misr Academy for Engineering and Technology
- **Lai Khin Wee**
Biomedical Engineering Department, University
Malaya
- **Latha Parthiban**
SSN College of Engineering, Kalavakkam
- **Lazar Stosic**
College for professional studies educators
Aleksinac, Serbia
- **Lijian Sun**
Chinese Academy of Surveying and Mapping,
China
- **Leandors Maglaras**
- **Leon Abdillah**
Bina Darma University
- **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**
- **M. Tariq Banday**
University of Kashmir
- **Mazin Al-Hakeem**
Research and Development Directorate - Iraqi
Ministry of Higher Education and Research
- **Md Rana**
University of Sydney
- **Miriampally Venkata Raghavendera**
Adama Science & Technology University,
Ethiopia
- **Mirjana Popvic**
School of Electrical Engineering, Belgrade
University
- **Manas deep**
Masters in Cyber Law & Information Security
- **Manpreet Singh Manna**
SLIET University, Govt. of India
- **Manuj Darbari**
BBD University
- **Md. Zia Ur Rahman**
Narasaraopeta Engg. College, Narasaraopeta
- **Messaouda AZZOUZI**
Ziane AChour University of Djelfa
- **Dr. Michael Watts**
University of Adelaide
- **Milena Bogdanovic**
University of Nis, Teacher Training Faculty in
Vranje
- **Miroslav Baca**
University of Zagreb, Faculty of organization and
informatics / Center for biomet
- **Mohamed Ali Mahjoub**
Preparatory Institute of Engineer of Monastir
- **Mohamed El-Sayed**
Faculty of Science, Fayoum University, Egypt.
- **Mohammad Yamin**
- **Mohammad Ali Badamchizadeh**
University of Tabriz
- **Mohamed Najeh Lakhoua**
ESTI, University of Carthage

- **Mohammad Alomari**
Applied Science University
- **Mohammad Kaiser**
Institute of Information Technology
- **Mohammed Al-Shabi**
Assisstant Prof.
- **Mohammed Sadgal**
- **Mourad Amad**
Laboratory LAMOS, Bejaia University
- **Mohammed Ali Hussain**
Sri Sai Madhavi Institute of Science & Technology
- **Mohd Helmy Abd Wahab**
Universiti Tun Hussein Onn Malaysia
- **Monji Kherallah**
University of Sfax
- **Mostafa Ezziyani**
FSTT
- **Mueen Uddin**
Universiti Teknologi Malaysia UTM
- **Mona Elshinawy**
Howard University
- **N Ch.Sriman Narayana Iyengar**
VIT University
- **Natarajan Subramanyam**
PES Institute of Technology
- **Neeraj Bhargava**
MDS University
- **Noura Aknin**
University Abdelamlek Essaadi
- **Nidhi Arora**
M.C.A. Institute, Ganpat University
- **Nazeeruddin Mohammad**
Prince Mohammad Bin Fahd University
- **Najib Kofahi**
Yarmouk University
- **Na Na**
NA
- **Om Sangwan**
- **Oliviu Matel**
Technical University of Cluj-Napoca
- **Osama Omer**
Aswan University
- **Ousmane Thiare**
Associate Professor University Gaston Berger of Saint-Louis SENEGAL
- **Pankaj Gupta**
Microsoft Corporation
- **Paresh V Virparia**
Sardar Patel University
- **Dr. Poonam Garg**
Institute of Management Technology,
Ghaziabad
- **Prabhat K Mahanti**
UNIVERSITY OF NEW BRUNSWICK
- **Qufeng Qiao**
University of Virginia
- **Rachid Saadane**
EE departement EHTP
- **Raghuraj Singh**
- **Raj Gaurang Tiwari**
AZAD Institute of Engineering and Technology
- **Rajesh Kumar**
National University of Singapore
- **Rakesh Balabantaray**
IIIT Bhubaneswar
- **RashadAl-Jawfi**
Ibb university
- **Rashid Sheikh**
Shri Venkateshwar Institute of Technology , Indore
- **Ravi Prakash**
University of Mumbai
- **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, Tasmaniia, Australia
- **Rowayda Sadek**
- **Ruchika Malhotra**
Delhi Technoogical University
- **Saadi Slami**
University of Djelfa
- **Sachin Kumar Agrawal**
University of Limerick
- **Dr.Sagarmay Deb**
University Lecturer, Central Queensland University, Australia
- **Said Ghoniemy**
Taif University
- **Samarjeet Borah**
Dept. of CSE, Sikkim Manipal University
University College of Applied Sciences UCAS-
Palestine
- **Santosh Kumar**
Graphic Era University, India
- **Sasan Adibi**
Research In Motion (RIM)
- **Saurabh Pal**
VBS Purvanchal University, Jaunpur

- **Saurabh Dutta**
Dr. B. C. Roy Engineering College, Durgapur
- **Sebastian Marius Rosu**
Special Telecommunications Service
- **Selem charfi**
University of Valenciennes and Hainaut
Cambresis, France.
- **Sengottuvelan P**
Anna University, Chennai
- **Senol Piskin**
Istanbul Technical University, Informatics Institute
- **Seyed Hamidreza Mohades Kasaei**
University of Isfahan
- **Shafiqul Abidin**
G S I P University
- **Shahanawaj Ahamad**
The University of Al-Kharj
- **Shawkl Al-Dubae**
Assistant Professor
- **Shriram Vasudevan**
Amrita University
- **Sherif Hussain**
Mansoura University
- **Siddhartha Jonnalagadda**
Mayo Clinic
- **Sivakumar Poruran**
SKP ENGINEERING COLLEGE
- **Shikha Bagui**
University of West Florida
- **Sim-Hui Tee**
Multimedia University
- **Simon Ewedafe**
Baze University
- **SUKUMAR SENTHILKUMAR**
Universiti Sains Malaysia
- **Slim Ben Saoud**
- **Sumit Goyal**
- **Sumazly Sulaiman**
Institute of Space Science (ANGKASA), Universiti
Kebangsaan Malaysia
- **Sohail Jabb**
Bahria University
- **Suhas Manangi**
Microsoft
- **Suresh Sankaranarayanan**
Institut Teknologi Brunei
- **Susarla Sastry**
J.N.T.U., Kakinada
- **Syed Ali**
SMI University Karachi Pakistan
- **T C. Manjunath**
HKBK College of Engg
- **T V Narayana Rao**
Hyderabad Institute of Technology and
Management
- **T. V. Prasad**
Lingaya's University
- **Taiwo Ayodele**
Infonetmedia/University of Portsmouth
- **Tarek Gharib**
- **THABET SLIMANI**
College of Computer Science and Information
Technology
- **Totok R. Biyanto**
Engineering Physics, ITS Surabaya
- **TOUATI YUCEF**
Computer sce Lab LIASD - University of Paris 8
- **Venkatesh Jaganathan**
ANNA UNIVERSITY
- **VIJAY H MANKAR**
- **VINAYAK BAIRAGI**
Sinhgad Academy of engineering, Pune
- **Vishal Bhatnagar**
AI&T&R, Govt. of NCT of Delhi
- **VISHNU MISHRA**
SVNIT, Surat
- **Vitus S.W. Lam**
The University of Hong Kong
- **Vuda Sreenivasarao**
St. Mary's College of Engineering & Technology
- **Wei Wei**
- **Wichian Sittiprapaporn**
Mahasarakham University
- **Xiaoqing Xiang**
AT&T Labs
- **Y Srinivas**
GITAM University
- **YASSER ATTIA ALBAGORY**
College of Computers and Information
Technology, Taif University, Saudi Arabia
- **YI FEI WANG**
The University of British Columbia
- **Yilun Shang**
University of Texas at San Antonio
- **YU QI**
Mesh Capital LLC
- **ZAIRI ISMAEL RIZMAN**
UiTM (Terengganu) Dungun Campus
- **ZENZO POLITE NCUBE**
North West University
- **ZHAO ZHANG**
Deptment of EE, City University of Hong Kong

- **ZHEFU SHI**
- **ZHIXIN CHEN**
ILX Lightwave Corporation
- **ZLATKO STAPIC**
University of Zagreb

- **ZUQING ZHU**
University of Science and Technology of China
- **ZURAINI ISMAIL**
Universiti Teknologi Malaysia

CONTENTS

Paper 1: An Autonomic Auto-scaling Controller for Cloud Based Applications

Authors: Jorge M. Londoño-Peláez, Carlos A. Florez-Samur

PAGE 1 – 6

Paper 2: Partition based Graph Compression

Authors: Meera Dhabu, Dr. P. S. Deshpande, Siyaram Vishwakarma

PAGE 7 – 12

Paper 3: Performance Analysis and Comparison of 6to4 Relay Implementations

Authors: Gábor Lencse, Sándor Répás

PAGE 13 – 21

Paper 4: Review on Aspect Oriented Programming

Authors: Heba A. Kurdi

PAGE 22 – 27

Paper 5: Fingerprint Image Segmentation Using Haar Wavelet and Self Organizing Map

Authors: Sri Suwarno, Subanar, Agus Harjoko, Sri Hartati

PAGE 28 – 31

Paper 6: A Concept-to-Product Knowledge Management Framework: Towards a Cloud-based Enterprise 2.0 Environment at a Multinational Corporation in Penang

Authors: Yu-N Cheah, Soo Beng Khoh

PAGE 32 – 39

Paper 7: Pedagogy: Instructivism to Socio-Constructivism through Virtual Reality

Authors: Moses O. Onyesolu, Victor C. Nwasor, Obiajulu E. Ositanwosu, Obinna N. Iwegbuna

PAGE 40 – 47

Paper 8: A Soft Processor MicroBlaze-Based Embedded System for Cardiac Monitoring

Authors: El Hassan El Mimouni, Mohammed Karim

PAGE 48 – 53

Paper 9: Recommender System for Personalised Wellness Therapy

Authors: Thean Pheng Lim, Wahidah Husain, Nasriah Zakaria

PAGE 54 – 60

Paper 10: A Survey of Network-On-Chip Tools

Authors: Ahmed Ben Achballah, Slim Ben Saoud

PAGE 61 – 67

Paper 11: Construction of Neural Networks that Do Not Have Critical Points Based on Hierarchical Structure

Authors: Tohru Nitta

PAGE 68 – 73

Paper 12: Acoustic Strength of Green Turtle and Fish Based on FFT Analysis

Authors: Azrul Mahfudz, Sunardi, Hamzah Ahmad, Syed Abdullah Syed Abdul Kadir, Nazuki Sulong

PAGE 74 – 78

Paper 13: A Distributed Key Based Security Framework for Private Clouds

Authors: Ali Shahbazi, Julian Brinkley, Ali Karahrudy, Nasseh Tabrizi

PAGE 79 – 83

Paper 14: An Open Cloud Model for Expanding Healthcare Infrastructure

Authors: Sherif E. Hussein, Hesham Arafat

PAGE 84 – 91

Paper 15: A Novel Algorithm for Improving the ESP Game

Authors: Mohamed Sakr, Hany Mahgoub, Arabi Keshk

PAGE 92– 96

Paper 16: A quadratic convergence method for the management equilibrium model

Authors: Jiayi Zhang

PAGE 97 – 102

Paper 17: The Bitwise Operations Related to a Fast Sorting Algorithm

Authors: Krasimir Yordzhev

PAGE 103 – 107

Paper 18: Efficient Role Assignment Scheme for Multichannel Wireless Mesh Networks

Authors: Sunmyeng Kim, Hyun Ah Lee

PAGE 108 – 114

Paper 19: A Diy Approach to Uni-Temporal Database Implementation

Authors: Haitao Yang, Fei Xu, Lating Xia

PAGE 115 – 123

Paper 20: The Impact of Cognitive Tools on the Development of the Inquiry Skills of High School Students in Physics

Authors: Mohamed I. Mustafa, Louis Trudel

PAGE 124 – 129

Paper 21: A System for Multimodal Context-Awareness

Authors: Georgios Galatas, Fillia Makedon

PAGE 130 – 136

Paper 22: Improving the Security of the Medical Images

Authors: Ahmed Mahmood, Tarfa Hamed, Charlie Obimbo, Robert Dony

PAGE 137 – 146

Paper 23: Mining Positive and Negative Association Rules Using FII-Tree

Authors: T Ramakrishnu, R B V Sbramanyam

PAGE 147 – 151

Paper 24: Link-Budget Design and Analysis showing Impulse-based UWB Performance Trade-Off flexibility as Integrator Solution for Different Wireless Short-Range Infrastructures

Authors: M.S Jawad, Othman A. R, Z. Zakaria, A.A.Isa, W. Ismail

PAGE 152 – 160

Paper 25: Comparative Study of the Software Metrics for the complexity and Maintainability of Software Development

Authors: Dr Sonal Chawla, Gagandeep Kaur

PAGE 161 – 164

Paper 26: Multithreading Image Processing in Single-core and Multi-core CPU using Java

Authors: Alda Kika, Silvana Greca

PAGE 165 – 169

Paper 27: EICT Based Diagnostic Tool and Monitoring System for EMF Radiation to Sustain Environmental Safety

Authors: K Parandham, RP Gupta, Amitab Saxena

PAGE 170 – 175

Paper 28: Facing the challenges of the One-Tablet-Per-Child policy in Thai primary school education

Authors: Ratchada Viriyapong, Antony Harfield

PAGE 176 – 184

Paper 29: A Remote Health Care System Combining a Fall Down Alarm and Biomedical Signal Monitor System in an Android Smart-Phone

Authors: Ching-Sung Wang, Chien-Wei Li, Teng-Hui Wang

PAGE 185 – 188

Paper 30: Detecting LinkedIn Spammers and its Spam Nets

Authors: V´ictor M. Prieto, Manuel A´lvarez, Fidel Cacheda

PAGE 189 – 199

Paper 31: Limit Cycle Generation for Multi-Modal and 2-Dimensional Piecewise Affine Control Systems

Authors: Tatsuya Kai

PAGE 200 – 206

Paper 32: Privacy-Preserving Clustering Using Representatives over Arbitrarily Partitioned Data

Authors: Yu Li, Sheng Zhong

PAGE 207 – 212

Paper 33: Texture Classification based on Bidimensional Empirical Mode Decomposition and Local Binary Pattern

Authors: JianJia Pan, YuanYan Tang

PAGE 213 – 222

Paper 34: A Bayesian framework for glaucoma progression detection using Heidelberg Retina Tomograph images

Authors: Akram Belghith, Christopher Bowd, Madhusudhanan Balasubramanian, Robert N. Weinreb, Linda M. Zangwill

PAGE 223 – 229

Paper 35: Towards a Seamless Future Generation Network for High Speed Wireless Communications

Authors: Kelvin O. O. Anoh, Raed A. A. Abd-Alhameed, Michael Chukwu, Mohammed Buhari, Steve M. R. Jones

PAGE 230 – 235

An Autonomic Auto-scaling Controller for Cloud Based Applications

Jorge M. Londoño-Peláez
Escuela de Ingenierías
Universidad Pontificia Bolivariana
Medellín, Colombia

Carlos A. Florez-Samur
Netsac S.A.
Medellín, Colombia

Abstract— One of the key promises of Cloud Computing is elasticity – applications have at their disposal a very large pool of resources from which they can allocate whatever they need. For any fair-size application the amount of resources is significant and both overprovisioning and under provisioning have a negative impact on the customer. In the first case it leads to over costs and in the second case to poor application performance with negative business repercussions as well. It is then an important problem to provision resources appropriately.

In addition, it is well known that application workloads exhibit high variability over short time periods. This creates the necessity of having autonomic mechanisms that make resource management decisions in real time and optimizing both cost and performance. To address these problems we present an autonomic auto-scaling controller that based on the stream of measurements from the system maintains the optimal number of resources and responds efficiently to workload variations, without incurring in over costs for high churn of resources or short duration peaks in the workload.

To evaluate the performance of our system we conducted extensive evaluations based on traces of real applications deployed in the cloud. Our results show significant improvements over existing techniques.

Keywords—autonomic resource management; cloud computing

I. INTRODUCTION

Cloud Computing has radically changed the way we provision computing resources. In The Cloud one can allocate resources in a matter of seconds, use them for as long as they are needed and then release them, making it possible a pricing model where you pay for what you use. The possibility of allocating and releasing resources on demand from what seems an unlimited pool is called elasticity. Both characteristics pose interesting new problems for system administrators: How to efficiently manage resources in a cloud environment.

Ideally, the burden of dynamically managing resources in a cloud environment should be automatized so that system administrators do not have to worry for the significant workload changes that may occur in a short term basis, typically minute by minute. To do so, some service providers offer some form of auto-scaling controller that takes as input a set of measurements from the system and makes decisions about allocating or releasing resources as needed. The auto-scaling controller itself depends on a set of control parameters that determine how fast it will react to changes, how well the

allocated resources will match the workload on the system, and how much will it cost to operate the whole system.

Determining an optimal set of parameters for the auto-scaling controller is not a trivial task and it commonly done by trial and error. The main contribution of this paper is posing the auto-scaling parameter determination problem as an optimization problem that can be numerically solved and used in practical systems so that the auto-scaling controller operates in a self-managed manner. We also present an implementation of the system and tests performed on real workload traces that show it benefits.

In §II we present our problem and its application scenario. §III describes the design of our auto-scaling controller and §IV explains the self-tuning technique adopted for our controller. In §V we present the evaluation of the system based on traces of actual cloud workloads. Finally, in §VI we review the related work and highlight the differences with respect to our mechanism.

II. CONTEXT

We consider the problem of dynamically allocating resources for an application deployed in the cloud when resource requirements are not known in advanced, exhibit high variability, and are difficult to predict. This context precludes the use of schedule-based allocation mechanisms. Henceforth, we concentrate in the problem of dynamically allocating resources in an Infrastructure as a Service (IaaS) cloud taking as input various performance measurements of the system.

A. Resource allocation and measurement

It is customary that IaaS providers handle resource allocation in the form of Virtual Machines (VMs). There are usually several VM sizes and the cloud provider has no access to the guest OS running on those VMs. Therefore the monitoring infrastructure only registers variables available at the hypervisor level. We refer to this kind of measurements as blackbox metrics. It is also possible to have monitoring agents running inside the VMs which would have access to application dependent metrics. In this case they would be called whitebox metrics.

For the purpose of our proposed system we will be using only blackbox metrics; although we hypothesize the general principles stated in this paper could be extended to handle the case of whitebox metrics as well. Validation of this hypothesis is left for future work.

B. Cost model

Standard per hour pricing policy is to charge a fixed amount per hour or fraction of resource usage. The hourly rate $\alpha_{T(i)}$ is determined by the type of the i^{th} resource, namely $T(i)$. Thus, the cost of n resources is the summation

$$cost = \sum_{i=1}^n \alpha_{T(i)} \left[\frac{t_{termination,i} - t_{start,i}}{1 \text{ hr}} \right] \quad (1)$$

where $t_{start,i}$ and $t_{termination,i}$ are the times at which the i^{th} resource is started and terminated, respectively. The ceil operation captures the fact that the customer is charged the hourly rate if the resource is used for a fraction of an hour. The key observation from this equation is that churn is very inefficient. For example, launching two instances for a few minutes each would cost twice as much as launching one instance for one hour. This insight will be fundamental in supporting the termination policy used by our controller, as discussed in the next section.

III. IMPROVED AUTO-SCALING TECHNIQUE

Fig. 1 shows the overall architecture of an auto-scaling system. Its main components are:

- 1) *Smoothing*: It is responsible of obtaining measurements from the monitoring service, applying a smoothing filter (e.g. SMA, EWMA), and computing an estimate of the workload to be used for auto-scaling decisions. Other proposals [1] include the use auto regression techniques to compute these estimates.
- 2) *Controller*: Given the estimated workload it determines the optimal number of instances required. The optimizer solves the problem of minimizing the total cost of the service while satisfying the application requirements.
- 3) *Resource allocator*: Takes the number of instances requested by the controller and instructs the management infrastructure to launch/terminate instances as required.

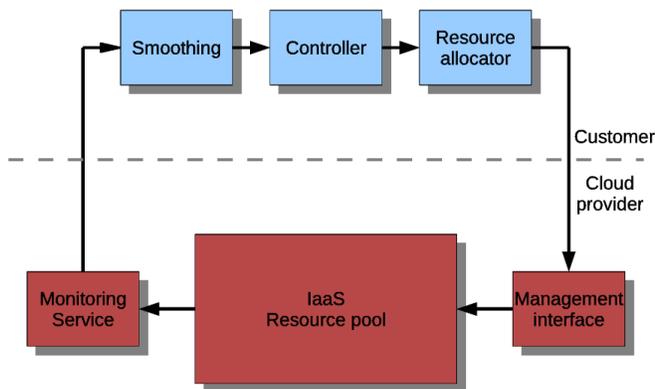


Fig. 1. Architecture of the auto-scaling system

A. Smoothing

Workload measurements typically exhibit high variability which makes resource management at small time-scales not feasible. Launching a new instance typically takes from tens of seconds to minutes and the workload contains many short duration spikes. Instead of making allocation decisions based

on short duration spikes the controller needs to identify workload variations that will persist for long enough periods of time in order to launch or terminate VMs.

Low-pass filters such as the Simple Moving Average (SMA) or the Exponentially Weighted Moving Average (EWMA) are commonly used to smooth the inputs. The SMA filter is defined as the average of the last m samples of a metric, being m the parameter to be optimized. Large values of m represent a large smoothing effect and slow response to changes. On the other hand, for small values of m the smoothed signal will closely follow the metric and give fast response to rapid changes. The EWMA filter weights the history of the signal by a series of exponentially decreasing factors. An exponential factor close to one gives a large weight to the first samples and rapidly makes old samples negligible. On the other size, a factor closer to zero gives little weight to the last samples and makes the contribution of old samples more significant, producing a more smoothed output.

With either filter, there is a tradeoff between how fast the filter reacts to persistent changes in the input and how well it smooth out short duration spikes. Hence, the optimal filter parameter is one of the key variables to be determined using the tuning technique to be described in §IV.

B. Controller

It is responsible for keeping the right number of VMs in the cluster so that the resource requirements of the application are satisfied. Overprovisioning increases the cost, so the controller's job is to find the minimum number of VMs needed by the application. In addition to the number of VMs, there may be additional type restrictions, e.g. on the amount of memory, or number of cores per VM.

For the controller, we adopt a simple hysteresis controller¹ with the following parameters:

- Thresholds: Upper and lower fractions of the resource's capacity used to trigger the launch/termination of a VM.
- Update policy: How to increase/decrease the number of instances when needed. In either case, the policy can be additive or multiplicative.

Determining the optimal parameters of the controller will be part of the job of our tuning technique.

C. Resource allocator

The resource allocator communicates with the cloud management services to launch/terminate virtual machines as indicated by the controller. As observed in §II.B, there is no point in terminating a VM if the amount of time it has been running for is not a multiple of an hour. For this reason we adopt a *lazy termination policy* that only terminates a VM if it has been running just below a multiple number of hours and the controller is requesting a lower number of VMs. In the meantime these instances contribute to handling the workload in the cluster, thus reducing response times and providing some buffer capacity to handle short load spikes at no extra cost.

¹ essentially the same one available in Amazon Web Services (AWS)

Another feature (found for example in AWS) is the cool-down period. The cool-down period prevents the resource allocator of making any changes to the system for certain amount of time. The motivation behind is to avoid frequent creation/termination of instances when the workload exhibits high variability, as this would have a negative impact on the cost of the service. We also implemented this feature in order to evaluate its importance and the associated tradeoffs.

IV. PARAMETER TUNING

Clearly, the performance of the auto scaling system depends on the setting of all the adjustable parameters. In our system we have one parameter for the smoothing filter, four parameters for the controller (fixing the update policy) and one extra parameter for the cooling period. We represent the parameters as the parameter vector τ .

In order to determine the optimal setting of these parameters, we pose the following optimization problem:

$$\operatorname{argmin}_{\theta} \{\operatorname{cost}(\theta) + \operatorname{penalty}(\theta)\}, \quad (2)$$

where $\operatorname{cost}(\theta)$ is the cost of the service as defined in (1) and $\operatorname{penalty}(\theta)$ is a measure of the workload that exceeded the allocated capacity, i.e. it measures by how far and for how long the presented workload exceeded the allocated capacity, as given by

$$\operatorname{penalty} = \beta \sum_{t=1}^T \max\{0, w_t - c_t\}, \quad (3)$$

with w_t being the workload at time t , c_t being the allocated capacity at time t , and β a weighting factor. For the purpose of the experimental evaluation we set $\alpha_{T(i)} = \beta = 1$ so that a unit of excess workload costs the same as a unit of capacity. In our model time is quantized and arrival of samples from the monitoring system and the action of the controller occur with the same sampling period (5 min in our traces).

Ideally, for tuning the controller all we need to do is to solve the optimization problem (2) for the optimal value of τ over all possible inputs to the system. Unfortunately, there is no closed form description of the input and it is easy to show that the optimal solution is not unique. Instead we adopted the following technique: Given a set of traces, we split them into a training set and a testing set.

The training set is used to conduct a numerical optimization in order to find the parameter vector θ , and the test set is used for evaluating its performance. The numerical optimization uses simulated annealing, and starting from some initial parameter vector θ_0 uses a temperature parameter to obtain neighbor parameter vectors.

The minimum among the neighbors is chosen as the starting point for the next iteration. At the end of each iteration, the algorithm reduces the temperature, thus reducing the range for choosing random neighbors. After a given number of iterations of not finding improved parameter vectors, the optimization finishes.

V. EXPERIMENTAL EVALUATION

For the evaluation of the system we used a set of traces taken from actual deployments of EC2 instances in AWS. The

traces cover a period of about 15 days and include all the metrics captured by the CloudWatch service, among others, CPU utilization, disk I/O, and network utilization. In these traces disk I/O activity was minimal, thus we made no further use of these data. From these traces we took a 50hr interval as training interval, and used other intervals as test cases. As the workload in the trace is relatively small, we scaled it up by a constant factor in order to simulate larger environments.

We implemented in Matlab the standard auto-scaling system offered by AWS and our improved system. The first one will give a baseline for comparison purposes. We then used our controller tuning technique (see §IV) to obtain the optimal parameter vector for both systems with the same training trace. Then, running the test trace in both systems we obtained the performance metrics for our analysis. The standard auto-scaling controller is setup to use a SMA smoothing filter and multiplicative increase/decrease policies. Our controller uses an EWMA filter and multiplicative increase/decrease policies. Both systems implement the cool-down policy.

A. Number of active instances over time

A first experiment evaluates the number of active instances over time and their liveness period, i.e. the amount of time they existed in the cloud. Fig. 2 shows the number of compute units (CU)² for an illustrative test case. It shows both cases, standard and improved auto-scaling.

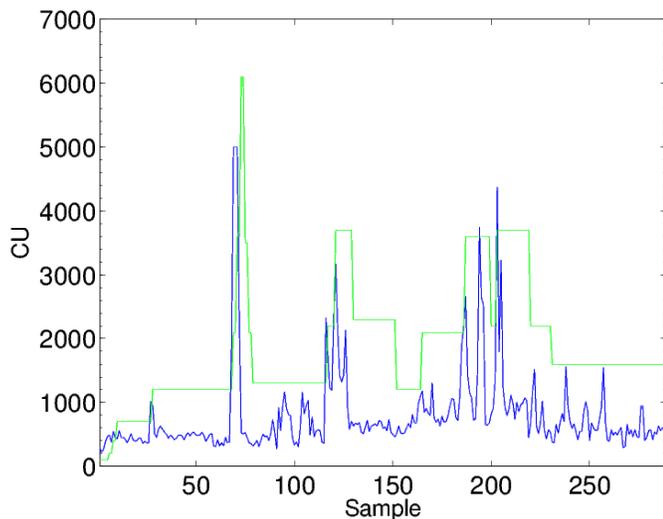
Although both track the workload (plus a safety margin as defined by the upper threshold), it is noticeable that the improved auto-scaling tracks more closely the peaks and the valleys and reacts faster to changes. This is especially remarkable in the valleys where the lazy termination policy keeps instances alive for up to a multiple of one hour, but still does a better job tracking the valleys than the cool-down policy.

The effect of the lazy termination policy is shown in Fig. 3 which shows the histogram of the liveness period of instances for the same test cases. For standard auto-scaling they take arbitrary lengths, and for improved auto-scaling they always take a multiple of twelve sampling periods minus one ($12n-1$). The minus one is due to our implementation terminating instances one sampling period before the hour to avoid using any fraction of the next hour.

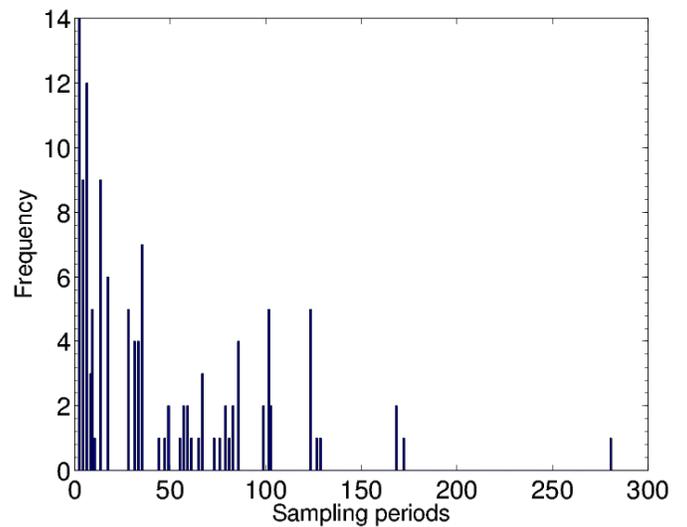
It is important to notice that our parameter tuning algorithm set the cool-down period to 3 sampling periods in the case of the standard auto-scaling and to 0 in the case of the improved auto-scaling technique. This is the result of solving the optimization (2), which in the first case is forced to keep the cool-down period larger than zero to void the problem of frequent termination/creation, which would significantly increase the cost.

In the second case, the cool-down period does not play any role in determining the cost of the service, as termination of instances is governed by the lazy termination policy.

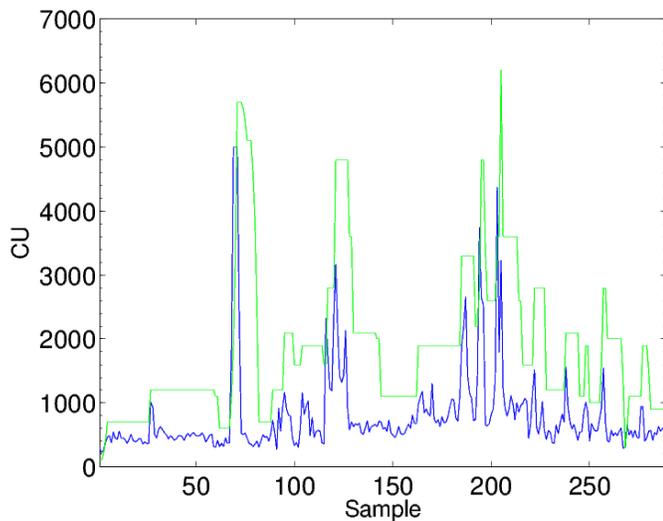
² For our analysis we set 100 CU equal to 1 Amazon's ECU



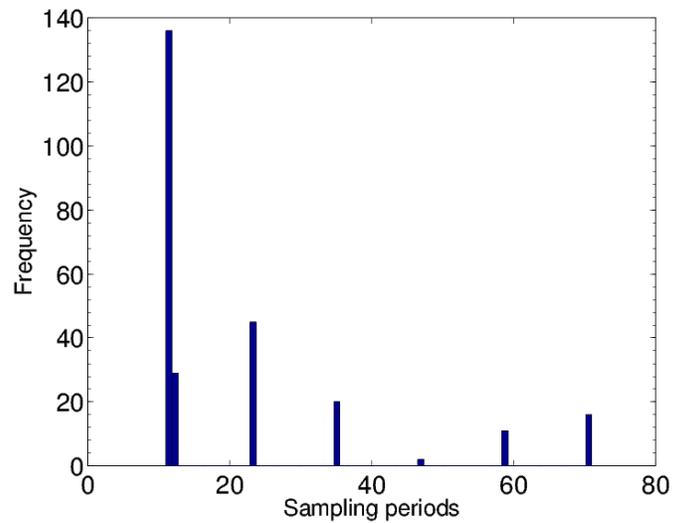
a) Standard auto-scaling



a) Standard auto-scaling



b) Improved auto-scaling



a) Improved auto-scaling

Fig. 2. Number of active instances over time

Fig. 3. Histogram of instance liveness period

B. Cost and penalty improvements

Fig. 4 shows the costs and penalties for all the test cases considered. Overall the improved technique has a small reduction of service cost, albeit a few exceptions. However, the comparison of penalties shows a significant reduction of the penalty with the improved algorithm. Considering all test cases, the average reduction of the cost was 6.3% and the average reduction of the penalty was 55.5%.

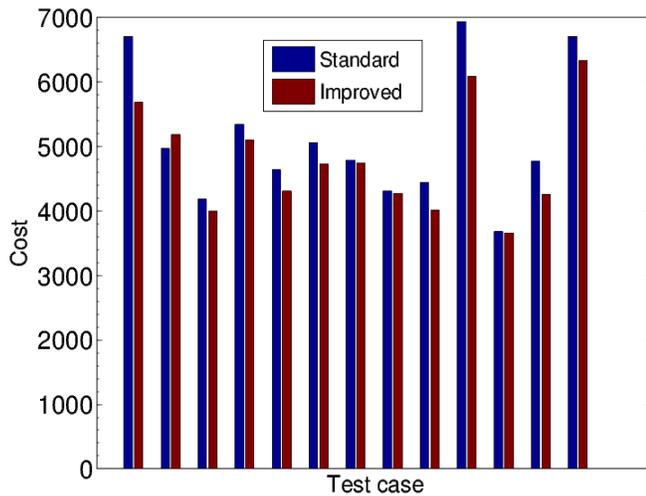
VI. RELATED WORK

Existing literature considers various approaches for handling the allocation of resources in a cloud computing environment. Although some approaches have some features in common with our solution, there are also important differences. Following we present a brief description of the most relevant approaches and highlight the main differences with respect to our work.

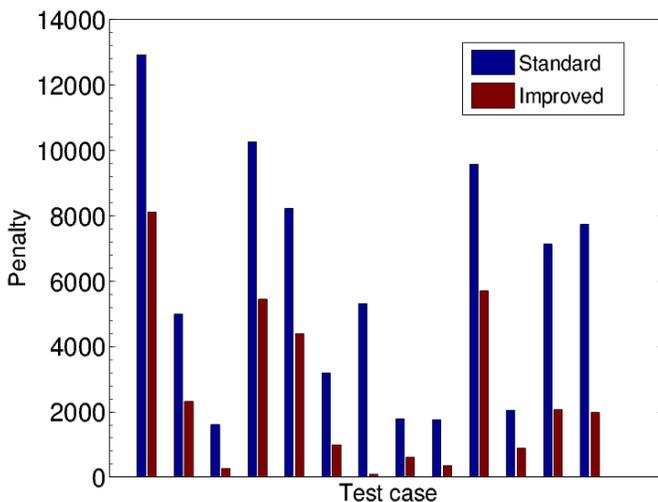
Bodik et al [2] present a technique that uses statistical machine learning to fit a non-linear performance model on the most recent set of samples. This model produces a target number of servers to satisfy the existing Service Level Agreements (SLAs), and this value is filtered through a hysteresis filter to avoid oscillations in the controller. The model captures the relationship between the number of request that fail the SLA's threshold and the current number of servers and workload. This technique does not take into account the cost of the service and under high variability of the workload would lead to frequent creation and termination of VMs, the churn problem that our technique avoids while minimizing the operational cost of the service.

Bi et al [3] developed a technique that uses a hybrid queueing model as the basis to provision resources for multi-tier applications running in a cloud data center. Their technique takes as inputs the request arrival rate, the service rate of the VMs for each tier, and the response time from the application.

Feeding this information into the model, makes it possible to determine the number of VMs required. The main difference with our system is that this technique relies on application level measurements, which may not always be available or could require additional development and integration effort.



a) Comparison of costs



b) Comparison of penalties

Fig. 4. Comparison of costs and penalties

Padala et al [4] adopt a blackbox approach that uses control theory to manage the virtual resources assigned to the application. The controller in this system assigns entitlements of the physical resources to the VMs they host, in such a way that the application performance meets the preset SLAs. The main difference with our work is the assumption that entitlements to physical resources can be adjusted online. Although this functionality is available in several virtualization frameworks, it is not commonly offered by public IaaS providers, which prefer to offer a set of predefined instance sizes.

Turner et al [5] explore a system that builds an empirical model of the application performance. Their system is tailored for multi-tier applications running on a virtualized

infrastructure. Data collected by the monitoring system includes resource consumption and application response time. A regression algorithm fits a model to the collected data and this model is used to adjust allocation of resources to the different virtual machines. Sangpetch et al [6] further develop a close-loop controller system that uses the model and a target Service Level Objective (SLO) to adjust the allocation of resources to each VM in the system. The controller uses both, a long term and a short term prediction to adjust the resource allocation to each of the VMs. These systems also rely on the assumption that the customer has control over the amount of resources assigned to each VM, which is not usually the case with IaaS providers.

Chandra et al [1] present a technique that combines measurements, a generalized processor sharing model, and time-series analysis to determine the fraction of the resources to assign to each of the application components. The allocated resources assure that the application meets its Quality of Service (QoS) constraints. This technique applies to the case of VMs sharing a host in which the entitlement of resources for each component is adjustable and the controller has access to internal performance metrics of the application. However, it is different to the problem we are handling because we deal with predetermined instance sizes, our goal is minimizing the service cost and penalty, and we restrict to blackbox metrics.

It is also worth noticing that there has been work on resource management mechanisms based on the idea of migrating VMs, as for example [7]. In this study migration was not considered as our focus is a public IaaS cloud, where migration services are not commonly available.

The problem of determining the optimal set of resources of various types with multiplicity been shown to be NP-Complete by Chang et al [8]. They also present an approximation algorithm. This algorithm considers the problem in a static setting, thus it is not applicable in the dynamic environment we consider. A related work by Dougherty et al [9] considers the auto-scaling problem from the point of view of minimizing the cost and energy consumption. In their work they used a Model Driven Engineering (MDE) approach combined with a constraint satisfaction technique to find the set of instances that supply the application requirements while minimizing cost and energy. This work assumes a static context where the application requirements remain stable over time.

VII. CONCLUSIONS

We have presented an autonomic auto-scaling controller specifically designed for allocating resources in a cloud datacenter under dynamic workloads. Our controller reduces the service cost and the performance penalties when compared to the optimized standard hysteresis controller commonly available from public cloud providers. Both characteristics are highly desirable for whoever deploys an application in an IaaS cloud.

Our controller departs from well-known auto-scaling controllers by incorporating a fast response smoothing filter, a numerical optimization technique for finely tuning the controller parameters, and implementing the lazy termination policy, which postpones the decision to terminate an instance

until the very last moment possible without extra charges. Also, our experiments showed that the lazy termination policy effectively makes the cool-down period unnecessary. The cool-down period limits the response time in the event of large workload changes, thus increasing performance penalties. On the other hand, the cool-down period does not reduce the service cost when using our improved controller

REFERENCES

- [1] Abhishek Chandra, Weibo Gong, and Prashant Shenoy, "Dynamic resource allocation for shared data centers using online measurements," SIGMETRICS Perform. Eval. Rev., vol. 31, pp. 300-301, 2003.
- [2] Peter Bodík et al., "Statistical Machine Learning Makes Automatic Control Practical for Internet Datacenters," in HotCloud, 2009.
- [3] Jing Bi, Zhiliang Zhu, Ruixiong Tian, and Qingbo Wang, "Dynamic Provisioning Modeling for Virtualized Multi-tier Applications in Cloud Data Center," in IEEE International Conference on Cloud Computing, 2010.
- [4] Pradeep Padala et al., "Adaptive control of virtualized resources in utility computing environments," SIGOPS Oper. Syst. Rev., vol. 41, pp. 289-302, 2007.
- [5] Andrew Turner, Akkarit Sangpetch, and Hyong S. Kim, "Empirical Virtual Machine Models for Performance Guarantees," in Large Installation System Administration Conference (LISA), 2010.
- [6] Akkarit Sangpetch, Andrew Turner, and Hyong Kim, "How to Tame Your VMs: An Automated Control System for Virtualized Services," in Large Installation System Administration Conference (LISA), 2010.
- [7] Mauro Andreolini, Sara Casolari, Michele Colajanni, and Michele Messori, "Dynamic load management of virtual machines in a cloud architectures," in IEEE 2009 International Conference on Cloud Computing, Los Angeles, CA, 2009.
- [8] F. Chang, J. Ren, and R. Viswanathan, "Optimal Resource Allocation in Clouds," in IEEE International Conference on Cloud Computing, 2010.
- [9] Brian Dougherty, Jules White, and Douglas C. Schmidt, "Model-driven auto-scaling of green cloud computing infrastructure," Future Gener. Comput. Syst., vol. 28, no. 2, pp. 371-378, 2012.

Partition based Graph Compression

Meera Dhabu, Dr. P. S. Deshpande, Siyaram Vishwakarma
Department of Computer Science & Engineering,
Visvesvaraya National Institute of Technology,
Nagpur – 440010 (India)

Abstract—Graphs are used in diverse set of disciplines ranging from computer networks to biological networks, social networks, World Wide Web etc. With the advancement in the technology and the discovery of new knowledge, size of graphs is increasing exponentially. A graph containing millions of nodes and billions of edges can be of size in TBs. At the same time, the size of graphs presents a big obstacle to understand the essential information they contain. Also with the current size of main memory it seems impossible to load the whole graph into main memory. Hence the need of graph compression techniques arises. In this paper, we present graph compression technique which partition graphs into subgraphs and then each partition can be compressed individually. For partitioning, proposed approach identifies weak links present in the graph and partition graph at those weak links. During query processing, the partitions which are required need to be decompressed, eliminating decompression of whole graph.

Keywords— graph compression; su graph; partitioning

I. INTRODUCTION

Today, numerous large-scale systems and applications need to analyze and store massive amounts of data that involve interactions between various entities – this data is best represented as a graph; for instance, the link structure of the World Wide Web, group of friends in social networks, data exchange between IP addresses, market basket data, etc., can all be represented as massive graph structures. As witnessed in the core tasks of these applications graph patterns could help build powerful, yet intuitive models for better managing and understanding complex structure. Some of these application domains are [19]:

- World Wide Web. The Web has a natural graph structure with a node for each page and a directed edge for each hyperlink. This link structure of the Web has been exploited very successfully by search engines like Google [18] to improve search quality. Other contemporary research works mine the Web graph to find dense bipartite cliques, and through them Web communities [16] and link spam [05]. Recent estimates from search engines put the size of the Web graph at around 3 billion nodes and more than 50 billion arcs [14]. (Note that these are clearly lower bounds since the Web graph has been growing rapidly over the years as more of the Web gets discovered and indexed.) Thus, the Web graph can easily occupy many terabytes of storage.
- Social Networking. Popular social networking websites like Facebook, MySpace and LinkedIn cater to millions

of users at a time, and maintain information about each user (nodes) and their friend-lists (edges). Mining the social network graph can provide valuable information on social relationships between users, the music, movies, etc. that they like, and user communities with common interests.

- IP Network Monitoring. IP routers export records containing source and destination IP addresses, number of bytes transmitted, duration, etc. for each IP communication flow. Recently, Iliofotou et. al. [12] proposed the idea of extracting Traffic Dispersion Graphs (TDGs) from network traces, where each node corresponds to an IP address and there is an edge between any two IP addresses who sent traffic to each other. Such graphs can be used to detect interesting or unusual communication patterns, security vulnerabilities, hosts that are infected by a virus or a worm, and malicious attacks against machines.
- Market Basket Data. Market basket data contains information about products bought by millions of customers. This is essentially a bipartite graph with an edge between a customer and every product that he or she purchases. Mining this graph to find groups of customers with similar buying patterns can help with customer segmentation and targeted advertising.

Several approaches have been proposed for the analysis and discovery of concepts in graphs in the context where graphs are used to model datasets. Modeling objects using graphs allows us to represent arbitrary relations among entities and capture the structural information. The utilization of richer and more elaborate data representations for improved discovery leads to larger graphs. The graphs are often so large that they cannot fit into the dynamic memory of conventional computer systems. Even if the data fits into dynamic memory, the amount of memory left for use during execution of the discovery algorithm may be insufficient, resulting in an increased number of page swaps and ultimately performance degradation. One of the main challenges for knowledge discovery and data mining systems is to scale up their data interpretation abilities to discover interesting patterns in large datasets. This paper addresses the scalability of graph-based discovery to monolithic datasets, which are prevalent in many real-world domains where vast amounts of data must be examined to find meaningful structures.

In [23], many challenges are faced by graph mining algorithms due to the huge size of graph. One issue is that a huge graph may severely restrict the application of existing

pattern mining technologies. Additionally, directly visualizing such a large graph is beyond our capability. In computer science, it is more important to understand the information embodied in abstract structures that are of our particular interests. For instance, how can we quantify the amount of information in the structure of graphs such as the Internet, social networks, and biological networks? How can we understand and utilize the “structure” of nonconventional data structures such as biological data, topographical maps, medical data, and volumetric data? Imagine a compressed graph, conserving the characteristics of the original graph. We can easily visualize it. The goal of compressing a graph is to make the high-level structure of the graph easily understood. Therefore, informative graph compression techniques are required and have wide application domains. Many graph compression techniques have been developed for compressing a web graph [7, 14, 25, 10, 4, 9]. In this paper we proposed partition based compression approach which helps in storing the compressed subgraphs on the systems that are located geographically apart. Thus it reduces the network traffic in distributed computing [6] since data will be available on local system itself. The aim of the proposed technique is to represent the data in compressed form while retaining the ability to answer the same queries as their uncompressed counterpart. We aim at representing graphs in highly compressed form, so as to manage huge instances in main memory.

The remainder of this paper is organized as follows. Section II reviews the background information as well as related work on graph compression. Section III presents the details of proposed partition based approach. Section IV presents the results of performance evaluation. Section V summarizes and concludes our paper.

II. BACKGROUND

The biggest challenge in graph compression is ever increasing demand of high compression ratio, which reduces memory requirement of a graph. A graph containing billions of nodes and trillions of edges cannot be stored in memory without compression and if we store it on disk then operations which need to be performed on this graph would require many disk I/O and disk seek operations which reduces algorithm performance drastically. Hence a graph needs to be divided to ensure that each partition is small enough to fit in main memory and thus reduces I/O operations significantly.

A. Problem definition

Given an undirected graph $G = (V, E)$, where V is set of vertices and E is set of edges in the graph G . We need to represent graph such that the compression ratio and bits per edge are maximum and minimum respectively. Compression ratio and bits per edge are given by the following formulae:

$$\text{Percentage Compression or Compression Ratio} = \frac{\text{input graph size} - \text{output graph size}}{\text{input graph size}} * 100$$

$$\text{Bits per edge} = \frac{\text{size of output or compressed graph}}{\text{total edges in the graph}}$$

B. Related work

In recent years many compression algorithms have been proposed. In [14] Gap encoding makes use of locality [8] property of web graph. Locality suggests that each list of successors should be represented as list of gaps. More precisely, if $S(x) = (s_1, s_2, \dots, s_k)$, then it can be represented as $(s_1 - x, s_2 - s_1 - 1, s_3 - s_2 - 1, \dots, s_k - s_{k-1} - 1)$. However; reference compression [14] technique exploits similarity property of web graphs. In this method, adjacency list $S(x)$, is represented as a “modified” version of some list $S(y)$, called the *reference list*. The difference $x - y$ is called the *reference number*. This results into *reference compression*, in which a sequence of bits, one bit for each successor in the reference list, tells whether the corresponding successor of y is also a successor of x . Nodes which are not covered by reference list are called *extra nodes*.

In differential compression, the differences with $S(y)$ are represented as a sequence of *copy blocks*. *Copy list* can be represented as an alternating sequence of 1 and 0-blocks, and specify the length of each block. This sequence of integers is preceded by a block count telling the number of blocks that will follow [14]. Consecutivity among extra nodes is frequent, hence to exploit this consecutivity, subsequences are isolated corresponding to integer intervals and number of integers in these intervals is called *length* [14].

In [8], an un-weighted graph $G = (V_G, E_G)$ can be represented as $R = (S, C)$ where $S = (V_s, E_s)$ is graph summary and C is set of edge corrections. Every node v in V_G belongs to a *super node* V in V_s which represents a set of nodes in G . A *super edge* $E = (V_i, V_j)$ in E_s represents the set of all edges connecting all pairs of nodes in V_i and V_j i.e. it simply collapse one bi-partite graph into two super nodes V_i and V_j and replaces all the edges by super edge between the super nodes. The edge correction C has parts $+e$ (edge to be added) and $-e$ (edge to be removed) which is considered during recreation of original graph.

In [7], Re-pair recursively finds pair of repeated symbols across all the lists and replace them by a new “non-terminal” symbol which has to be expanded later when extracting the lists. In [3], a directed bipartite clique $G = (V, E)$ can be transformed into a directed star. A directed bi-partite clique (S, T) is a pair of two disjoint set S and T such that $u \in S$ and $v \in T$ and there is a directed link from u to v in G . For a bi-clique (S, T) a new compressed graph $G' = (V', E')$ is formed by adding a new vertex x to the graph, removing all the edges in (S, T) and adding a new edge $ux \in E'$ for each $u \in S$ and new edge $xv \in E'$ for each $v \in T$.

In [13], an undirected graph $G = (V, E)$, where V is a set of nodes and E is set of edges, is represented using adjacency list method and thus $m + n$ space is required, where $n = |V|$ and $m = |E|$. But for simple undirected graphs G , it should be noted that the complement graph G^c of G is sufficient for representing G . For a very dense graph G , the size of the edge set of the complement graph may be much less than m . Therefore, the original graph is store in a data structure if $m \leq \frac{n(n-1)}{4}$, and the complement graph if $m > \frac{n(n-1)}{4}$. this

method requires $n + \min(m, m^c)$ space, where $m^c = \frac{n(n-1)}{2} - m$.

TABLE I. NOTATIONS

Notation	Description
w	Maximum no. of edges in bridge
neb_i	Set of neighbors of node i
S_1, S_2	Sets S_1 store neighbor of node i and set S_2 stores rest nodes
deg_i	Degree of node i
$G = (V, E)$	A graph G where V is set of nodes and E is set of edges

III. PROPOSED PARTITION BASED APPROACH

Compression allows more efficient storage and transfer of graph data, and may improve the performance of various algorithms by allowing computation to be performed in faster levels of computer memory hierarchies. Good compression requires using the structural properties of the graph, and hence first important step is to understand this structure. For example, in Web graphs, there appear to be natural clusters of related pages with similar connections.

In this paper we restrict our discussion to undirected graph but can be easily extended for directed graph. We use an undirected graph for modeling complicated structures which contains dense clusters and these dense clusters are connected with weak links called bridges. Proposed partition based compression algorithm exploits graph property *locality*.

Link locality has been independently observed and reported by several authors. For instance, Suel and Yuan [16] observe that on average, around three-quarters or 75% of the links from a page are to other page on the name domain/host. Given this observation, we attempt to partition graph into dense clusters and then these dense clusters are further compressed using reference compression technique.

We employ breadth first search algorithm starting from a randomly chosen node, say $x \in V$ which returns a connected component C . Now take a node say i from C and make two sets S_1 and S_2 . Set S_1 contains all neighbors of node and set S_2 contains all the nodes in C except the nodes in S_1 . A node x in S_1 with degree d is a bridge node of width w if the following conditions are satisfied:

- 1) $d_x - w$ neighbors of node x are in set S_1 and exactly w neighbors are in set S_2 .
- 2) $(|neb_y| - 1)$ neighbors of node y are in S_2 , where $y \in neb_x$ and $y \in S_2$.

If both the conditions are not satisfied then node may be shifted to set S_2 or if more than half neighbors of node x are in set S_2 . We repeat this process for all the nodes in set S_1 and S_2 until we find a bridge between the two sets or no change in set S_1 and S_2 . In this way we find bridge between set S_1 and S_2 which results into two subgraphs. Repeat the same procedure by choosing another random node from C until we get sufficiently small subgraphs. These subgraphs are

compressed sequentially using reference compression technique [14].

Each subgraphs thus obtained after partitioning is compressed by applying reference compression algorithm. In this method, instead of representing adjacency list $S(x)$ for node x directly, it is represented as a "modified" version of some next list $S(y)$, called the *reference list*. The difference $y - x$ is called the *reference number*. Thus the reference compression results in a sequence of bits, one for each successor in the reference list, which tells whether the corresponding successor of node y is also a successor of node x . The representation of $S(x)$ with respect to $S(y)$ is made of two parts: a sequence of $|S(y)|$ bits, called the *copy list*, and the list of integers $S(x)/S(y)$, called the *list of extra nodes*. *Copy list* specifies which of the links contained in the reference list should be copied: it will contain 1 at the i^{th} position; iff the i^{th} entry of $S(y)$ also appears in $S(x)$ [14].

For each node i in a subgraph s , we find reference node j (node which has maximum number of common successors with node i). we consider *reference_width* for finding reference node. *reference_width* can be fixed or can be equal to size of subgraph.

For reference node, we calculate *reference number*, *copylist* and *extra nodes*. *copylist* is further compressed as a sequence of *copyblock* which contains the information about the number of 1's and 0's appearing in *copylist* alternatively. Further *extranodes* are compressed since there is consecutively among *extranodes*. Once all the nodes are covered in subgraph we take next subgraph for compression.

A. Algorithm

In this section pseudo-code for partitioning the large graph is given. Function *check_condition_1* and *check_condition_2* will return "1" if *condition_1* and 2 mentioned in section III is true for node k .

Arrange all nodes of G in decreasing order of *degree*.

Procedure PartitionGraph(G, w)

```

begin
while ( $w > 0$ )
{
for each node  $i \in V$ 
{
 $S_1 = i \cup neb_i$ ;
 $S_2 = V - S_1$ ;
while (no change in  $S_1$  and  $S_2$  or bridge is not found)
{
for each node  $k \in S_1$ 
{
flag_1 = check_condition_1( $S_1, S_2, k$ )
if (flag_1 == true)
{
flag_2 = check_condition_2( $S_1, S_2, k$ )
if (flag_2 == true)
{
a. Node  $k$  is a bridge node of width  $w$ .
b. Remove all  $w$  edges between  $k$  and  $S_2$ .
}
}
}
}
}
}

```

```

else if (more than  $\frac{|neb_k|}{2}$  nodes are in  $S_2$ )
{
     $S_1 = S_1 - k;$ 
     $S_2 = S_2 \cup k;$ 
}
for each node  $k \in S_2$ 
{
     $flag\_1 = check\_condition\_1(S_2, S_1, k)$ 
    if ( $flag\_1 == true$ )
    {
         $flag\_2 = check\_condition\_2(S_2, S_1, k)$ 
        if ( $flag\_2 == true$ )
        {
            a. Node  $k$  is a bridge node of width  $w$ .
            b. Remove all  $w$  edges between  $k$  and  $S_1$ .
        }
    }
else if (more than  $\frac{|neb_k|}{2}$  nodes are in  $S_1$ )
{
     $S_2 = S_2 - k;$ 
     $S_1 = S_1 \cup k;$ 
}
}
}
w--;
}end while;
end-begin.

```

IV. PERFORMANCE EVALUATION

In this section, we present experimental results. We have performed experiments on 2.10 GHz Intel core i3 processors with 4GB main memory, running on 32-bits Windows 7 platform. Proposed algorithm is implemented in Java. We performed experiment on synthetic datasets generated using graph generator. Details of the graph dataset used for experiments are given in Table II.

A. Graph Partitioning

A graph having 9985 nodes and 123416 edges is partitioned into 354 subgraphs, size of each subgraph varies from 26 to 30 nodes both inclusive with bridge width equal to 3. Whereas number of bridges is 510, among these 358 bridges are of width one, 98 bridges are of width two and 54 bridges are of width three. On the other hand, a graph of 1979 nodes and 24340 edges is partitioned into 71 subgraphs, size of each subgraphs again vary from 26 to 30 both inclusive where bridge width is three. Number of bridges is 98 among these 66 bridges are of width one, 20 bridges are of width two, 10 bridges are of width three.

TABLE II. DETAILS OF GRAPH DATASET

Id	No. of nodes	No. of edges
G ₁	1979	24340
G ₂	4993	61392
G ₃	9985	123416

B. Effect of different parameters on compression ratio

In Fig. 1, y-axis represents compression ratio and x-axis represents reference width w [9]. Different reference widths are 3, 5, 7 and a *subgraph*. Reference width equal to the subgraph means all the nodes in the subgraph will be considered in search of reference node [10]. Fig.1. shows compression ratio without copy blocks for different graph size. Compression ratio increases slowly with the increase in reference width. Fig. 2 shows compression ratio with copy blocks for different graph size. Fig. 3 shows compression ratio with copy blocks and extra nodes for different graph size. From Fig. 2 and 3 can observe that the compression ratio increases rapidly with the increase in reference width. When reference width is equal to subgraph compression ratio is maximum. For all reference width, compression ratio of the graph with 9985 nodes and 123416 edges is higher among all graphs. Hence ratio increases with increase in number of nodes and edges i.e. we get better compression ratio for dense graphs.

Boldi and Vigna [14] have given the best algorithm ever which takes 2 to 3 bits per edge for a graph of size 18.5 million nodes and 300 million edges. S. Raghavan [21] has shown that super node and super edge representation takes 5.07 bits per edge for average over 25 million, 50 million, 100 million nodes. Broder [1] showed that a graph of 200M nodes and 1.5G edges requires 37.87 bits per edge.

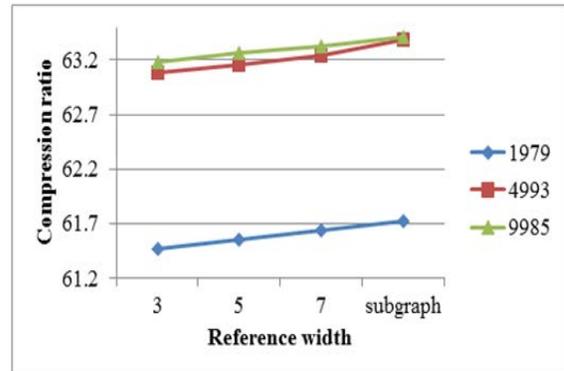


Fig. 1. Compression ratio (without copy block) v/s reference width w .

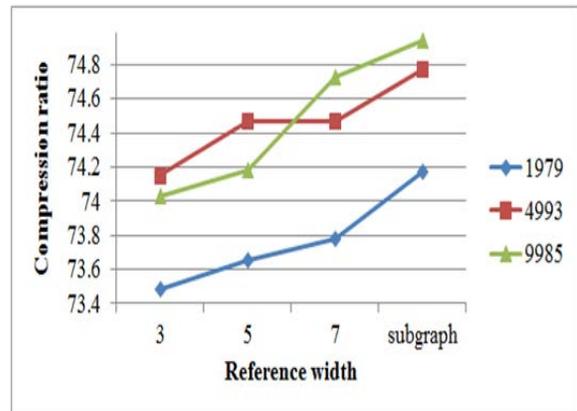


Fig. 2. Compression ratio (with copy block) v/s reference width w .

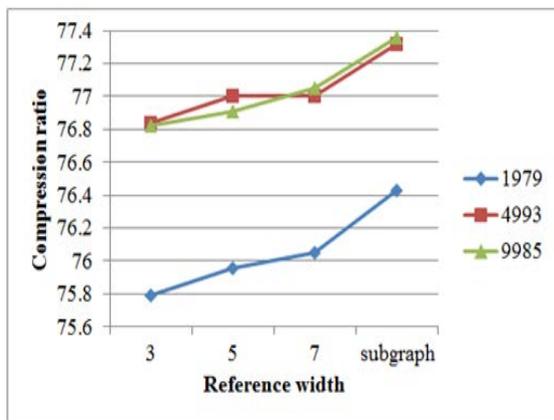


Fig. 3. Compression ratio (with copy block and extra nodes) v/s reference width w.

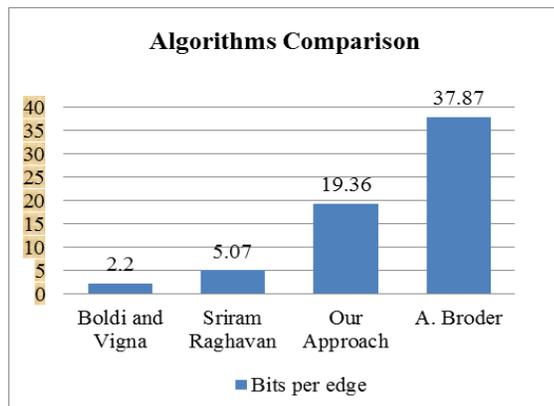


Fig. 4. Comparison of Boldi and Vigna, Sriram Raghvan, and A. Broder with Partition based reference compression approach.

V. CONCLUSION & FUTURE WORK

In this paper we proposed an effective solution in the form of a partitioning approach, to one of the main challenges for graph-based knowledge discovery and data mining systems, which is to scale up their data interpretation abilities to discover interesting patterns in large graph datasets. We observed that for partition based reference compression approach, compression ratio increases with increase in reference width and it is maximum when reference width is equal to size of subgraph. Moreover it helps in distributed computing by reducing network traffic and storage burden on single system.

Possible future enhancement to the proposed approach is reducing partitioning time which increases sharply with the increase in graph size. Since we run BFS algorithm for each node in the graph which gives connected component but we can ignore the nodes which are in the partition and cannot be partitioned further.

Our algorithm is sequential i.e. first graph partitioning is done and then reference compression algorithm is applied. This causes re-loading of each partition for the compression. It can be improved by compressing the partition when it cannot be partitioned further.

REFERENCES

- [1] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the web", *Journal of Computer Networks*, pp. (1-6), 2000.
- [2] A. Inokuchi, T. Washio and H. Motoda, "An apriori-based algorithm for mining frequent substructures from graph data", In *PKDD'00*, pp. 13-23, 2000.
- [3] C. Cranor, T. Johnson, O. Spatschek, and V. Shkapenyuk, "Gigascope: A stream database for network applications" In *SIGMOD*, pp. 647-651, 2003.
- [4] D. Blandford, G. Blelloch, I. Kash, "Compact representation of separable graphs" In *Proc. of the 14th SODA*, pp. 679-688, 2003.
- [5] D. Gibson, R. Kumar, and A. Tomkins, "Discovering large dense subgraphs in massive graphs" In *VLDB*, pages 721- 732, 2005.
- [6] F. Zhou, "Graph Compression", Department of Computer Science and Helsinki Institute for Information Technology HIIT, pp. 1-12.
- [7] F. Claude and G. Navarro, "Fast and compact web graph representations", In *Proceedings of 14th International Symposium on String Processing and Information Retrieval*, Santiago, Chile, pp. 105-116, October 29-31, 2007.
- [8] H. Khalili, A. Yahyavi and F. Oroumchian, "Web graph pre compression for similarity based algorithms", *Proceedings of the 3rd International Conference on Modeling, Simulation and Applied Optimization Sarjah, U.A.E.*, pp. 20-22, January 2009.
- [9] J. Guillaume, M. Latapy, L. Viennot, "Efficient and simple encodings for the web graph", In Meng, X., Su, J., Wang, Y. (eds.) *WAIM 2002 LNCS*, vol. 2419, pp. 328-337, Springer Heidelberg.
- [10] K. Bharat, A. Broder, M. Henzinger, P. Kumar, S. Venkatasubramanian, "The connectivity server: fast access to linkage information on the web", In *Proc. of the 7th WWW*, pp. 469-477, 1998.
- [11] K. Randall, R. Stata, R. Wickremesinghe and J. Wiener, "The link database: fast access to graphs of the web", *Research Report 175*, Compaq Systems Research Center, Palo Alto, CA, 2001.
- [12] M. Iliofotou, P. Pappu, M. Faloutsos, M. Mitzenmacher, S. Singh, and G. Varghese, "Network monitoring using traffic dispersion graphs (tdgs)" In *IMC*, pp. 315-320, 2007.
- [13] N. Larsson and A. Moffat, "Off-line dictionary-based compression", *Proceedings of the IEEE*, Vol. 88, No. 11, November 2000.
- [14] P. Boldi and S. Vigna, "The WebGraph framework I: compression techniques", In *Proceedings of the 13th international conference on World Wide Web*, ACM, pp. 595-602, USA, 2004.
- [15] R. Kumar, J. Novak, and A. Tomkins, "Structure and evolution of online social networks" In *KDD*, 2006.
- [16] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, "Trawling the web for emerging cyber-communities" In *WWW*, pp. 1481-1493, 1999.
- [17] R. Rathi, D. J. Cook and L. B. Holder A Serial Partitioning Approach to Scaling Graph-Based Knowledge Discovery. *American Association for Artificial Intelligence*, pp. 188-193. 2002.
- [18] S. Brin and L. Page, "The anatomy of a large search hypertextual web engine" In *WWW*, pp. 107-117, 1998.
- [19] S. Navlakha, R. Rastogi and N. Shrivastava, *Graph Summarization with Bounded Error*. *SIGMOD'08*, June 9-12, 2008, Vancouver, BC, Canada.
- [20] S. Perugini, M. Gon Calves, and E. Fox, "Recommender systems research: A connection-centric survey". *J. Intell. Inf. Syst.*, 23(2):107-143, 2004.
- [21] S. Raghvan and H. Garcia-Molina, "Representing web graphs", In *Proc. of the IEEE International Conference on Data Engineering*, 2003.
- [22] T. Suel and J. Yuan, "Compressing the graph structure of the web", In *Proc. of the IEEE Data Compression Conference*, pp. 213-222, 2001.
- [23] X. Yan and J. Han, "gSpan: Graph Based Substructure Pattern Mining", *Technical Report UIUCDCS-R-2002-2296*, Department of Computer Science, University of Illinois at Urbana-Champaign, 2002.
- [24] Y. Asano, Y. Miyawaki and T. Nishizeki, "Efficient compression of web graphs", In *Proc. 14th Conference on Computing and Combinatorics (COCOON)*, LNCS 5092, pp. 1-11, 2008.

- [25] Y. Asano, T. Ito, H. Imai, M. Toyoda, M. Kitsuregawa, "Compact encoding of the web graph exploiting various power laws: statistical reason behind link database". In: Dong, G., Tang, C.-j., Wang, W.(eds.) WAIM 2003 LNCS, vol. 2762, pp. 37–46, Springer Heidelberg.

AUTHORS PROFILE

P. S. Deshpande received M.Tech in Computer Science from IIT, Bombay (India) and received Ph.D. from Nagpur University (India). He is currently working as a Associate Professor in the Department of Computer Science & Engineering at Visvesvaraya National Institute of Technology, Nagpur (India).

His research interests include Data Bases, Data Mining, and Pattern Recognition.

Meera Dhabu is working as a Assistant Professor in the Department of Computer Science & Engineering at Visvesvaraya National Institute of Technology, Nagpur (India) and pursuing her Ph. D. in Computer Science. Her areas of interests are graph mining and soft-computing.

Siyaram Vishvakarma received M. Tech. in Computer Science from Department of Computer Science & Engineering, Visvesvaraya National Institute of technology, Nagpur (India). His area of interests are graph mining and pattern recognition.

Performance Analysis and Comparison of 6to4 Relay Implementations

Gábor Lencse

Department of Telecommunications
Széchenyi István University
Győr, Hungary

Sándor Répás

Department of Telecommunications
Széchenyi István University
Győr, Hungary

Abstract—the depletion of the public IPv4 address pool may speed up the deployment of IPv6. The coexistence of the two versions of IP requires some transition mechanisms. One of them is 6to4 which provides a solution for the problem of an IPv6 capable device in an IPv4 only environment. From among the several 6to4 relay implementations, the following ones were selected for testing: *sit* under Linux, *stf* under FreeBSD and *stf* under NetBSD. Their stability and performance were investigated in a test network. The increasing measure of the load of the 6to4 relay implementations was set by incrementing the number of the client computers that provided the traffic. The packet loss and the response time of the 6to4 relay as well as the CPU utilization and the memory consumption of the computer running the tested 6to4 relay implementations were measured. The implementations were tested also under very heavy load conditions to see if they are safe to be used in production systems.

Keywords—IPv6 deployment; IPv6 transition solutions; 6to4; performance analysis

I. INTRODUCTION

The majority of the current Internet still uses the Internet Protocol version 4 (IPv4) for forwarding the packets of the communication of the applications. Even though IPv6 was defined in 1998 [1], it has not replaced IPv4 yet. As of Aug. 31, 2013, only 1.88% of the Internet traffic reaching Google used IPv6 [2]. The coexistence of the two versions of IP results in different issues (e.g. the two endpoints of the communication use different IP versions, or the endpoints use the same IP version but the communication path between the endpoints supports the other version only). Several *transition mechanisms* were developed to solve the different issues of the coexistence of the two versions of the Internet Protocol. These *theoretical solutions* are defined in different RFCs. There are a number of *implementations* for each solutions. When a network operator decides to support some of the IPv6 transition mechanisms, it can be a difficult task to choose the right implementations because there can be security, reliability and performance issues. Several papers were published in the topic of performance analysis of different IPv6 transition implementations.

One of the most important driving forces of the deployment of IPv6 is the depletion of the global IPv4 address pool. IANA

delegated the last five “/8” IPv4 address blocks to the five Regional Internet Registries in 2011 [3]. Therefore an important upcoming coexistence issue is the problem of an IPv6 only client and an IPv4 only server, because internet service providers (ISPs) can still supply the relatively small number of new servers with IPv4 addresses from their own pool but the huge number of new clients can get IPv6 addresses only. DNS64 [4] and NAT64 [5] are the best available techniques that make it possible for an IPv6 only client to communicate with an IPv4 only server. Another very important coexistence issue comes from the case when the ISP does not support IPv6 but the clients do and they would like to communicate with IPv6 servers. The most matured solution for this problem is called 6to4 [6]. The stability and performance analysis of different 6to4 implementations is the topic of this paper.

The remainder of this paper is organized as follows: first, a short survey of the results of the most current publications is given, second, 6to4 is introduced, third, the selection of the 6to4 relay implementations is discussed, fourth, our test environment is described, fifth, the performance measurement method of the 6to4 relay implementations is detailed, sixth, the results are presented and discussed, seventh, the validity of our results is considered and our plans for the future research are presented, finally, our conclusions are given.

This topic was identified as being of high importance for network administrators.

II. A SHORT SURVEY OF CURRENT RESEARCH RESULTS

A. The Problem of an IPv6 Only Client and an IPv4 Only Server

Several papers were published in the topic of the performance of DNS64 and NAT64 in 2012 and 2013. The performance of the TAYGA NAT64 implementation (and implicitly of the TOTD DNS64 implementation) is compared to the performance of NAT44 in [7]. The performance of the Ecdysis NAT64 implementation (that has its own DNS64 implementation) is compared to the performance of the authors’ own HTTP ALG in [8]. The performance of the Ecdysis NAT64 implementation is compared to the performance of both the NAT-PT and an HTTP ALG in [9]. All of these papers deal with the common performance of a given DNS64 implementation with a given NAT64 implementation. The performance of the BIND DNS64 implementation and performance of the TAYGA NAT64 implementation are

This research was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0012: “Smarter Transport” – IT for co-operative transport system – The Project is supported by the Hungarian Government and co-financed by the European Social Fund.

analyzed separately and also their stability is tested in [10]. A good survey of the most recent DNS64 and NAT64 research results is given in [11]. They also demonstrated that the DNS64+NAT64 system is a viable solution for an internet service provider. Our results about the stability and performance of different DNS64 and NAT64 implementations were published in [12] and [13], respectively.

B. The Problem of an IPv6 Capable Client in an IPv4 Only Environment

A good survey of IPv6 transition mechanisms including both translation and tunneling solutions can be found in [14]. It discusses 6to4 among the tunneling mechanisms. Ref. [15] named 6to4 and Teredo [16] the two most widely used transition solutions on the bases of the IPv6 prefixes in use. The performance of 6to4 is addressed in [17]. They prepared a controlled environment and compared the performance characteristics (round trip time and throughput) of 6to4 to the native IPv4 and IPv6 using both TCP and UDP between the endpoints. They used Cisco routers in the test network. In contrast, we have chosen different free software [18] (also called open source [19]) implementations for stability testing and performance comparison.

III. INTRODUCTION TO 6TO4 IN A NUTSHELL

The aim of 6to4 is to help those IPv6 capable devices that are residing in an IPv4 environment to connect to other devices being in the same situation and to the native IPv6 internet. The solution is an “automatic tunnel” that encapsulates the IPv6 packets into IPv4 packets (using protocol number 41, as the configured IPv6 over IPv4 tunnel [20]).

The IPv6 capable device can be a single host having a 6to4 pseudo-interface that performs the encapsulation of the IPv6 packets into IPv4 packets and also the decapsulation in the opposite direction. This is called *6to4 host*. It is also possible that there are multiple IPv6 devices in an IPv6 network behind a so-called *6to4 (border) router* that performs the encapsulation of the IPv6 packets into IPv4 packets and the decapsulation in the opposite direction. These *6to4 IPv6 devices* can communicate with other 6to4 IPv6 devices or with IPv6 devices on the native IPv6 internet. In the latter case, they need a 6to4 relay at the border of the IPv4 internet and the IPv6 internet, see Fig. 1.

It is a precondition of the applicability of the 6to4 solution that a 6to4 host or a 6to4 router must have a public IPv4 address. The IPv6 addresses for the IPv6 capable devices will get IPv6 addresses from the 2002::/16 prefix. The next 32 bits of their IPv6 addresses are the 32 bits of the public IPv4 address of the 6to4 host or 6to4 router and still there are 16 bits for subnetting. (It can be filled with 0 or chosen randomly if no subnetting is needed.) The last 64 bits of the IPv6 address may be generated in the usual way from the MAC addresses of the hosts using the modified EUI-64 algorithm.

If the communication occurs between two IPv6 capable devices that both use 6to4 then the route of the encapsulated packet in the IPv4 internet is exactly determined by the public IPv4 addresses of the two 6to4 hosts/routers. If one of the communication endpoints resides in the native IPv6 internet then the route of the packet must go through a 6to4 relay. There

are multiple 6to4 relays having the same 192.88.99.1 anycast address and the network will use the nearest one.

The forthcoming example scenario can be followed in Fig. 1. Let our client having the 2002:c000:208::2 6to4 IPv6 address communicate with the server having the 2001:db8::2 global IPv6 address. The client sends out its IPv6 packet containing its own IPv6 address as the source address and the IPv6 address of the server as the destination address. The packet arrives to the 6to4 router as the default gateway. The 6to4 router encapsulates the IPv6 packet in an IPv4 packet using its own IP address, 192.0.2.8 as the source address and the 192.88.99.1 anycast address as the destination address. The protocol type is set to 41, which indicates that an IPv6 packet was embedded. The packet arrives to the nearest 6to4 relay at the 192.88.99.1 anycast address. The relay recognizes the 41 protocol type and thus it decapsulates the IPv6 packet and sends towards its destination. The server receives the packet and replies in the normal way addressing its reply packet to the client. In the global public IPv6 internet, the 2002::/16 prefix is routed (using anycast addressing) towards the nearest 6to4 relay, which may be different from the one that was used by the packet travelling from the client to the server (asymmetric routing). The 6to4 relay receives the IPv6 packet and encapsulates it in an IPv4 packet. It determines the target IPv4 address from the destination IPv6 address as it contains the 192.0.2.8 IPv4 address of the 6to4 router next to its 2002::/16 prefix: 2002:c000:208::2. When the 6to4 router receives the packet it simply decapsulates the embedded IPv6 packet and sends it to the client.

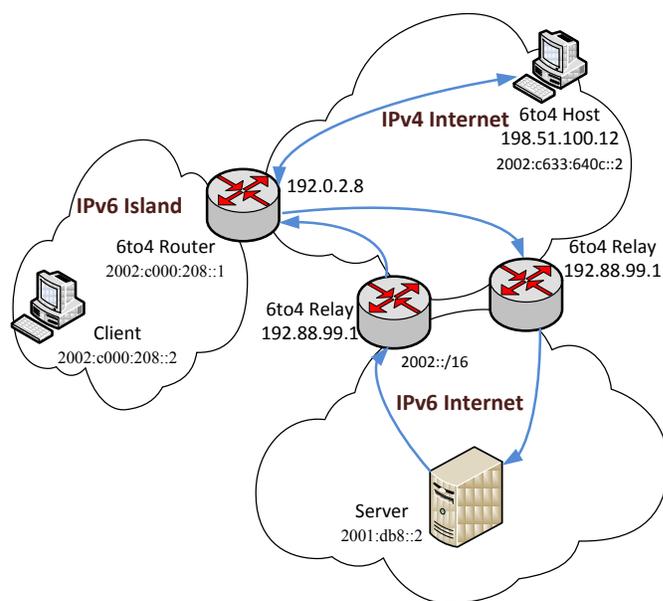


Fig. 1. Sample network for the demonstration of 6to4

Note that if two devices using both 6to4 IPv6 addresses communicate with each other then the 6to4 routers or hosts never send the IPv4 packet (containing the encapsulated IPv6 packet) to a 6to4 relay rather they always take the 32 bits next to the 2002::/16 prefix of the destination IPv6 address and use these 32 bits as the target IPv4 address. A 6to4 target address can be easily recognized from its 2002::/16 prefix. All other

target addresses are treated as global IPv6 addresses and the destination address of the IPv4 packet is set to the 192.88.99.1 anycast address of the 6to4 relays.

Let us consider the performance requirements of the devices performing 6to4 operations.

- A 6to4 host performs encapsulation and decapsulation of the traffic from and to the local host.
- A 6to4 router performs encapsulation and decapsulation of the traffic from and to a limited number of hosts behind the router.
- A 6to4 relay may be responsible for the traffic of a huge number of hosts. Because of the anycast addressing, their number and thus the load of the 6to4 relay depend on the location of other 6to4 relays.

Therefore, if a 6to4 relay using the 192.88.99.1 anycast IPv4 address is set up then it may receive huge load. Consequently, it should be a stable system that does not collapse even in serious overload situation rather complies with the graceful degradation principles [21].

IV. SELECTION OF IMPLEMENTATIONS FOR TESTING

As it was mentioned before, only free software [18] (also called open source [19]) implementations were considered for stability testing and performance comparison. We had multiple reasons for this decision:

- The licenses of certain vendors (e.g. [22] and [23]) do not allow the publication of benchmarking results.
- Free software can be used by anyone for any purposes thus our results can be helpful for anyone.
- Free software is free of charge for us, too.

In our previous research efforts of performance and stability analysis of IPv6 transition solutions (DNS64 [12] and NAT64 [13]), we used Linux, OpenBSD and FreeBSD as host operating systems. Unfortunately 6to4 is not implemented in OpenBSD for security concerns [24]. Thus, the following implementations were selected for testing: sit under Linux, stf under FreeBSD and stf under NetBSD.

V. TEST ENVIRONMENT

The aim of our tests was to examine and compare the performance of the selected 6to4 implementations. We were also interested in their stability and behavior under heavy load conditions. (For testing the software, some hardware had to be used, but our aim was not the performance analysis of any hardware.)

A. The Structure of the Test Network

The topology of the network is shown in Fig. 2. The central element of the network is the 6to4 relay. The ten Dell workstations at the bottom of the figure played the role of the 6to4 hosts for the 6to4 relay performance measurements. The two Dell computers at the top of the figure responded to all the 6to4 hosts.

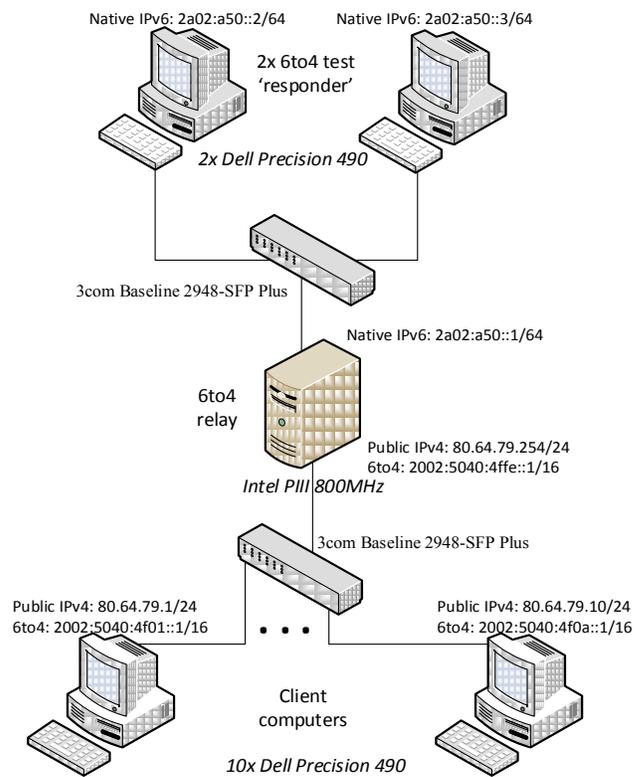


Fig. 2. Topology of the 6to4 test network.

Even though their hardware was much more powerful than that of the 6to4 relay (see details later), we used two of them to be sure that the responder part of the system was never a bottleneck during our tests.

B. The Hardware Configuration of the Computers

A test computer with special configuration was put together for the purposes of the 6to4 relay so that the 6to4 hosts will be able to produce high enough load for overloading it. The CPU and memory parameters were chosen to be as little as possible from our available hardware base in order to be able to create an overload situation with a finite number of 6to4 hosts, and only the network cards were chosen to be fast enough. The configuration of the test computer was:

- Intel D815EE2U motherboard
- 800MHz Intel Pentium III (Coppermine) processor
- 256MB, 133MHz SDRAM
- Two 3Com 3c940 Gigabit Ethernet NICs

For the 6to4 host purposes, standard *DELL Precision Workstation 490* computers were used with the following configuration:

- DELL 0GU083 motherboard with Intel 5000X chipset
- Two Intel Xeon 5140 2.33GHz dual core processors
- 4x1GB 533MHz DDR2 FB-DIMM SDRAM (quad channel)

- Broadcom NetXtreme BCM5752 Gigabit Ethernet controller (PCI Express)

Note that these computers were the same as those used in the DNS64 and NAT64 tests ([12] and [13]) but with a little faster CPU and four identical RAM modules which were able to operate quad channel.

The responder computers were similar to the 6to4 computers but they had somewhat slower CPUs:

- Two Intel Xeon 5130 2GHz dual core processors

Debian Squeeze 6.0.3 GNU/Linux operating system was installed on all the computers including the Pentium III test computer acting as a 6to4 relay when it was used under Linux. The version number of FreeBSD and NetBSD were 9.0 and 6.0.1, respectively. The responder computers had the OpenBSD 5.1 operating system. OpenBSD was chosen because it supports NAT66, which was needed for our experiments.

C. The Software Configuration of the Computers

Fig. 2 shows the IP addresses of the Ethernet interfaces. The 6to4 hosts had IPv4 and 6to4 IPv6 addresses. The 6to4 relay had two Gigabit Ethernet interfaces: eth1 was used for communication with the 6to4 hosts and it had the 80.64.79.254 public IPv4 address and the 2002:5040:4ffe::1 6to4 address; eth2 was used for communication with the responder computers and it had the 2a02:a50::1 global IPv6 address. The

responder computers had the 2a02:a50::2 and 2a02:a50::3 global IPv6 addresses.

To make the results comparable, the same Pentium III computer was used for 6to4 purposes under Linux, FreeBSD and NetBSD. The network settings were also identical. Fig. 3 shows the exact settings of the interfaces under Linux containing the starting of the sit 6to4 pseudo interface and also the routing. Note that some of the tests were also performed with one responder only. Fig. 3 contains the settings for both cases. The network settings and the routing of FreeBSD are shown in Fig. 4 and the FreeBSD stf tunnel was started by the commands shown in Fig 5. The configuration files for the network interfaces under NetBSD are shown in Fig. 6 and Fig 7. The routing under NetBSD was set and the stf tunnel was started by the script in Fig. 8.

As it can be seen in Fig. 3 and Fig. 4 the packets to the 2001:0738:2c01:8001:ffff:0:a00::/104 network were directed to the responder computer(s). On the responder computers, NAT66 was used to redirect these packets to the computer itself (see Fig. 9), to be able to respond to them.

Note that the CPU utilization was monitored during the measurements, and even when using only one responder, its maximum CPU utilization was about 10%.

The network setting of the 6to4 hosts were done as shown in Fig. 10.

```
#to the 6to4 hosts
auto eth1
iface eth1 inet static
address 80.64.79.254
netmask 255.255.255.0

#to the responders
auto eth2
iface eth2 inet6 static
address 2a02:a50::1
netmask 64
pre-up echo 1 > /proc/sys/net/ipv6/conf/all/forwarding
pre-up echo 0 > /proc/sys/net/ipv6/conf/eth2/autoconf
pre-up echo 0 > /proc/sys/net/ipv6/conf/eth2/accept_ra
#In the case of TWO responders:
post-up ip -f inet6 route add 2001:0738:2c01:8001:ffff:0000:0a00::/105 via 2a02:a50::2
post-up ip -f inet6 route add 2001:0738:2c01:8001:ffff:0000:0a80::/105 via 2a02:a50::3
# In the case of ONE responder:
post-up ip -f inet6 route add 2001:0738:2c01:8001:ffff:0000:0a00::/104 via 2a02:a50::2

#6to4
auto sit0
iface sit0 inet6 static
address 2002:5040:4ffe::1
netmask 16
```

Fig. 3. Configuration of the Linux interfaces, routing and the sit tunnel (/etc/network/interfaces).

```
ifconfig_sk0_ipv6="inet6 2a02:a50::1/64"  
ifconfig_sk1="inet 80.64.79.254 netmask 255.255.255.0"  
ipv6_static_routes="responder1 responder2"  
ipv6_route_responder1="2001:0738:2c01:8001:ffff:0000:0a00::/105 2a02:a50::2"  
ipv6_route_responder2="2001:0738:2c01:8001:ffff:0000:0a80::/105 2a02:a50::3"  
ipv6_gateway_enable="YES"
```

Fig. 4. Configuration of the FreeBSD interfaces and routing when using two responders (/etc/rc.conf).

```
ifconfig stf0 create  
ifconfig stf0 inet6 2002:5040:4ffe::1 prefixlen 16 alias
```

Fig. 5. Starting script of the FreeBSD 6to4 tunnel (start6to4).

```
up  
media autoselect  
inet6 2a02:a50::1 prefixlen 64 alias
```

Fig. 6. Configuration of the interface towards the clients, NetBSD (/etc/ifconfig.sk0).

```
up  
media autoselect  
80.64.79.254 netmask 0xfffff000 media autoselect
```

Fig. 7. Configuration of the interface towards the responders, NetBSD (/etc/ifconfig.sk1).

```
route add -inet6 2001:0738:2c01:8001:ffff:0000:0a00:: -prefixlen 105 2a02:a50::2  
route add -inet6 2001:0738:2c01:8001:ffff:0000:0a80:: -prefixlen 105 2a02:a50::3  
ifconfig stf0 create  
ifconfig stf0 inet6 2002:5040:4ffe::1 prefixlen 16 alias
```

Fig. 8. Configuration script under NetBSD (start6to4).

```
set timeout interval 2  
set limit states 400000  
match in on em1 inet6 to 2001:0738:2c01::/48 rdr-to em1
```

Fig. 9. Redirection on the responder computers under OpenBSD (/etc/pf.conf).

```
auto eth0  
iface eth0 inet static  
address 80.64.79.{1..10} # that is 80.64.79.1 on client1, 80.64.79.2 on client2, etc.  
netmask 255.255.255.0  
  
auto tun6to4  
iface tun6to4 inet6 v4tunnel  
address 2002:5040:4f0{1..a}::1 # {} is to be interpreted as above  
netmask 16  
gateway ::80.64.79.254  
endpoint any  
local 80.64.79.{1..10} # {} is to be interpreted as above
```

Fig. 10. Configuration of the network interfaces of the 6to4 hosts under Linux (/etc/network/interfaces).

VI. PERFORMANCE MEASUREMENT METHOD

In order to be able to exactly tune the measure of the load from moderate to serious overload in controlled steps, the number of the clients (6to4 hosts) was tuned. First, a series of measurements was conducted by a single client. Second, the test system was restarted and two clients were used, etc. The measurements were done by the execution of the script in Fig. 11. Different destination IPv6 addresses were used to simulate real-life situation. In a given series of experiments, each active client sent $256*64*8*11=1,441,792$ ICPMv6 echo requests (sending eleven of them to each one of the different

$256*64*8=131,072$ IPv6 addresses). Using the “&” sign for asynchronous command execution, eight ping6 commands were executed quasi-parallel utilizing the computing power of the two dual core CPUs. The target address range, 2001:0738:2c01:8001:fff:0000:0a00::/104 was cut into two halves. The lower half side of the address range, 2001:0738:2c01:8001:ffff:0000:0a00::/105 was used by the commands containing the variable *i*, and the higher half side of the address range, 2001:0738:2c01:8001:fff:0000:0a80::/105 was used by the commands containing the variable *j*. Thus the application of two responders could be easily done by using only two lines in the routing table of the 6to4 relay.

```
#!/bin/bash
i=`cat /etc/hostname|grep -o .$`
j=$((i+128))
for b in {0..255}
do
  rm -r $b
  mkdir $b
  for c in {0..252..4}
  do
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$i.$b.$c
    >$b/6to4p-10-$i-$b-$c &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$j.$b.$c
    >$b/6to4p-10-$j-$b-$c &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$i.$b.$((c+1))
    >$b/6to4p-10-$i-$b-$((c+1)) &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$j.$b.$((c+1))
    >$b/6to4p-10-$j-$b-$((c+1)) &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$i.$b.$((c+2))
    >$b/6to4p-10-$i-$b-$((c+2)) &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$j.$b.$((c+2))
    >$b/6to4p-10-$j-$b-$((c+2)) &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$i.$b.$((c+3))
    >$b/6to4p-10-$i-$b-$((c+3)) &
    ping6 -c11 -i0 -q
    2001:738:2c01:8001:ffff:0000:10.$j.$b.$((c+3))
    >$b/6to4p-10-$j-$b-$((c+3))
  done
done
```

Fig. 11. 6to4 relay performace test script (ping-test-8d.sh).

The CPU utilization and the memory consumption of the 6to4 relay was measured under BSD with the following command:

```
vmstat -w 1 > load.txt
```

Under Linux, the command line was the following:

```
dstat -t -c -m -l -p -n --unix --output load.csv.
```

VII. MEASUREMENT RESULTS

A. Linux sit tunnel with two responders

The results can be found in Table 1. (The tables were put on the same page for the synoptic view and easy comparison of the results of the 6to4 implementations.) This table and all the other tables with the results are to be interpreted as follows. Row 1 shows the number of clients that executed the test script.

(The load of the 6to4 relay was proportional with the number of the clients.) The packet loss ratio is displayed in the second row. Rows 3 and 4 show the average and the standard deviation of the response time (expressed in milliseconds), respectively. The following two rows show the average and the standard deviation of the CPU utilization of the test computer. Row 7 shows the number of forwarded packets per seconds. The last row shows the estimated memory consumption measured at the test computer. Note that this parameter can be measured with high uncertainty, as its value is very low and other processes than the 6to4 relay implementation may also influence the size of free/used memory of the test computer.

The number of forwarded packets per seconds and the CPU utilization are graphically displayed in Fig. 12.

Evaluation of the results:

- Though packet loss occurred for eight or more clients, the packet loss ratio was always very low (under 0.03 percent, which means that 3 packets were lost from 10,000 packets).
- The average response time (given in milliseconds) was low even at 10 clients. Its measure showed nearly linear increase in the function of the load from 4 to 10 clients.
- The number of forwarded packets per second could increase nearly linearly in the function of the number of clients for small number of clients (1-3). It showed less than linear growth for 4-6 clients, and saturation for 7-10 clients, where there was not enough free CPU capacity.
- The CPU utilization showed a linear increase in the function of the number of clients in the case of 1-3 clients (3.0%, 6.1%, 10.1%), as expected. However it showed a radical increase in the case of 4-7 clients (16.9%, 28.1%, 51.0%, 86.4%) which was unexpected and seems to be groundless.
- The memory consumption was always very low (note: it was measured in kB) and it was only slightly increasing in the function of the load with some fluctuations.

To sum up the findings above, we can lay down that the Linux sit 6to4 relay performed quite well, its memory consumption was found to be very low and its average response time increased approximately linearly with the load at high load conditions, that is, it seems to comply with the *graceful degradation* principle [21], however from 4 to 7 clients, the CPU utilization increased higher than linearly in the function of the number of clients. This is a strange phenomenon and it should be investigated before Linux sit 6to4 relay is being actually used in environments with strong response time requirements.

TABLE I. LINUX 6TO4 RELAY PERFORMANCE RESULTS USING TWO RESPONDERS

1	Number of Clients	1	2	3	4	5	6	7	8	9	10	
2	Packet loss (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.03	
3	Response Time of ping6 (ms)	Average	0.274	0.308	0.241	0.480	0.576	0.7	0.867	0.998	1.168	1.311
4		Std. dev.	0.017	0.029	0.053	0.082	0.105	0.126	0.148	0.176	0.192	0.196
5	CPU Utilization (%)	Average	3.0	6.1	10.1	16.9	28.1	51.0	84.6	94.1	92.5	94.3
6		Std. dev.	1.0	1.3	2.5	5.6	10.3	22.2	23.8	13.5	12.8	8.9
7	Traffic Volume (packets/sec)	11177	21360	29424	35166	39829	42615	43502	45411	45691	46812	
8	Memory Consumption (kB)	72	96	120	124	148	124	108	164	212	292	

TABLE II. FREEBSD 6TO4 RELAY PERFORMANCE RESULTS USING TWO RESPONDERS

1	Number of Clients	1	2	3	4	5	6	7	8	9	10	
2	Packet loss (%)	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
3	Response Time of ping6 (ms)	Average	0.491	0.936	1.454	2.050	2.635	3.234	3.866	4.456	5.059	5.629
4		Std. dev.	0.034	0.093	0.143	0.179	0.181	0.139	0.144	0.115	0.156	0.178
5	CPU Utilization (%)	Average	69.0	91.4	97.8	98.7	99.7	100.0	99.9	100.0	100.0	100.0
6		Std. dev.	3.9	2.9	1.4	1.1	0.6	0.0	0.4	0.0	0.0	0.0
7	Traffic Volume (packets/sec)	8900	11916	12873	13018	13179	13248	13176	13212	13214	13288	
8	Memory Consumption (kB)	536	892	1548	892	892	1300	2320	2060	2012	2648	

TABLE III. NETBSD 6TO4 RELAY PERFORMANCE RESULTS USING TWO RESPONDERS

1	Number of Clients	1	2	3	4	5	6	7	8	9	10	
2	Packet loss (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
3	Response Time of ping6 (ms)	Average	2.191	2.270	2.668	2.514	2.543	3.033	2.786	3.320	3.267	3.398
4		Std. dev.	0.574	0.628	0.809	0.718	0.719	0.768	0.686	0.634	0.617	0.557
5	CPU Utilization (%)	Average	9.9	18.9	25.0	34.9	42.9	44.5	55.8	55.0	62.5	67.4
6		Std. dev.	2.0	4.0	6.3	7.9	9.5	10.3	11.6	10.1	11.3	10.7
7	Traffic Volume (packets/sec)	3068	5958	7836	11006	13628	14021	17644	17241	19721	21079	
8	Memory Consumption (kB)	936	144	156	160	160	176	240	228	236	592	

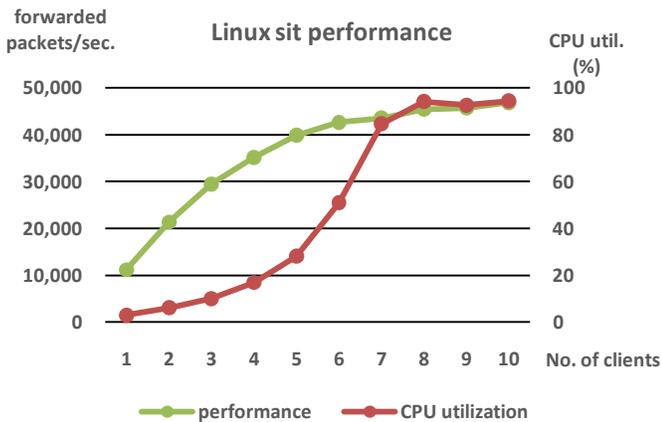


Fig. 12. Linux sit performance and CPU utilization

B. Linux sit tunnel with one responder

The aim of this series of measurements was to see if there are significant differences in the results compared to the case with two responders therefore the measurements were taken only for certain number of clients (1, 4, 5, 8, 10). The comparison can be seen in Fig. 13. The results of the measurements with the one responder were very close to that of the measurements with two responders. For this reason, we do

not include measurement results with one responder for the other two implementations.

C. FreeBSD stf tunnel with two responders

The results can be found in Table 2. The number of forwarded packets per seconds and the CPU utilization are graphically displayed in Fig. 14. Evaluation of the results:

- Though packet loss occurred for 1-3 clients, but the packet loss ratio was always very low (under 0.02 percent, which means that 2 packets were lost from 10,000 packets).
- The average response time (given in milliseconds) was acceptable even at 10 clients. Its measure showed nearly linear increase in the function of the load.
- The number of forwarded packets per second in the function of the number of clients showed saturation from 3 clients because there was not enough free CPU capacity.
- The CPU was practically fully utilized from 3 clients.
- The memory consumption was always low (note: it was measured in kB) and it was only slightly increasing in the function of the load with some fluctuations.

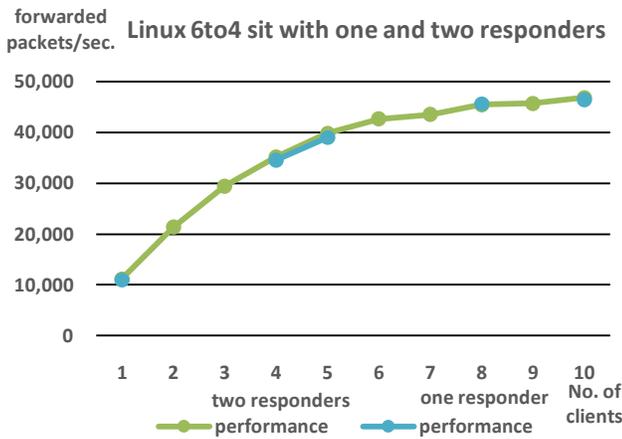


Fig. 13. Comparison of the Linux sit performance with one and two responders.

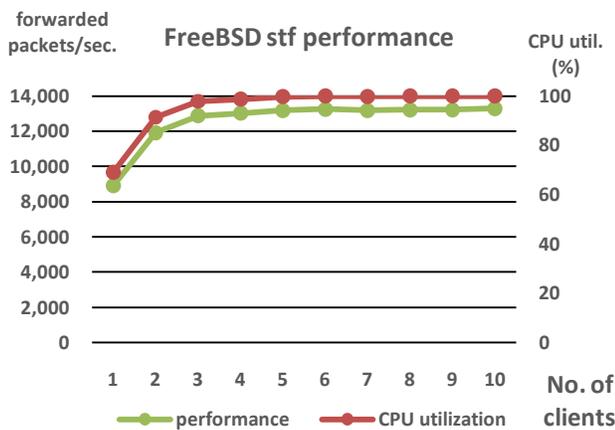


Fig. 14. FreeBSD stf performance and CPU utilization

To sum up the findings above, we can lay down that the FreeBSD stf 6to4 relay performed quite well, its memory consumption was found to be low and its average response time increased approximately linearly with the load, that is, it complies with the *graceful degradation* principle [21].

D. NetBSD stf tunnel with two responders

The results can be found in Table 3. Evaluation of the results:

- The packet loss ratio was always less than 0.005% (thus it was rounded to 0.00%).
- The average response time (given in milliseconds) was acceptable even at 10 clients. Its measure showed only a very slight increase in the function of the load.
- The number of forwarded packets per second in the function of the number of clients did not show saturation even in the case of 10 clients.
- The CPU was not fully utilized even in the case of 10 clients.

- The memory consumption was always low (note: it was measured in kB), but is showed serious fluctuations.

To sum up the findings above, we can lay down that the NetBSD stf 6to4 relay performed quite well, its memory consumption was found to be low and its average response time increased less then linearly with the load. As the CPU was never fully utilized we can state only that as far as we could test it, NetBSD stf complied with the *graceful degradation* principle [21].

E. Comparison of the Results and Final Evaluation

As for the number of forwarded packets per second, Linux sit showed much better performance than FreeBSD/NetBSD stf at any investigated load conditions. As for 10 clients, Linux sit processed $46812/13288=3.52$ times more packets per seconds than FreeBSD stf and $46812/21079=2.22$ times more packets per seconds than NetBSD stf. As for the average response time, Linux sit was also much better than FreeBSD/NetBSD stf. The average response times of Linux sit, FreeBSD stf and NetBSD stf at 10 clients were: 1.3ms, 5.6ms and 3.4ms, respectively. Thus we can say that as for their measured performance, Linux sit seems to be much better than FreeBSD/NetBSD stf. However, Linux sit showed the phenomenon of a super linear CPU consumption in the function of the load at certain conditions. This makes us cautious and we advise to test Linux sit further before using it in mission critical environments with strong response time requirements.

The comparison of the performance results of the FreeBSD stf and of the NetBSD stf is also very interesting. On the one hand, FreeBSD stf performed much better than NetBSD stf both in the number of packets processed (FreeBSD: 8900, NetBSD: 3068) and in the average response time (FreeBSD: 0.5ms, NetBSD: 2.2ms) with one client. Note that its CPU utilization was much higher, too. On the other hand, NetBSD stf performed significantly better than FreeBSD stf both in the number of packets processed (NetBSD: 21079, FreeBSD: 13288) and in the average response time (NetBSD: 3.4, FreeBSD: 5.6) with ten clients. NetBSD had even free CPU capacity. Therefore if one prefers the BSD platform (e.g. for security reasons) FreeBSD can be a good choice if low traffic is expected (and ensured) because of its shorter response time. If high traffic is expected (or if there is a possibility of high traffic), than the more robust NetBSD is our recommendation.

VIII. DISCUSSION OF THE VALIDITY OF THE RESULTS AND FUTURE WORK

As 6to4 operates at network level and IP can carry the data units of both TCP and UDP, no care was taken to the transport layer protocol or the type of applications. This approach can be justified by the fact that 6to4 operates on the basis of the IP header and does not take care to the payload of the datagram. The payload was actually ICMPv6, because it was simple to generate by using the `ping6` Unix command.

On the one hand, an IP datagram is just like another, thus our results characterize the investigated 6to4 relay implementations in general with no regard to the transport layer protocol, applications or even network topology.

On the other hand, the length of the used ICMPv6 echo request / echo reply packets was 100 bytes, which is quite short, thus our tests were a kind of worst case tests for the implementations in the sense that much more payload bytes may be carried by the same amount of work of the 6to4 relay when using longer packets. (Our purpose was to provide the highest possible workload for the tested 6to4 relay implementations to be able to test their behavior under serious overload situations.)

As real life applications use longer packets up to 1500 bytes – the MTU (Maximum Transmission Unit) size of the Ethernet link layer protocol – we plan to perform our tests using different (higher) packet sizes, too.

It is also our plan to check the possible interference between the 6to4 relay implementations and the transport layer protocols. Unlike ICMP and UDP, TCP provides two reliable streams (one per direction) between the endpoints. Therefore TCP resends the lost packets and thus generates higher load for unreliable channels. The packet loss ratio was low in our experiments and thus we do not expect significant differences, but we plan to perform the tests both using UDP and TCP, too.

IX. CONCLUSIONS

The 6to4 IPv6 transition method was introduced. Linux sit, FreeBSD stf and NetBSD stf 6to4 relay implementations were selected for performance and stability testing. The test environment and the measurement method were described.

It was found that Linux sit gave the best performance results in both the number of forwarded packets per second and the average response time under all investigated load conditions. However Linux sit showed super linear CPU consumption in the function of the load under certain conditions. Therefore we advise systems administrators to be cautious and to test Linux sit further before using it in mission critical environments with strong response time requirements.

Within the BSD platform, which can be a choice for security reasons, FreeBSD gave shorter response time at low load conditions and NetBSD could process more packets per second at high load conditions.

REFERENCES

- [1] S. Deering and R. Hinden, "Internet protocol, version 6 (IPv6) specification", IETF, December 1998. (RFC 2460)
- [2] Google, "IPv6 statistics", <http://www.google.com/ipv6/statistics.html>
- [3] The Number Resource Organization, "Free pool of IPv4 address space depleted" <http://www.nro.net/news/ipv4-free-pool-depleted>
- [4] M. Bagnulo, A Sullivan, P. Matthews and I. Beijnum, "DNS64: DNS extensions for network address translation from IPv6 clients to IPv4 servers", IETF, April 2011. ISSN: 2070-1721 (RFC 6147)
- [5] M. Bagnulo, P. Matthews and I. Beijnum, "Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers",

- IETF, April 2011. ISSN: 2070-1721 (RFC 6146)
- [6] B. Carpenter, K. Moore, "Connection of IPv6 domains via IPv4 clouds", IETF, February 2001. (RFC 3056)
- [7] K. J. O. Llanto and W. E. S. Yu, "Performance of NAT64 versus NAT44 in the context of IPv6 migration", in *Proc. International MultiConference of Engineers and Computer Scientists 2012 Vol. I.* (IMECS 2012, March 14-16, 2012), Hong Kong, pp. 638-645
- [8] C. P. Monte et al, "Implementation and evaluation of protocols translating methods for IPv4 to IPv6 transition", *Journal of Computer Science & Technology*, vol. 12, no. 2, pp. 64-70
- [9] S. Yu, B. E. Carpenter, "Measuring IPv4 – IPv6 translation techniques", Technical Report 2012-001, Department of Computer Science, The University of Auckland, January 2012
- [10] G. Lencse and G. Takács, "Performance Analysis of DNS64 and NAT64 Solutions", *Infocommunications Journal*, vol. 4, no 2, June, 2012. pp. 29-36.
- [11] E. Hodzic, S. Mrdovic, "IPv4/IPv6 Transition using DNS64/NAT64: deployment issues", *Proc. 2012 IX International Symposium on Telecommunications* (BIHTEL, Oct. 25-27, 2012), Sarajevo, Bosnia and Herzegovina
- [12] G. Lencse and S. Répás, "Performance analysis and comparison of different DNS64 implementations for Linux, OpenBSD and FreeBSD", *Proc. 27th IEEE International Conference on Advanced Information Networking and Applications* (AINA-2013, March 25-28, 2013) Barcelona, Spain, pp. 877-884.
- [13] G. Lencse and S. Répás, "Performance analysis and comparison of the TAYGA and of the PF NAT64 implementations" *Proc. 36th International Conference on Telecommunications and Signal Processing* (TSP-2013, July 2-4, 2013) Rome, Italy, pp. 71-76.
- [14] P. Wu, Y. Cui, J. Wu, J. Liu, and C. Metz, "Transition from IPv4 to IPv6: a state-of-the-art survey" *Communications Surveys & Tutorials*, IEEE, vol. 15, no. 3, 2013, pp. 1407 - 1424
- [15] E. Karpilovsky, A. Gerber, D. Pei, J. Rexford, and A. Shaikh A, "Quantifying the extent of IPv6 deployment," in *Passive and Active Network Measurement*. vol. 5448/2009: Springer Berlin/Heidelberg, 2009, pp. 13-22.
- [16] C. Huitema, "Teredo: tunneling IPv6 over UDP through network address translations (NATs)", IETF, February 2006. (RFC 4380)
- [17] N. Bahaman, E. Hamid, A.S Prabuwno, "Network performance evaluation of 6to4 tunneling" *Proc. 2012 International Conference on Innovation Management and Technology Research* (ICIMTR, May 21-22, 2012) Malacca, Malaysia, pp. 263-268
- [18] Free Software Foundation, "The free software definition", <http://www.gnu.org/philosophy/free-sw.en.html>
- [19] Open Source Initiative, "The open source definition", <http://opensource.org/docs/osd>
- [20] E. Normark, R. Gillian, "Basic transition mechanisms for IPv6 hosts and routers", IETF, October 2005. (RFC 4213)
- [21] NTIA ITS, "Definition of „graceful degradation“" http://www.its.bldrdoc.gov/fs-1037/dir-017/_2479.htm
- [22] Cisco, "End user licence agreement", http://www.cisco.com/en/US/docs/general/warranty/English/EU1KEN_.html
- [23] Juniper Networks, "End user licence agreement", <http://www.juniper.net/support/eula.html>
- [24] Federico Biancuzzi, "The man in the machine", <http://www.securityfocus.com/columnists/459>

Review on Aspect Oriented Programming

Heba A. Kurdi

Computer Science Department
Imam Muhammad Ibn Saud Islamic University
Riyadh, Saudi Arabia

Abstract—Aspect-oriented programming (AOP) has been introduced as a potential programming approach for the specification of nonfunctional component properties, such as fault-tolerance, logging and exception handling. Such properties are referred to as crosscutting concerns and represent critical issues that conventional programming approaches could not modularize effectively leading to a complex code. This paper discusses AOP concept, the necessity that led to it, how it provides better results in code quality and software development efficiency, followed by stating challenges that developers and researchers face when dealing with this approach. It has been concluded that AOP is promising and deserves more attention from developers and researchers. However, more systematic evaluation studies should be conducted to better understand its implications.

Keywords—*Aspect Oriented Programming; software engineering; AspectJ*

I. INTRODUCTION

This A typical program code is composed of several distinct components. Each of these components is responsible for accomplishing a core function required by the system. Some concerns, though, such as error handling, security and synchronization, are important for the entire system and they therefore crosscut multiple components. Implementing these crosscutting concerns is considered to be a challenging issue that conventional programming approaches, such as Object-Oriented Programming (OOP) and Procedural-Oriented Programming (POP), can not modularize very effectively. Lack of code modularity usually results in a tangled and complex code. As a result, Aspect-Oriented Programming (AOP) has recently emerged as a promising new approach to handle this issue. The term was coined by Gregor Kiczales in 1997 [1] as a complement to the OOP rather than as a replacement to it [2].

From the linguistic meaning of the word “aspect”, a general idea of the technical meaning would arise. AOP is a programming approach that aims to solve crosscutting concerns throughout better modularization of the code. It enhances system features such as modularity, readability and simplicity by better handling of crosscutting concerns [3]. Based on this definition, it is clear that AOP makes a clear distinction between two types of concerns in the software development process:

- Primary concern: represents real world components or objects. In OOP, a class represents each of these components.
- Crosscutting concerns: refers to a programme design feature that is required by multiple software

components. Therefore, its implementation is scattered and/or repeated among them, severely affecting code modularity [4].

For instance, in a banking system, primary concerns include customer and account management, statement generation, transaction tracking ... etc. These concerns are usually implemented as procedures (operations), or classes in conventional programming approaches, i.e. OOP and POP. Examples of crosscutting concerns would include exception handling, authentication and security aspects, which are usually considered essential parts of many procedures or classes in conventional approaches. Therefore, they are handled in multiple locations within the same program, causing a drastic decrease in the quality, readability and modularity of the software [12]. Aspects are treated differently in AOP. They are considered an extended version of the class with additional features [5]. Figure 1 shows the central concepts in each of the three programming approaches and how they are related to each other.

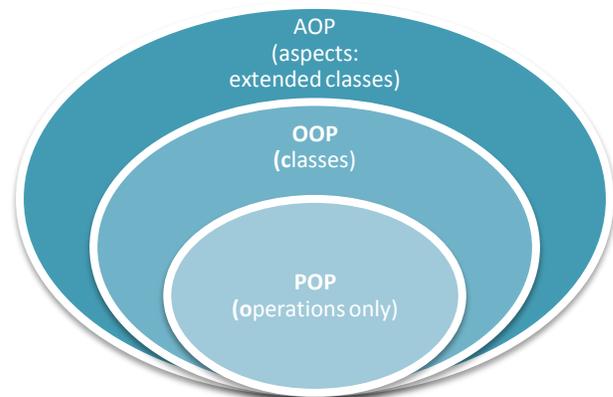


Fig. 1. The relationship between POP, OOP and AOP

Even though an increasing number of programmers and software engineers started adopting the AOP approach, a lot of concerns and challenges are still hindering wider adoption [2]. Therefore, this paper reviews the state-of-the-art in AOP and sheds some light on its related issues, starting with its terminologies and implementation approaches in section 2. The needs that led to the introduction of AOP and its potential benefits are presented in section 3. Section 4 then goes on to provide an overview of previous works that conducted evaluation studies of AOP. In section 5, possible threats and challenges of AOP are discussed and, finally, section 6 provides the conclusion, summarizing the paper and spotting some future research directions.

II. AOP IMPLEMENTATION APPROACHES

Unlike traditional programming approaches AOP provides explicit support for modularizing programs; rather than scattering the code related to a non-functional requirement or a concern throughout a program [19], developers can place it within a separate segment [15]. This required introducing new programming concepts and terminologies such as:

- **Crosscutting concern:** is a purpose that a program wants to achieve. However, this purpose should be scattered among many classes or methods.
- **Aspect:** is a modularized implementation for a crosscutting concern. It amalgamates the distributed code that of a crosscutting concern in one module.
- **Join point:** is a well-defined position in a program, such as throwing an exception or invoking a method.
- **Advice:** is a class of functions that can modify other functions. It is applied at a given join point of a program.
- **Pointcut:** is a general term for a set of joint points whenever reached the corresponding advices will be executed.
- **Weaving:** is the process in which an aspect is added into an object. It can be executed in the compiling time or during the running of the program [6].

There are two approaches for implementing AOP:

A programming language that has been developed specifically for AOP, such as AspectJ: AspectJ [22][23] is the first and most popular tool that AOP developers use for creating software. It is an extension for the Java programming language and uses a Java-like syntax [13]. It is available for download as part of Java software development kit (SDK) that supports it from the official website. All Java programmes are valid in AspectJ, in addition to a special extended version of a class, which is called an aspect [17]. An aspect contains all components of a regular class, as well as some additional entities such as pointcuts and advices [4]. AspectJ needs a special compiler to generate Java byte code. The java class file generated by AspectJ compiler has no difference compared to general Java byte code files [6]. Figure 2 presents an example of AOP in AspectJ.

- Techniques provided by already available programming languages to support aspect implementation:
- Many programming framework have released additions to support ASP[18][20], such as .NET [8] and Spring. Figure 3 illustrates an example of ASP in Spring AOP.

A detailed survey of AOP implementation techniques is provided in [6].

III. AOP ADVANTAGES

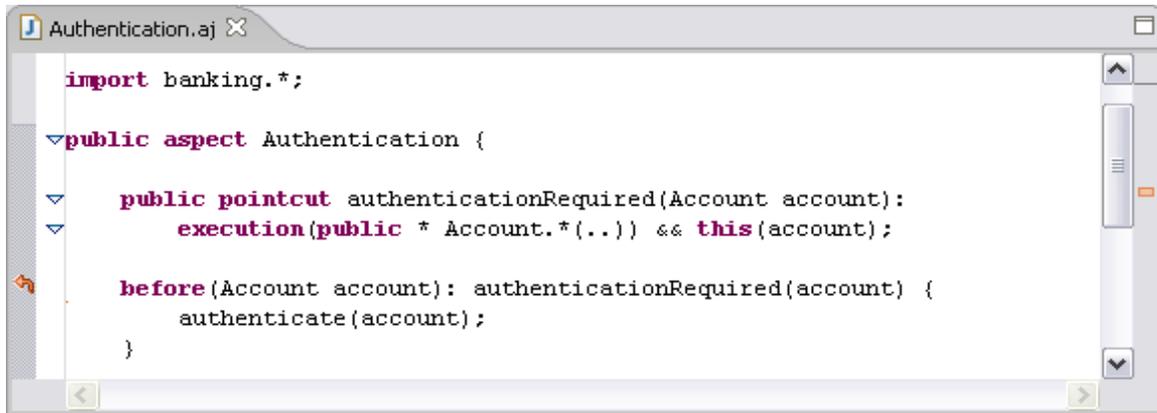
According to Kiczales [1] the OOP and POP have many programming problems that did not allow these approaches to

clearly capture some design elements which are important for software implementation. Therefore, AOP presented itself as a promising approach and as a solution for conventional programming approaches problems. However, solutions provided by AOP do not necessarily come in terms of lower compilation time or less memory usage. Rather, according to Laddad [9], using AOP for implementing software systems will certainly enhance software quality in many ways including:

- **Clear responsibilities for individual modules:** AOP offers better modularisation, by gathering the code that deals with the same aspect in one module avoiding the redundancy of crosscutting concerns. This also leads to a better programming development process because each developer could use his/her expertise with the module he/she knows better.
- **Consistent implementation:** Unlike traditional implementations of crosscutting concerns, which are conspicuous in their inconsistency, AOP provides consistent implementation by having each aspect handled once.
- **Improved reusability:** AOP isolates core concerns from the crosscutting ones, enabling more mixing and matching, and therefore improving the overall reusability in both modules. In contrast, traditional methods do not have this kind of separation between concerns.
- **Improved skill transfer:** The concepts of AOP are reusable and transferable. Therefore, developers training time and cost will be minimised even if they need to learn more than one language. This is because core concerns and design patterns are universal. However, this is not the situation in other frameworks, where developers have to learn from the beginning each time, wasting considerable time and money on training.
- **System-wide policy enforcement:** AOP allows programmers to enforce a variety of contracts and provide guidance in following “best” practices by creating reusable aspects.
- **Logging-fortified quality assurance:** The disability of replicating a bug is one of the major disappointments for traditional methods’ developers, because it is such a ponderous process and thus barely used. On the other hand, AOP enables quality-assurance persons to attach the bug paper with its log, easing the reproduction of the behaviour by the developer.
- **Better simulation of the real world through virtual mock objects:** Software quality testing is enhanced in AOP application by using mock objects. Some scenarios often are not tested because of their complexity that requires an effort to simulate faults such as a network failure. AOP makes the difficult and cumbersome testing process easier without the need to compromise the core design for testability.
- **Nonintrusive what-if analysis:** Dissimilar to non-AOP approaches, AOP does not waste time and space by

checking whether functionality is needed by running what-if analysis every time before changing the system

behaviour.

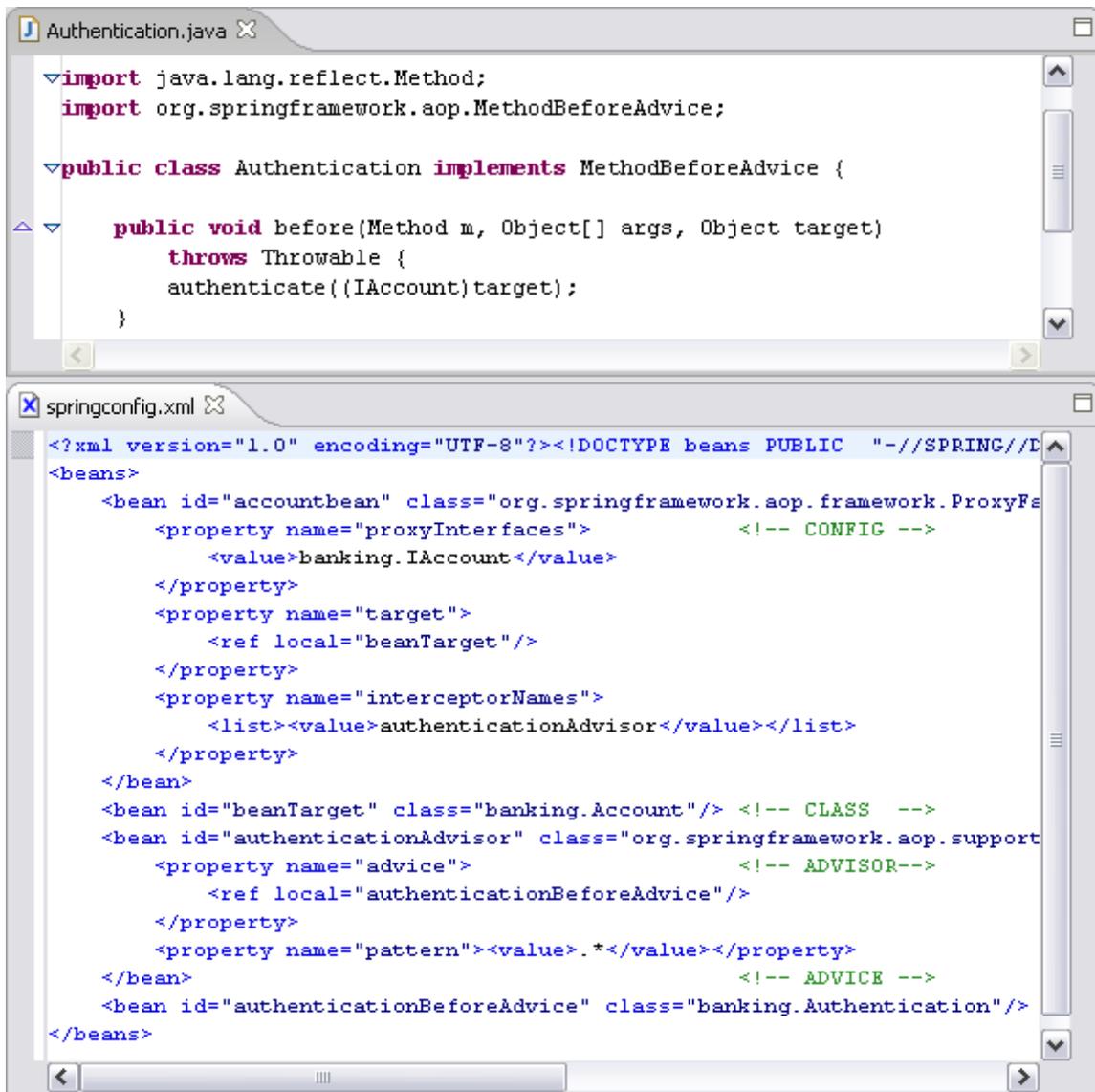


```
Authentication.aj
import banking.*;

public aspect Authentication {
    public pointcut authenticationRequired(Account account):
        execution(public * Account.*(..)) && this(account);

    before(Account account): authenticationRequired(account) {
        authenticate(account);
    }
}
```

Fig. 2. An Aspect for papering unhandled exception in AspectJ [7]



```
Authentication.java
import java.lang.reflect.Method;
import org.springframework.aop.MethodBeforeAdvice;

public class Authentication implements MethodBeforeAdvice {
    public void before(Method m, Object[] args, Object target)
        throws Throwable {
        authenticate((IAccount) target);
    }
}

springconfig.xml
<?xml version="1.0" encoding="UTF-8"?><!DOCTYPE beans PUBLIC "-//SPRING//D
<beans>
  <bean id="accountbean" class="org.springframework.aop.framework.ProxyFe
    <property name="proxyInterfaces"> <!-- CONFIG -->
      <value>banking.IAccount</value>
    </property>
    <property name="target">
      <ref local="beanTarget"/>
    </property>
    <property name="interceptorNames">
      <list><value>authenticationAdvisor</value></list>
    </property>
  </bean>
  <bean id="beanTarget" class="banking.Account"/> <!-- CLASS -->
  <bean id="authenticationAdvisor" class="org.springframework.aop.support
    <property name="advice"> <!-- ADVISOR-->
      <ref local="authenticationBeforeAdvice"/>
    </property>
    <property name="pattern"><value>.*</value></property>
  </bean> <!-- ADVISOR -->
  <bean id="authenticationBeforeAdvice" class="banking.Authentication"/>
</beans>
```

Fig. 3. An Aspect for papering unhandled exception in Spring AOP [7].

IV. EVALUATION APPROACHES

Due to the potential benefits of AOP in software engineering and the tremendous advantages claimed by its supports, many studies have emerged to systematically evaluate the AOP approach and compare it to conventional programming approaches.

Ali et al. [10] have made a systemic review of comparative evidence of aspect-oriented programming. They discussed, in detail, the benefits and limitations of AOP based on the following criteria: performance, code size, modularity, evolvability, cognition and language mechanism. Each criterion was studied and was concluded with one of four possible results:

- Positive – when they note enhancement of the criterion with AOP compared to non-AOP implementations.
- Negative – when the implications of introducing aspects are not advantageous in the context.
- Insignificant – when AOP solution does not produce better results than earlier solutions, or there is no noteworthy evidence of enhancement.
- Mixed – when the study concludes with a combination of above three statement types and does not deliver any aggregated statement about the effect that AOP had on the studied characteristic.

The outcomes after evaluation each criteria are as follows:

- Performance: The results were Mixed results having AOP generating positive outcomes in regards to execution performance by improved response time and minimising the usage of both memory and hardware costs. However, the results were Insignificant when AOP was tested in Unix OS to evaluate runtime cost. The result of using AOP for optimising a network simulator was the same. This outcome made some researchers question if AOP can influence the performance.
- Code size: From the beginning, the founder of AOP, Kiczales [1], promised that his approach would create a tangible reduction in the size of code because of the separation of crosscutting concerns as mentioned in earlier sections. According to the research finding in this matter, there was a notable reduction in code size by approximately 40%, which means that there was a reduction in the line of code (LoC) as well. In addition, there was a reduction in certain types of codes such as exception handling. However, in some particular cases, AOP did not remarkable affect the LoC numbers. This led to the conclusion that AOP is actually effective in minimising the code size positively most of the time. If not, it will be more or less the same as non-AOP approaches.
- Modularity: Modularity results were positive, especially in Separation of Concerns (SoC). However, there was a lack of evidence in some studies, which suggests the need of more research in this area.

- Evolvability: Evolvability means AOP's ability to adapt to the continuous change in the user requirements and operational environment. Results were positive for this matter.
- Cognition: The cognitive outcomes were measured through looking at the development time and understandability, which is the degree to which developers/evaluators understand a system or component. Obtained results were insignificant so three studies were reviewed but results are not encouraging
- Language mechanism: The way that AOP deals with the code is certainly different from traditional approaches. Exception handling was taken as an example and compared in both OOP and AOP approaches. Results found were positive.

To evaluate the effectiveness of AOP in separation of concerns, Tsang et al. [14] applied a code quality metrics suite, developed in [18] to compare between real systems developed based on AOP and OOP in terms of system properties. They used the amount of reduction in coupling and cohesion values of the CK metrics as performance measures. The results showed better modularity of AOP systems over OOP systems,

Madeyski and Szala [4] have also made an empirical study of the impact of AOP on software development efficiency and design quality. Although their study has an obvious weakness, which is the small sample size (three programmers, only one of which is using AOP while the other two used OOP), it does give some research background for future studies. They asked the programmers to develop a web-based application for manuscript submission and reviewing. The goals of the study include:

- Evaluating the AOP impact on code quality.
- Evaluating the AOP impact on software development efficiency.

The researchers concluded their study by stating that the impact of AOP in software development efficiency was not confirmed. This is firstly because of the disability of applying statistical tests to analyze it due to the limited number of participants, as mentioned earlier. Secondly, it is because the statistical tests that they could execute for internal metrics showed insignificant results. That was also the case for the AOP impact on code quality: according to the researcher, the only positive impact in code quality was modularity.

Recently, Boticki et al. [2] investigated the educational benefits of introducing AOP paradigm into programming courses for undergraduates software engineering students. The study discusses how using the AOP paradigm, affects students' programs, their exam results, and their overall perception of the theoretically claimed benefits of AOP. The research methodology consisted of analyzing of students' programs, administering surveys, and collecting exam results. The results showed that the use of AOP as a supplement to object-oriented programming enhances the productivity of the students and leads to increased understanding of theoretical concepts.

V. CHALLENGES

So far, AOP has not gained wide adoptions. In addition to the possible reason related to it still being in infancy stage, some other disadvantages and challenges associated with it were highlighted in the following studies.

According to Laddad [9] there are two common oppositions to AOP, the first being that it makes the debugging process much harder. The second opposition is the fact that crosscutting modules implementation requires understanding the core module implementation details and vice versa. This is not the case in the OOP approach, though, where understanding is only required of the exposed abstraction between two classes. Moreover, Luca and Depsi [11] have discussed the challenges that AOP faces as a new programming approach in the following points:

- Lack of expertise: The community members of AOP are approximately only 2000 programmers worldwide, and only 10-15% of them are experienced enough to use AOP in an OOP environment.
- Concerns: Although AOP came to provide and to deliver a better separation of concerns (SoC), in reality, when a system reaches a certain degree of complexity, such separation is very hard to achieve, if not impossible.
- Standardisation: AOP introduced new dimensions and standards to programming. This, in general, creates complexity and possible resistance, but it was also the case when the OOP was introduced after the POP, which indicates that this is a normal scenario.

VI. CONCLUSION

AOP is a programming approach that aims to solve crosscutting concerns by offering better modularization of the code. This paper provided a brief overview of the state-of-the-art in AOP, starting with its definitions and example usages. It then went on to highlight the needs that led to the introduction of AOP. These can be summarized as the desperate demand for improved software quality. After that, an overview of previous works that have conducted evaluation studies of AOP were presented. The studies discussed the benefits and limitations of AOP based on performance, code size, modularity, evolvability, cognition, language mechanism and efficiency.

However, obtained results could not prove or disprove the effectiveness of AOP, except in two measures: language mechanisms and code size. AOP showed positive outcomes in these two measures. Possible threats and challenges associated with AOP were also discussed. They included making the debugging process harder and requiring more understanding of the core module and crosscutting concerns implementation. All these issues were not presented in conventional programming approaches.

All of the referenced research had a common conclusion, declaring the need of further in-depth studies and more research of AOP and its impact, which shows that this approach is still relatively new and unpopular. However, the developers who used this approach feel very confident and they

talk assertively about its enrichment to software quality. The empirical studies, though, had another thing to say, and it was not always in favor of AOP.

To conclude, it has been found that AOP is a very interesting topic that needs to take its righteous place in the programming community. Only then could researchers study AOP effectively and efficiently.

REFERENCES

- [1] G. Kiczales, J. Lamping, A. Mendhekar, C. Maeda, C.Videira Lopes, J.-M. Loingtier, J. Irwin, "Aspect Oriented Programming", In Proc. Europ. Conf. on Object-Oriented Prog.(ECOOP), Finland, Springer Verlag LNCS 1241, June 1997.
- [2] I. Boticki, M. Katic, S. Martin, "Exploring the Educational Benefits of Introducing Aspect-Oriented Programming Into a Programming Course," , IEEE Transactions on Education, vol.56, no.2, pp.217-226, May 2013.
- [3] T. Zukai, P. Zhiyong, "Survey of Aspect-Oriented Programming Language , Journal of Frontiers of Computer Science and Technology, 2010, vol.4, no.1, pp 1-19.
- [4] L. Madeyski, L. Szala, "Impact of aspect-oriented programming on software development efficiency and design quality: an empirical study," IET Software, 2007 , vol. 1, no.5, pp. 180-187.
- [5] J. Viegas, J. Vuas, "Can aspect-oriented programming lead to more reliable software?," IEEE Software, 2000, vol.17, no.6, pp. 19-21.
- [6] D. Zhengyan, "Aspect Oriented Programming Technology and the Strategy of Its Implementation," In Proceedings of International Conference on Intelligence Science and Information Engineering (ISIE), 2011, pp.457,460, 20-21.
- [7] M. Kersten, AOP@Work: AOP tools comparison, Part 1, accessed [26/8/2013] [online] available: <http://www.ibm.com/developerworks/library/j-aopwork1/>
- [8] H. Bing, C. Jiaying G. Jianye, "An Approach to Implement AOP Framework Under .NET Platform, Journal of Computer and Modernization," 2009, vol.11.
- [9] R. Laddad, "Aspect-oriented programming will improve quality," IEEE Software, 2003, vol.20, no. 6, pp. 90-91.
- [10] M. Ali, M. Babar, L. Chen, K. Stol, "A systematic review of comparative evidence of aspect-oriented programming," Information and Software Technology, 2010, vol.52, no.9, pp. 871-887.
- [11] L. Luca, I. Despi, "Aspect Oriented Programming Challenges," Anale Seria Informatica, 2005.vol. 2, no. 1, pp. 65-70.
- [12] R. D. Dechow (2005), "Advanced Separation of Concerns and the Compatibility of Aspect- Orientation. accessed [26/8/2013] [online]
- [13] [http://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/29757/Dec howDouglasR2005.pdf?sequence=1](http://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/29757/Dec%20DouglasR2005.pdf?sequence=1)
- [14] G.Kiczales, E. Hilsdale, J. Hugunin, M. Kersten, J. Palm,W. G. Griswold, "Getting Started with AspectJ", CACM, 2001, vol. 44, no. 10.
- [15] S.L. Tsang, S. Clarke, E.L.A. Baniassad, "An evaluation of aspect-oriented programming for Java-based real-time systems development," In Proceedings of the International Symposium of Object-Oriented Real-Time Distributed Computing ISORC, 2004, Vienna, Austria, pp. 291-300.
- [16] K. Lieberherr, D. Orleans, J. Ovlinger, "Aspect-oriented programming with adaptive methods," ACM Communication, 2001, vol. 44, no. 10, pp 39-41.
- [17] T. Xie, J. Zhao, "A framework and tool supports for generating test inputs of AspectJ programs," In Proceedings of AOSD, 2006, pp. 190-201,
- [18] A. Colyer , A. Clement , G. Harley, M. Webster, Eclipse AspectJ: Aspect-Oriented Programming With AspectJ and the Eclipse AspectJ Development Tools. Addison-Wesley pp.504, 2004.
- [19] S.R. Chidamber, C.F. Kemerer, "A metrics suite for object oriented design," IEEE Transaction on Software Engineering, 1994, vol. 20, no. 6, pp. 476-493.

- [20] E. M. Novikov, "An approach to implementation of aspect-oriented programming for C," *Programming and Computer Software*, 2013, vol.39, no. 4, pp 194-206.
- [21] E. M. Novikov, "One approach to aspect-oriented programming implementation for the C programming language, Proc. of the 5th Spring/Summer Young Researchers' Colloquium on Software Engineering, Yekaterinburg, 2011, pp. 74-81.
- [22] G. Kiezales, E. Hilsdale, J. Hugunin, M. Kersten, J. Palm, W.G. Griswold, "An overview of AspectJ," In *Proceedings of the 15th European Conference on Object-Oriented Programming (ECOOP'01)*, 2001, pp. 327-353.
- [23] AspectJ: an aspect-oriented extension to Java. <http://www.eclipse.org/aspectj/doc/released/progguide/language.html>
- [24] Introduction to AspectJ. <http://eclipse.org/aspectj/doc/released/progguide/starting-aspectj.html>

Fingerprint Image Segmentation Using Haar Wavelet and Self Organizing Map

Sri Suwarno

Information Technology Faculty,
Duta Wacana Christian University, Yogyakarta
Indonesia

Subanar¹, Agus Harjoko², Sri Hartati³

Computer Science Post Graduate Program, Faculty of
Mathematics and Natural Sciences, Gadjahmada
University, Yogyakarta, Indonesia

Abstract—Fingerprint image segmentation is one of the important preprocessing steps in Automatic Fingerprint Identification Systems (AFIS). Segmentation separates image background from image foreground, removing unnecessary information from the image. This paper proposes a new fingerprint segmentation method using Haar wavelet and Kohonen's Self Organizing Map (SOM). Fingerprint image was decomposed using 2D Haar wavelet in two levels. To generate features vectors, the decomposed image was divided into nonoverlapping blocks of 2x2 pixels and converted into four elements vectors. These vectors were then fed into SOM network that grouped them into foreground and background clusters. Finally, blocks in the background area were removed based on indexes of blocks in the background cluster. From the research that has been carried out, we conclude that the proposed method is effective to segment background from fingerprint images.

Keywords—Fingerprint Segmentation; AFIS; background image; foreground image; Haar wavelet; SOM

I. INTRODUCTION

Automatic Fingerprint Identification Systems (AFIS) have been widely used in modern offices, as well as in police departments. The success of AFIS highly depends on the quality of fingerprint images that were inputted to the systems. High quality inputs might be recognized or verified more accurately than low quality inputs. By high quality inputs, we mean fingerprint images that contain a lot of useful information. Unfortunately, fingerprint images are not only composed of useful information but also noises and unnecessary information such as background. In this case, background exists when fingerprint is captured using digital devices that have wider sensor surface than finger area [1] or when latent or rolled fingerprints are scanned using digital devices. To exclude background from computation process, it is necessary to remove background from the image. The process of removing background from image may be defined as segmentation. Zhang et al [2] stated that the goal of segmentation is to discard the background and to reduce the number of false features. This means that segmentation might improve AFIS performance.

Researchers used a number of methods and a variety of features to build fingerprint segmentation algorithms. Segmentation methods might be classified in some ways, and so did feature generation. Based on feature source, segmentation methods can be grouped into block-wise based and pixel-wise based. In block-wise based method, features are

extracted from blocks, while in pixel-based method features are extracted from pixels. Commonly, the features extracted from the blocks for this purpose are the values of coherence, mean, variance and field direction. Block-wise segmentation method is fast, but it creates blocking effect in the segmented image edges [3]. To avoid blocking effect, some researchers such as [4] chose pixel-wise method. Theoretically, the method produced finer segmented image, but this method is sensitive to noise and needs longer computation. In [3][5] block-wise and pixel-wise methods are combined to overcome the weaknesses of both methods.

Other researchers proposed classification of segmentation methods from different perspective. Some papers, such as [3][6][7] categorized segmentation methods into supervised and unsupervised, based on the way the decisions are taken. In unsupervised methods, features are extracted from blocks and classification of background and foreground is decided based on a threshold chosen empirically [3]. By choosing an empirical threshold, blocks can be classified as background or foreground. In supervised methods, the decision is made using simple classifier based on training samples [7].

Although some features have been chosen by researchers, finding simple and discriminative features is still a challenging problem. This paper proposed a new segmentation method based on block-wise features generated by Haar wavelet decomposition and using Kohonen's Self-Organizing Maps to classify the blocks into background or foreground. Our method of segmentation would be expected to segment fingerprint images adaptively without user intervention. Although the method produces blocking effect, but the segmented image still reserve a lot of important information that can be used in the next processes, such as classification or recognition.

The rest of the paper is organized as follows. Section 2 discusses feature extraction, Section 3 describes segmentation method, Section 4, presents experiment procedure and some experiment results, Section 5 presents the conclusion of the paper, and finally the future work is presented in Section 6.

II. FEATURES EXTRACTION

Theodoridis and Koutroumbas [8] stated that features should be available prior to the design of classifier. Furthermore, they considered that the goal of feature extraction is to generate a vector that uniquely identifies a single object. The size of feature vector should be smaller than the size of the data from which the vector is generated, so that processing of

feature vectors would be faster than processing the original data itself. Therefore, it is advantageous to select methods that can generate feature vectors as small as possible without losing important information.

There are some choices to generate feature vectors based on the nature of the fingerprint image. Logically, we need to examine what properties of background that differ from foreground. At least there are three properties of background and foreground that can be extracted to form features, namely intensity, homogeneity and pattern. Background intensity is usually brighter than foreground intensity. It means that pixel values in the background area are higher than in the foreground area. Related to homogeneity, background area is more homogenous than foreground so that its variant is smaller than foreground variant. Patterns of background and foreground are more difficult to be measured numerically. Some measurements have been proposed to define patterns in fingerprint image, such as orientation or direction of ridges, the number of ridges and the thickness of ridges. These properties have been used extensively, but they are sensitive to noise and need long computation. To overcome the drawbacks we utilized feature generator that indirectly detects intensity, homogeneity and pattern as well. The generator that we chose is Haar wavelet decomposition. We used 2D Haar wavelet decomposition in two levels that decomposed original image into approximation and detail coefficients. Theoretically, all of these coefficients are resulted from linear transformation from the same data. So if we selected only one coefficient, it could reduce computation complexity without degrading the performance. In this method we chose the elements of approximation coefficient as vector feature. This feature consists only four elements. Sometimes the intensity of background pixels are close to furrows' pixels intensity. It means that if the block size is too small, the furrows will be classified as background either. This problem might be solved by considering the size of the furrows. We observed that in 512 dpi fingerprint images, the furrows size are around 6 to 9 pixels. Therefore we chose blocks of size 8x8 pixels by considering that when this kind of blocks reside in foreground area, they always contain part of furrows, so that those blocks will be classified as foreground.

III. SEGMENTATION METHOD

As mentioned in previous discussion, segmentation may be seen as a classification problem. It is why in [8] and [9] the writers stated that when features have been selected, any classifier can be used to segment background from foreground. Indeed, every classifier has its own advantages and disadvantages. In this research, segmentation process utilized a simple SOM network with only four input nodes and four clusters. Fig. 1 is the diagram of the method that we proposed. It can be seen in Fig 1 that fingerprint image is firstly decomposed using 2D Haar wavelet in two levels. The process produces four coefficients, i.e. approximation, horizontal, vertical and diagonal coefficients. Among these four coefficients, only approximation coefficient is chosen to generate feature vectors.

After this decomposition, the size of the data image is a quarter of the original size. As a result, the size of data blocks in the decomposed image is also a quarter of data blocks size of the original image. Based on this calculation, decomposed image was divided into nonoverlapping blocks of 2x2 pixels. These blocks are then converted into vectors of 4 elements. All of feature vectors are then fed into SOM network to be clustered.

In this method, the SOM network composed of four input neurons and four clusters. Input of the network are feature vectors resulted from the previous process. In this research, only three alternatives of cluster number, namely two, three and four, were chosen. If the number of cluster is two, theoretically background and foreground blocks would be separated into different clusters. To the contrary, if the cluster number is three or four, there will be more than two alternatives where blocks are clustered. In this research, the number of cluster was tested to find the best performance.

Theoretically, there is no mechanism to determine where a block is grouped. To estimate where background blocks were grouped, we used two considerations,

(1) the number of background blocks is far fewer than the number of foreground blocks, and (2) pixel intensity values of background are higher than pixels intensity of foreground. In SOM training algorithm, higher value inputs tend to be the winners from the first epoch. By considering these two hypothesis, we may predict that background blocks might be cumulated in a cluster that contains fewer number of blocks. Furthermore, background blocks might also be cumulated in cluster with small numbers, namely cluster number 1 or number 2. By using the indexes of blocks in background cluster, the background blocks could be removed from the original fingerprint images.

IV. EXPERIMENT AND RESULT

The proposed method was tested using NIST-4 database [10]. In this research, we conducted two experiments by examining the number of epochs and the number of clusters in the SOM training. For the number of epochs, we tested five values, namely 100, 125, 150, 175, and 200 to find the optimum epoch value.

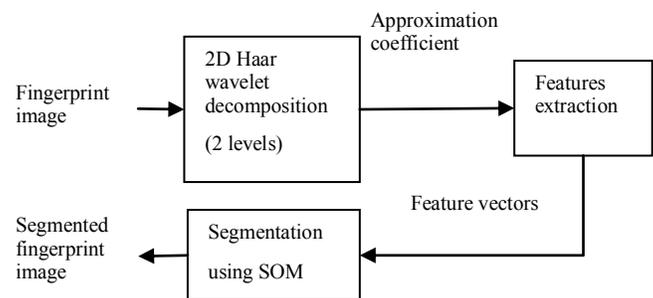


Fig. 1. Diagram of Proposed Method

TABLE 1. NUMBER OF EPOCH VS NUMBER OF CLUSTER MEMBER

File Name	Epochs = 100		Epochs = 125		Epochs = 150		Epochs = 175		Epochs = 200	
	Number of cluster members									
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 1	Cluster 2
F0001_01.tif	1396	2700	710	3386	710	3386	712	3384	711	3385
F0002_05.tif	1491	2605	668	3428	674	3422	670	3426	668	3428
F0003_10.tif	2104	1992	2111	1985	2111	1985	2111	1985	2111	1985
F0004_05.tif	2037	2059	2031	2065	2031	2065	2031	2065	2031	2065
F0005_03.tif	1673	2423	705	3391	704	3392	702	3394	694	3402
F0006_09.tif	1459	2637	1151	2945	1151	2945	1150	2946	1150	2946
F0007_09.tif	1391	2705	1216	2880	1217	2879	1216	2880	1216	2880
F0008_10.tif	2319	1777	2374	1722	2372	1724	2373	1723	2374	1722
F0009_08.tif	1293	2803	805	3291	806	3290	801	3295	800	3296
F0010_10.tif	1212	2884	591	3505	586	3510	592	3504	587	3509

Furthermore, for cluster numbers we tested three values, namely two, three and four, using the epoch value resulted in the previous experiment.

A. Finding the optimum epoch

It is desirable that the number of epoch should be as small as possible; however, the SOM training should reach stable condition. To find the optimum number of epoch we tested 100 fingerprint images from NIST-4 database. We used two, three and four clusters as dependent variables to find the optimum epoch numbers. By considering the clarity of visualization, we only presented 10 of them, as seen in Table 1 and Fig. 2. In this experiment, we did not need to examine cluster 2 because it should be foreground cluster. To save the space, we did not present the table results and related figures for experiment with three dan four clusters. The results were almost similar as Table 1 and Fig. 2.

It can be seen in Table 1 that the number of blocks that were cumulated in cluster-1s tend to stable starting from epochs 125 upwards, and Fig. 2 clarify this finding.

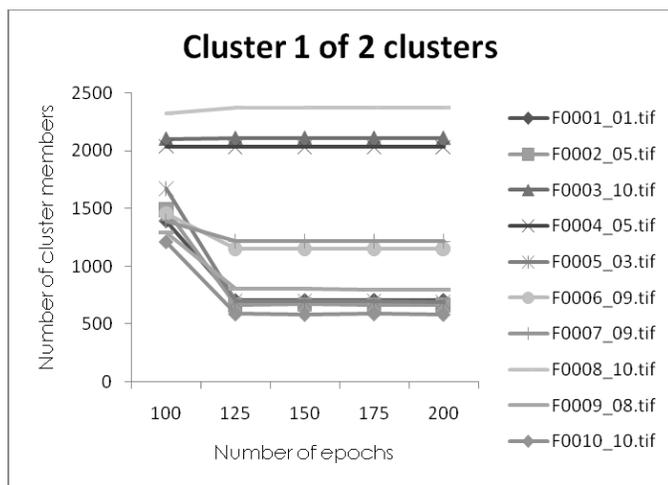


Fig. 2. Number of cluster vs number of epoch for two clusters

From these experiments, it could be concluded that for this method, 125 is the smallest epoch number in order to get converged result.

B. Finding the best cluster number

After the optimum epoch number was found, namely 125, then we used this number to find the best cluster numbers. As mentioned in previous explanation, we used two, three and four clusters. It is difficult to measure the best result numerically, so we examined them visually. Fig. 3 and Fig. 4 are the examples of the results.

From these two figures it might be concluded that when the cluster number is two, some blocks inside in the foreground area are considered as background, and showed as white blocks. From our experiments of 100 fingerprint images, we concluded that the best result was when the number of cluster was four.

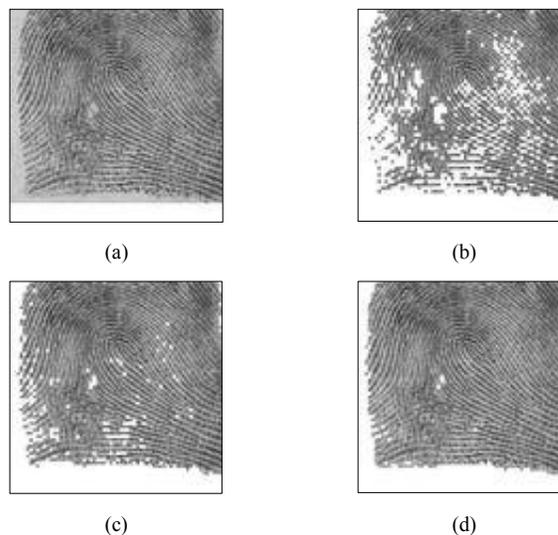


Fig. 3. Result examples (a) original image (b) cluster number = 2 (c) cluster number = 3 (d) cluster number = 4

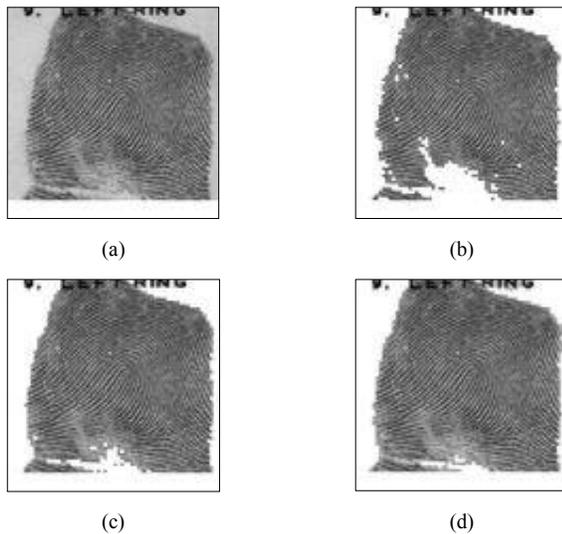


Fig. 4. Fingerprint with annotation. (a) original image, (b) cluster number = 2, (c) cluster number = 3, (d) cluster number = 4

However, due to the white 32-bit-strip in the bottom of fingerprint images of NIST-4 database, our method using four clusters could not separate background from foreground. In this case, background cluster moved from cluster number 1 to cluster number 2. Therefore, the algorithm should be modified slightly.

V. CONCLUSIONS

We have proposed a new method for fingerprint image segmentation based on block-wise features. We used 100 fingerprint images of NIST-4 database to test the method. The method used Haar wavelet decomposition to generate feature vectors. The vectors were then clustered by Self Organizing Map (SOM) to classify background blocks and foreground blocks. Using block indexes in the background cluster, the background blocks in the original images were removed. The result of our experiments showed that the proposed method is simpler and more adaptive compared to other methods. The

method does not need any preprocessing such as enhancement, and it does not need user intervention such as to select threshold either. Based on our experiments, it can be concluded that epoch number of 125 upwards would give stable condition of SOM training. Moreover, the optimum number of cluster would be 4.

VI. FUTURE WORK

Our future work will continue to implement the method in fingerprint processing applications such as orientation field estimation and classification.

REFERENCES

- [1] H. Jia and P. Tang, "An Improved Fingerprint Image Segmentation Algorithm," in International Conference on Computer Science and Electronics Engineering, 2012, pp. 118-120.
- [2] J. Zhang, R. Lai, and C.C.J. Kuo, "Latent Fingerprint Segmentation with Adaptive Total Variation Model," in IAPR International Conference on Biometrics, 2012, pp. 189-195.
- [3] S. Zhang, X. Jing, B. Zhang, and S. Sun, "An Adaptive Fingerprint Image Segmentation Algorithm Based on Multiple Features," in Second International Conference on Advanced Computer Control, 2010, pp. 191-194.
- [4] E. Ruiz-Echartea, M. Morales-Sandoval, M., A. Nuno-Maganda, and Y. Hernandez-Mier, "A Novel Strategy for Image Segmentation of Latent Fingerprint," in 22nd International Conference on Electrical Communications and Computers, 2012, pp. 196-201.
- [5] C. Chen, D. Zhang, L. Zhang, and Y. Zhao, "Segmentation of Fingerprint Image by Using Polarimetric Feature," in International Conference on Autonomous and Intelligent Systems, 2010, pp. 1-4.
- [6] M.U Akram, A. Ayaz, and J. Intiaz, "Morphological and Gradient Based Fingerprint Image Segmentation," in International Conference on Information and Communication Technologies, 2011, pp. 1-4.
- [7] J. Ma, X. Jing, Y. Zhang, S. Sun, and H. Huang, "Simple Effective Fingerprint Segmentation Algorithm for Low Quality Images," in IEEE International Conference on Broadband Network and Multimedia Technology, 2010, pp. 855-859.
- [8] S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th ed. London, UK: Academic Press, 2009.
- [9] D. Maltoni, D. Maio, A. K Jain, and S. Prabhakar, Handbook of Fingerprint Recognition. London: Springer-Verlag, 2009.
- [10] C.I Watson and C.L. Watson, "NIST Special Database 4 Fingerprint Database," National Institute of Standards and Technology Advanced Systems Division, Image Recognition Group, 1992.

A Concept-to-Product Knowledge Management Framework: Towards a Cloud-based Enterprise 2.0 Environment at a Multinational Corporation in Penang

Yu-N Cheah

School of Computer Sciences
Universiti Sains Malaysia
11800 USM Penang, Penang, Malaysia

Soo Beng Khoh

A Multinational Corporation
Penang, Malaysia

Abstract—Knowledge management initiatives of a multinational corporation in Penang are currently deployed via its enterprise-wide portal and Intranet. To improve knowledge management initiatives from its current strength, efforts could now be focused on synergizing organizational workflow as well as computing resources and repositories. This paper proposes a concept-to-product knowledge management framework to be deployed in a cloud-based environment. It aims to provide effective support for collaborative knowledge management efforts at all stages of projects. The multi-layered framework is built upon the organizational memory which drives relevant processors and applications. The framework manifests itself in the form of a cloud-based concept-to-product dashboard from which employees can access applications and tools that facilitate their day-to-day tasks in a seamless manner.

Keywords—knowledge management framework; organizational memory; Enterprise 2.0; workflow management; cloud computing; concept to product

I. INTRODUCTION

Efforts to manage organizational knowledge have come a long way. From the days when the focus was on document management and policies for knowledge sharing, current knowledge management trends are shifting towards providing search and exploration facilities, and semantic capabilities. Web 2.0 is also a prominent feature with the concept of Enterprise 2.0, i.e. an effort that places importance on the social dimension of the Web. With effective knowledge management also comes effective visualization and mobility [1][2].

A multinational corporation (MNC) (the name of the corporation is withheld for confidentiality purposes) based in Penang (henceforth called MNC Penang) was set up in 1976 as one of the MNC's design centers worldwide. Operations in MNC Penang mainly concerns the design and manufacturing of the MNC's line of portable and mobile radio communication devices. In MNC Penang, knowledge management efforts have been initiated and made operable through its enterprise-wide portal and Intranet, particularly for lessons learned.

Although MNC Penang values knowledge management initiatives, the overall effort could still be improved from its current strength. These improvements primarily involve

knowledge management initiatives in synergized organizational workflow that span across various functional groups, and their associated computing resources and repositories. Therefore, this paper proposes an integrated concept-to-product knowledge management framework with the following key features:

- Knowledge management at all project stages: Knowledge management support, in terms of applications or tools, is provided from conceptual design until the product is shipped.
- Effective support for collaborative efforts: Work practices are captured in Web 2.0 fashion that allows collaboration and interactivity between project members.
- Computing resources and repositories are centralized: All computing resources and repositories are centrally located in a cloud-based environment to minimize maintenance cost and maximize accessibility and efficiency.

The framework ultimately manifests itself in the form of a cloud-based concept-to-product dashboard from which employees can access applications and tools that facilitate their day-to-day tasks at any project stage in a seamless manner.

II. RELATED WORK

There are many on-going efforts to realize knowledge clouds. One such effort to realize knowledge in a cloud [3] involves incorporating support for semantically rich data or knowledge, collaboration, and self-organization. In practical terms, these are achieved by providing seamless access to data (cloud storage), self-organization of data, distributed query processing, as well as ensuring data quality and security.

Work on the Enterprise Knowledge Cloud [4] views that the clouds would interconnect business partners and suppliers to customer and consumers. The clouds are potentially driven by social network, blogging, and wiki-type applications and these would be grouped into private (corporate), partner, and public clouds. The combination of these three clouds forms the Enterprise Knowledge Cloud.

Besides connectivity, cloud-based architectures have also been explored for specific purposes such as knowledge discovery [5]. For this purpose, three-layer architecture was developed. A service selection layer provides a variety of (knowledge discovery) services to be assembled into a workflow at the service workflow composition layer. Finally, the workflow is executed at the application execution layer on the cloud.

From a domain-specific perspective, a large-scale knowledge cloud called caBIG [6] was proposed for integrated cancer-related biomedical research, translational medicine, and personalized health management. It allows for various research initiatives from different institutions to connect to a grid-based network. This allows for the exchange of resources including those from federal establishments.

In general, cloud computing appears to be the way forward, building on the strengths of service computing, as well as grid and distributed computing. The challenge here would be the migration of legacy systems over to the cloud environment, as well as the reuse of existing data, applications, and services.

An overview of the layers is as follows:

- **Object:** The object layer consists of MNC Penang’s organizational memory. This contains all formalized organizational data, information, and knowledge. The organizational memory is controlled or described by structural and domain taxonomies or ontologies. These control the structure of data, information, and knowledge that is stored in the organizational memory, as well as standardize the vocabulary that is used in the content. MNC Penang’s organizational memory will be further described in Section III.
- **Processors:** Processors are basically processes and techniques that work upon the object layer content, i.e. the organizational memory, and taxonomies or ontologies. One or more processors are used to implement certain applications. The processors include those for data mining, machine learning, computational social and cultural dynamics, as well as for knowledge and language processing.
- **Applications:** The series of applications aim to assist

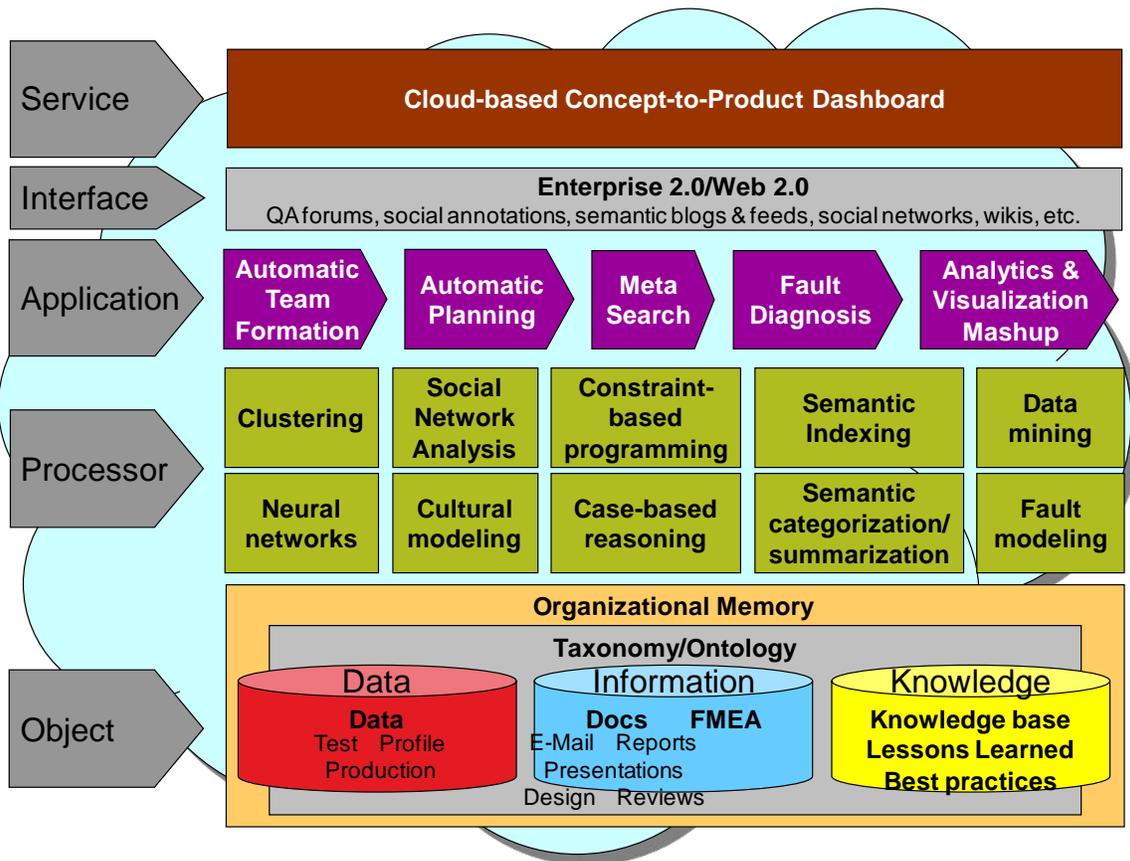


Fig. 1. Concept-to-product knowledge management framework

III. CONCEPT-TO-PRODUCT KNOWLEDGE MANAGEMENT FRAMEWORK

The concept-to-product knowledge management (C2P-KM) framework is a multi-tiered framework. Fig. 1 illustrates the framework.

MNC Penang’s employees throughout the course of a particular project, i.e. from a product’s conceptual design to shipment. For example, at the commencement of the project, the proposed application would be automatic team formation, while at the shipment of the product and eventual conclusion of the project, the

analytics and visualization mashup would facilitate reporting and evaluation. Details of the applications would be provided in Section IV.

- Interface: The interface layer defines the interaction of the users with the available applications.
- Service: This defines the service that is available to the users, i.e. a cloud-based concept-to-product (C2P) dashboard. The dashboard is implemented in a cloud computing environment, allowing seamless access to applications and organizational memory.

IV. MNC PENANG'S ORGANIZATIONAL MEMORY

The wide range of activities carried out at MNC Penang results in the generation of a variety of data, information, knowledge, and wisdom [7] that is distributed across many

physical locations. Collectively, all these can be viewed as MNC Penang's organization memory (see Fig. 2).

MNC Penang's data includes test data from research and development activities. It would also consist of manufacturing-related data, i.e. materials, quality issues, costs, etc. Administrative data would also be included, i.e. employee profile and records, facilities management, etc.

At the information level, data would be analyzed and interpreted in an effort to ascertain the causes of certain test results. Subsequently, documents would be generated. Therefore, failure modes and effects analysis (FMEA) are generated at this level, and so are documents, e.g. reports, presentations, etc. All these provide information on the reason behind the data as well as to provide details on day-to-day activities.

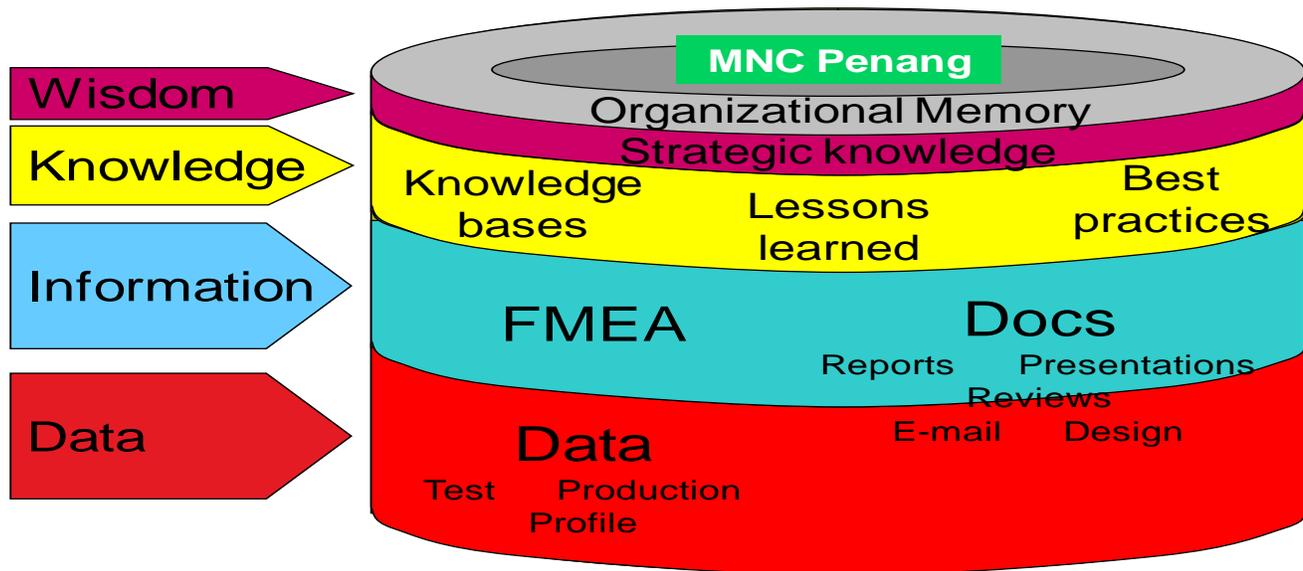


Fig. 2. MNC Penang's organizational memory

The knowledge level serves to form an abstraction or generalization of the information, allowing actions to be taken as a result. Information from the FMEA and documents could be used to build knowledge bases, compile lessons learned, or even establish best practices for the organization. While the information contained in FMEAs and documents serve as a basis for the knowledge level, the propagation and reuse of knowledge level items stand to benefit the organization the most. On the long-term, effective management of knowledge would help reduce the effects of brain-drain and employee turn-over.

Ultimately, at the level of wisdom, this would consist of strategic knowledge drawn from the experience of MNC Penang's senior employees. This would be the most unstructured level in the organizational memory, and includes tacit knowledge stored in the minds of the experts in the organization. Strategic knowledge would be drawn upon during strategic planning, decision making, and other activities to gain competitive advantage.

As an illustration, mechanical engineers measure a series of part dimensions (e.g. O-ring for water sealing in volume knobs)

in the course of their work. These measurements are data. Over a period of time, trends in these measurements are analyzed. Compliance of the O-ring measurements to specifications is also noted (with those stored in manuals and reports). These processing of measurements and results of compliance are at the level of information. The trend in O-ring measurements and its compliance serve as triggers for knowledge-use. Deviation in O-ring measurements could have been due to parts that are out of specifications from a supplier.

In this case, the knowledge bases could be referred to in order to determine the best cause of action, e.g. to halt production, or to use alternatives. Finally, at the level of wisdom, senior management could decide on the severity of the non-compliance of the O-rings, and whether or not to continue procurement from a particular supplier on the long-term.

V. KNOWLEDGE ENGINEERING AT MNC PENANG

The organizational memory is a rich source of data, information, and knowledge objects to be processed or "engineered", i.e. acquired, represented, and reasoned/processed. Fig. 3 illustrates how knowledge

engineering-related processes work on the organizational memory.

In terms of purpose, knowledge engineering processes, especially acquisition, and representation, allow organizational memory items to be brought in, thus building the object layer of the framework shown in Fig. 2. These processes make the organizational memory items operable, adding value to their existing state. Additionally, the processes of reasoning and processing allow the organizational memory to be made operable in the form of applications, thus building the application layer of the framework.

In the context of MNC Penang, some of the processes, or processors (processing structures) and technologies that are relevant for organization-wide usage include clustering, neural networks, social network analysis, cultural modeling, constraints-based programming, case-based reasoning, semantic indexing, semantic categorization/summarization, data mining, and fault modeling (see processor layer in Fig. 1). The utilization of these processors in various applications would be further described in the Section VI.

VI. FROM CONCEPT TO PRODUCT: FLOW OF APPLICATIONS AND SUPPORTING PROCESSORS

The organization is “alive” with activities, executing projects that have been defined. To facilitate the execution of the projects, a list of applications is proposed to assist MNC Penang’s employees along the way (see application layer in Fig. 1). It should also be noted that in order to carry out the activities, relevant inputs are needed in order to produce the necessary outputs. In this regard, the organizational memory at the object layer serves to fuel the activities of the organization while at the same time acts as a repository for outputs of the activities.

A. Automatic Team Formation

Projects are executed by teams, and the success of a project depends on the performance of the team. Therefore, ideally,

team members should be carefully chosen to ensure they are available, have the right skill set, work well together, and are motivated.

The current practice in many organizations, and MNC Penang is no exception, is to manually assign employees to a particular project. This is largely due to the differing project complexities and the engineers’ skill-sets matching. However, if given a larger pool of employees and projects, automatic team formation would be an advantage.

Automatic team formation, inspired by coalition formation [8] and computational cultural dynamics [9] efforts, allow teams to be formed based on the profiles of employees stored in the organizational memory. Automatic team formation, having previously been carried out on learning environments [10] allows specifications of teams to be defined by managers. The employee profiles are then matched to fulfill the needs of the specifications while ensuring competency and cohesiveness. For this purpose, clustering and neural network processors would be relevant.

B. Automatic Planning

Planning is an essential step in the initial phases of a project as it facilitates looking ahead in an effort to ensure products are delivered on time. Within an organization, the core business activities are pretty much repetitive in nature in view that newer products and services are often modifications of existing ones.

Similarly in MNC Penang, where the nature of its business is in the design and manufacture of radio communication devices, projects do not differ very much in terms of the activities that need to be carried out. In this MNC, phase-gates for product development are set at specific intervals and relevant activities are carried out to meet the respective deadlines.

In view of this, planning can be automated where plans for previous projects could be used as a reference in planning for

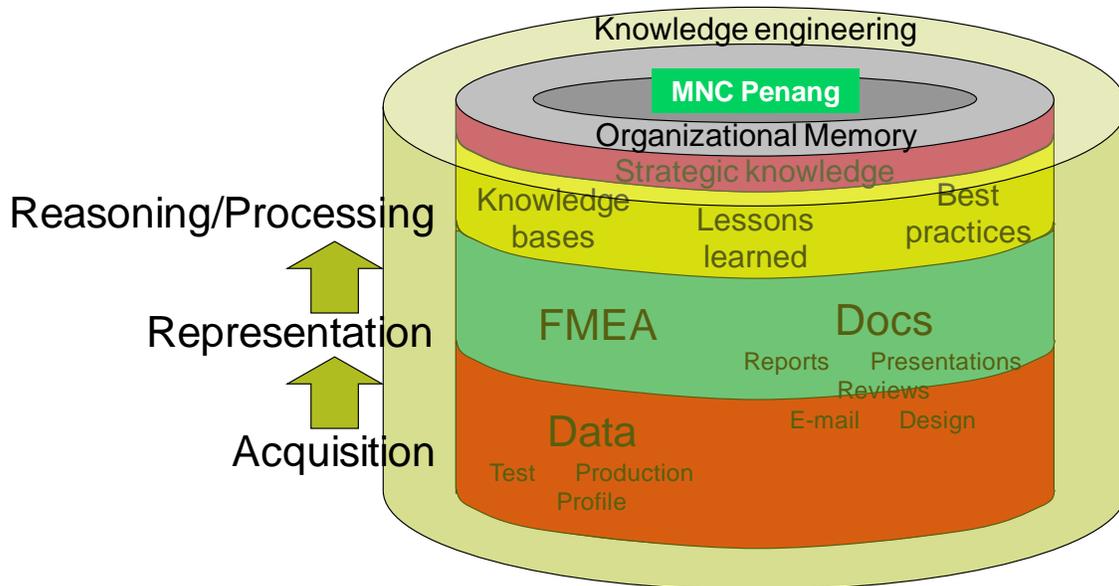


Fig. 3. MNC Penang’s knowledge engineering processes

new projects [11]. Case- and constraint-based approaches may be suitable for such tasks. Additionally, re-planning may be necessary should serious problems arise, causing delays; or there may be instances when tasks may complete ahead of schedule. In such situations, dynamic planning strategies may be employed [12].

C. Meta Search

During the course of a project, many references are made to existing documentation and reports, manuals, correspondences, etc. In large organizations with multiple simultaneous projects, it is a challenge to search for the right resources in time as the collection or repository of documents, i.e. the organizational memory, is very large. An added challenge is that the required resources may be in different formats, and located in different locations. These locations could be on different servers, or they may even be in the personal possession of different employees.

Employees of MNC Penang, being part of a multinational corporation, have access to an even larger organizational memory. The current search facility employed in the MNC's web portal (based on OpenText's Livelink [13]) is powerful in the sense that it searches through a wide range of the MNC's online resources. The current search experience could be enhanced in the areas of search results, improving search speeds, presenting updated or most recent version of requested documents, as well as predicting and recommending relevant adjacent search results to the user.

MNC Penang's employees wanted a search facility with the following capabilities:

- with features similar to Google's
- able to search all databases
- a one-stop search portal
- results that are properly classified
- presents results intuitively

In order to improve the search facility, proper semantic indexing as well as semantic categorization and summarization could be carried out. It must be stressed that these are done "semantically" in order to understand the meaning and intention of the user, and the context in which the search is made.

Therefore, the way forward would be to incorporate semantic web technologies in the organizational memory itself [14][15][16]. This would lead to a search experience that:

- is personalized, e.g. Hunch (www.hunch.com)
- is incorporated in the documents that are being viewed, e.g. Zemanta (www.zemanta.com)
- incorporates natural language processing, e.g. Hakia (www.hakia.com)
- incorporates intuitive visualization to search results, e.g. Quintura (www.quintura.com)
- is entity-based, e.g. Freebase (www.freebase.com)

D. Fault Diagnosis

As much as the organization's reputation rests on the reliability of the products offered, customers also consider the quality of customer support when making purchase decisions. Therefore, after the delivery of products, organizations would be concerned about providing the best after-sales service to their customers.

In the context of MNC Penang, this is a critical business area in view that the MNC's radio communication devices are often deployed in situations where down-time has to be minimal, e.g. incident scene and security management, high-volume logistics and fleet management, event management, etc. Therefore, there is a need for an expert-like system that could diagnose faults in a timely manner in both back-office or on-site scenarios. This could be achieved by using a hybrid of case base [17], knowledge base, and fault modeling approaches [18]. Also current are the use of fuzzy logic, neural networks, and neuro-fuzzy approaches [19]. Through these, it is hoped that "no stones are left unturned" in coming up with effective solutions.

E. Analytics and Visualisation Mashup

For strategic decision making purposes, data that has been collected during the course of the project needs to be analyzed and presented in a manner that would assist decision makers in their task. The challenge here is that there is a variety of data that is located at various locations that need to be scoured through.

MNC Penang's design and manufacturing activities undergo various milestones. Data and reports are generated at each stage by different departments. At the end of the project, these need to be consolidated in an effort to measure the outcome of the project.

A mashup that analyses data, and produces intuitive visual representation of the results would benefit decision makers. As opposed to traditional portals that provide various tools and functions separately, mashups aim to aggregate multiple tools and functions to present a seamless value-added output [20]. Ultimately, mashups could serve as a source of data mining for business intelligence [21].

VII. COLLABORATION AND AGILITY VIA ENTERPRISE 2.0

Web 2.0 technologies are now explored in a good number of organizations. It can be considered one of the main driving forces behind the concept of Enterprise 2.0. Enterprise 2.0 covers the interface layer of the C2P-KM framework (see Fig. 1) and it addresses the user experience of the C2P dashboard. Enterprise 2.0 technologies are centered on the following concepts [22]:

- Search: Users must be able to find information they are looking for. Many employees are finding Intranet navigation/site maps to be less useful than keyword-based searches. Expert finders are also popular.
- Links: The usefulness of a document or item of information can be better measured by the number of links or references pointing to it. In return, links help

users arrive at a desired piece of information more quickly.

- **Authoring:** Writing should be made less daunting. Tools should be available to facilitate writing both individually and collaboratively, e.g. blogs and wikis.
- **Tags:** These help to make sense of the organizational memory and help users keep track of content. The tags form a folksonomy, i.e. a categorization system built by users, for users.
- **Extensions:** These make use of business intelligence and data mining within the organizational memory for the benefit of users. These include attempts to present results or information that are a step ahead of the user, e.g. search results that may be of interest to the user besides those actually requested.
- **Signals:** There needs to be a balance between push and pull of information, and hence, the right “signals” need to be given to users to prevent information overload. Current feed mechanisms, e.g. RSS and aggregators, are efforts in this direction.

While traditional desktop applications such as word processors, spreadsheets, and e-mail would continue to be popular tools, the level of flexibility and collaboration that can be achieved is limited. With social web technologies prevalent in Enterprise 2.0 frameworks, organizations could work towards being more open, transparent, agile, collaborative, and “social”-driven [23].

One can imagine the use of blogs instead of traditional word processing applications to document personal ideas and reports. Not only would the blog entry be available to a wider audience, it would also allow feedback and comments for further action if required. Communication via social networking applications instead of e-mail would simplify group discussions and facilitate the identification of experts. Another example would be the co-authoring of documents using wikis instead of sending documents back-and-forth via e-mail between the authors.

While all these seem attractive, the challenge in adopting Enterprise 2.0 would be to make the transition from “Enterprise 1.0” as seamless as possible. Traditional tools could be integrated as much as possible into the organization’s Enterprise 2.0 framework.

VIII. CLOUD-BASED CONCEPT-TO-PRODUCT DASHBOARD

In view that the MNC’s organizational memory is distributed across many physical locations, current client-server-based architectures could pose certain challenges in terms of maintenance and cost. In making computing services leaner and more affordable, cloud computing is perceived as the solution, i.e. by moving computing hardware and software to a centralized facility [24]. The location of the cloud or the computing resources is transparent to the user, and the users would use applications in the form of Software as a Service (SaaS) [25].

In the context of MNC Penang, certain online applications are hosted by the MNC headquarters (as a global organization)

at a particular location. Issues to consider in such a setup include server load and bandwidth. With a cloud-based implementation of the MNC’s computing hardware and online applications, sharing of computing power and applications would be made easier and achieved at a lower cost.

Another advantage of cloud computing is that the organizational memory would be centralized. When data and information are centralized, data processing and mining could be achieved more inclusively, thus allowing business intelligence to be better achieved in the cloud [26], and also facilitating knowledge management [27][28]. Therefore, MNC Penang’s C2P dashboard could be made operable in the cloud as enterprise-wide services [29]. Not only would this facilitate infrastructural implementation, it would also encourage better software engineering practices through the reuse of existing services to form new application pipelines or process flows.

IX. AN EXAMPLE: FAULT DIAGNOSIS

Fig. 4 provides an overview of the processes of knowledge engineering for fault diagnosis and its relationship with the C2P dashboard and organizational memory.

Senior customer support engineers are required to share their knowledge via the C2P dashboard. This would be in a structured or semi-structured manner. Their knowledge would be used to produce various fault models. Besides knowledge from human experts, data from the organizational memory would be mined in order to discover fault and diagnostic patterns, and potentially new diagnostic knowledge. The fault models and mined knowledge would then be formalized, represented, and stored in the knowledge bases in the organizational memory.

When customer support engineers are out in the field supporting customers, or when they are at their home base trying to reproduce and resolve faults reported by customers, the fault diagnosis application would be utilized. The engineers would input observable and measurable signs of the complaints via the C2P dashboard and the application would then consult the knowledge bases for a likely solution to the problem.

While the fault diagnosis application would be the main application used by the engineers, they would still have the other applications (e.g. the meta-search, and analytics and visualization mashup) and supporting Web 2.0 tools in order to receive additional information and knowledge, as well as to collaborate, and receive feedback and support from their colleagues.

X. DISCUSSION

The organizational memory would serve as the fuel for the C2P dashboard, minimizing the impact of the knowledge gaps of individual employees. While the organizational memory could be organized according to the data-information-knowledge-wisdom continuum, it is in actuality quite chaotic.

Therefore, the importance of the ontology cannot be discounted in such a cloud-based implementation. The ontology must successfully enforce standards in terms of document structure and vocabulary to facilitate effective knowledge sharing and reuse.

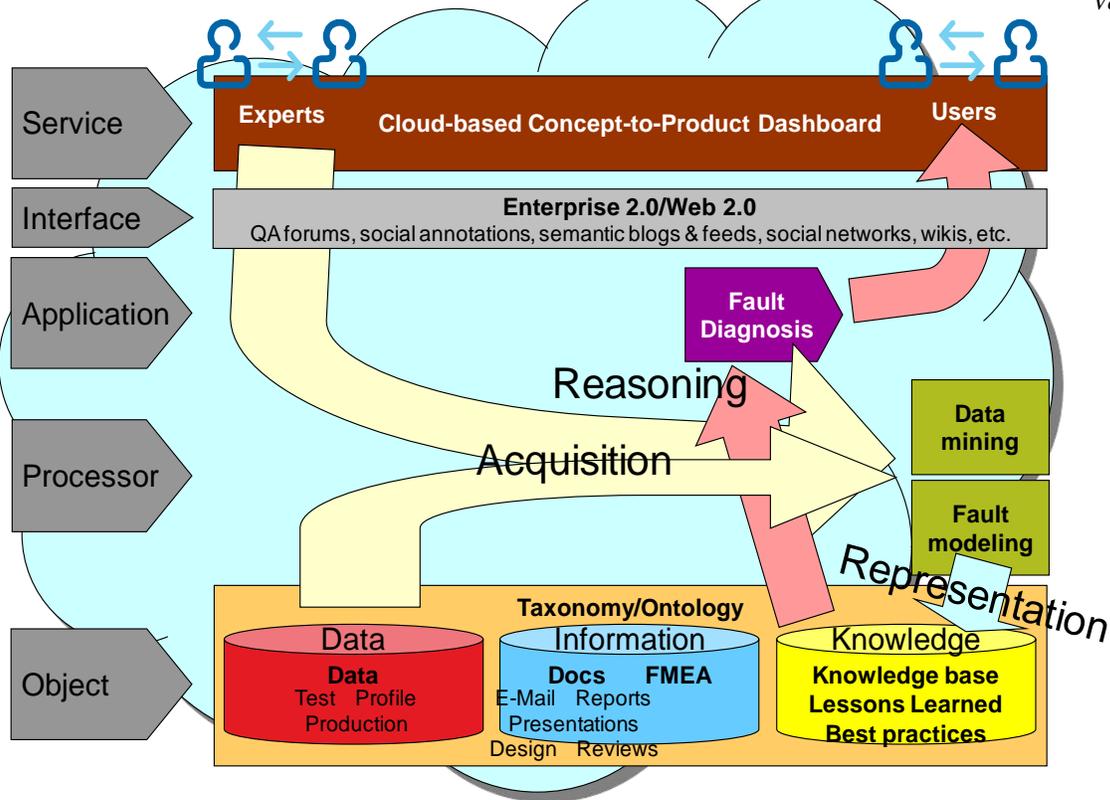


Fig. 4. Fault diagnosis knowledge engineering overview

The C2P dashboard’s flow of applications described in Section VI serves as a starting point for more applications or project activities to be added. Applications such as project monitoring and logging could be inserted in the flow in the future. The existing applications such as automatic planning could be deployed in stages, focusing on planning for design and manufacturing, and perhaps extended later to include planning for human resource, etc. The analytics and visualization mashup could be implemented in a similar way, focusing first on mashups for engineers and later include those for senior management.

While Web 2.0 seems to be the main driving force behind the concept of Enterprise 2.0, the C2P-KM framework is already pointing towards Web 3.0 with the need for applications to be semantically aware. While it is important for the C2P-KM framework to capitalize on Web 2.0 technologies to encourage collaboration within the organization, Web 3.0 is expected to deliver results in a more semantically accurate manner. In this way, manual effort in fact-finding from multiple sources could be reduced by deploying semantically aware searches. These searches understand the intent and context of the query and therefore could proceed automatically. It is anticipated that Enterprise 3.0 is not far on the horizon.

The proposed C2P-KM framework for MNC Penang has far reaching implications in terms of content, and applications or services. The cloud-based deployment allows flexibility in creating content which in turn fuels the services. The maturity of cloud computing would further drive the Enterprise 2.0 idea.

Nevertheless, the success of a cloud-based deployment would depend on the need for computing power. Similarly, the reduction in computing cost would only be justified with the

increase in demand for organization-wide computing power. Hence, MNC Penang needs to fully deploy computing intensive business intelligence applications to justify the C2P-KM framework.

XI. CONCLUSION

Many organizations strive to be lean and agile in today’s very competitive business environment. In order to achieve this, many existing policies and procedures need to be simplified and streamlined, i.e. lean, and new policies need to be able to adjust to changes and evolving needs, i.e. agile. These needs can potentially be translated into powerful but straightforward applications that employees could use as part of their day-to-day work.

In this paper, the C2P-KM framework was proposed and described as an attempt to map out MNC Penang’s existing operational and knowledge needs against future application and architectural demands. In this effort, it was realized that an integrated and seamless way to manage the organizational memory at MNC Penang and to design applications around it that could better assist its engineers and management in decision-making would be advantageous.

Going forward, the C2P-KM framework could be taken beyond the context of MNC Penang’s needs but also that of the MNC in general. In view that cloud computing would be mainstream in the near future, it is hoped that any organization, irrespective of business domain, would eventually be able to capitalize and customize the proposed framework for its own use. This would result in a higher level of reusability, paving way for an open architectural standard for cloud-based knowledge management.

REFERENCES

- [1] N. Nikesh, "Emerging trends in knowledge management," Knowledge Management Term Paper, International School of Information Management, University of Mysore, 2009, <http://www.slideshare.net/nikeshn/emerging-trends-in-knowledge-management>. Accessed: 1 February 2013.
- [2] T. Wieringa, "Trends in knowledge management," 2010, http://www.greenchameleon.com/gc/blog_detail/trends_in_knowledge_management. Accessed: 1 February 2013.
- [3] D. Cerri, E. Della Valle, D. De Francisco Marcos, F. Giunchiglia, D. Naor, L. Nixon, K. Teymourian, P. Obermeier, D. Rebholz-Schuhmann, R. Krummenacher, and E. Simperl, "Towards knowledge in the cloud," OTM 2008 Workshops: On the Move to Meaningful Internet Systems, Lecture Notes in Computer Science, 5333/2008, 2008, pp. 986-995.
- [4] K.A. Delic, and J.A. Riley, "Enterprise knowledge clouds: Next generation KM systems?" 2009 International Conference on Information, Process, and Knowledge Management (eKnow 2009), Cancun, Mexico, 2009, pp. 49-53.
- [5] D. Talia, "Towards an open service framework for cloud-based knowledge discovery," Cloud Futures 2010 Workshop, Redmond, USA, 2010.
- [6] J.R. Beck, "The Cancer Biomedical Informatics Grid (caBIG): An evolving community for cancer research," in M.F. Ochs, J.T. Casagrande, and R.V. Davuluri, Eds., Biomedical Informatics for Cancer Research, Springer, 2010, pp. 177-200.
- [7] R.L. Ackoff, "From data to wisdom," Journal of Applied Systems Analysis, vol. 16, pp. 3-9, 1989.
- [8] M. Pechoucek, V. Marik, and J. Bárta, "A knowledge-based approach to coalition formation," IEEE Intelligent Systems, vol. 17, no. 3, pp. 17-25, 2002.
- [9] D. Nau, and J. Wilkenfeld, "Computational cultural dynamics," IEEE Intelligent Systems, vol. 23, no. 4, pp. 18-19, 2008.
- [10] N. Rubens, M. Vilenius, and T. Okamoto, "Automatic group formation for informal collaborative learning," 2009 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology, Milan, Italy, 2009.
- [11] P. Liu, H. Yu, and Q. Miao, "Automated planning for incident response based on CBR," 2010 IEEE International Conference on Information Theory and Information Security (ICITIS 2010), Beijing, China, 2010, pp. 403-406.
- [12] N. Mahiddin, Y.-N. Cheah, and F. Haron, "An ontology- and constraint-based approach for dynamic personalized planning in renal disease management," International Journal of Advanced Computer Science and Applications, vol. 2, no. 10, pp. 60-69, 2011.
- [13] OpenText, "Livelink," 2005, <http://www.opentext.com/2/global/products/products-all/livelink-landing.htm>, Accessed: 1 February 2013.
- [14] R. MacManus, "Top 10 semantic web products of 2008," 2008, http://readwrite.com/2008/12/02/top_10_semantic_web_products_2008, Accessed: 1 February 2013.
- [15] R. MacManus, "Top 10 semantic web products of 2009," 2009, http://readwrite.com/2009/12/02/top_10_semantic_web_products_of_2009, Accessed: 1 February 2013.
- [16] R. MacManus, "Top 10 semantic web products of 2010," 2010, http://readwrite.com/2010/12/29/top_10_semantic_web_products_of_2010, Accessed: 1 February 2013.
- [17] C. Pous, J. Colomer, J. Meléndez, and J.L. de la Rosa, "Case based reasoning as an extension of fault dictionary methods for linear electronic analog circuits diagnosis," in Recent Advances in Artificial Intelligence Research and Development, J. Vitrià, P. Radeva, and I. Aguiló, Eds., Amsterdam, The Netherlands: IOS Press, 2004.
- [18] J. Vajpai, and S. Dhoot, "Probabilistic intelligent fault diagnosis in television receiver circuit using visual symptoms," International Journal of Applied Engineering Research, vol. 4, no. 8, pp. 1535-1542, 2009.
- [19] V. Palade, C.D. Bocaniala, and L. Jain, Eds., Computational Intelligence in Fault Diagnosis. London, UK: Springer-Verlag, 2006.
- [20] S. Peenikal, "Mashups and the enterprise," Mphasis White Paper, 2009.
- [21] A. Wille, "Dashboard mashup: Business intelligence with data integration," Information Management Newsletters, 2010, 24 September 2010.
- [22] A. McAfee, "Enterprise 2.0: The dawn of emergent collaboration," MIT Sloan Management Review, vol. 47, no. 3, pp. 21-28, 2006.
- [23] O. Marks, and S. Patel, "Accelerating business performance," SOVOS Group White Paper, 2010.
- [24] M. Creeger, "Cloud computing: An overview," ACM Queue, vol. 7, no. 5, 2009.
- [25] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Communications of the ACM, vol. 53, no. 4, pp. 50-58, 2010.
- [26] H. Baars, and H.-G. Kemper, "Business intelligence in the cloud?" The 14th Pacific Asia Conference on Information Systems 2010 (PACIS 2010), Taipei, Taiwan, 2010, pp. 1528-1539.
- [27] R. Yasin, "Knowledge management in the cloud: Catalyst for open government?" Federal Computer Week, 2010, 3 May 2010.
- [28] L. Li, Y. Zheng, F. Zheng, and S. Zhong, "Cloud computing support for personal knowledge management," 2009 International Conference on Information Management, Innovation Management and Industrial Engineering, Xi'an, China, 2009, pp. 171-174.
- [29] Y. Wei, and M.B. Blake, "Service-oriented computing and cloud computing: Challenges and opportunities," IEEE Internet Computing, vol. 14, no. 6, pp. 72-75, 2010.

AUTHORS PROFILE

Yu-N Cheah is currently an associate professor at the School of Computer Sciences, Universiti Sains Malaysia. His research interests include knowledge management, intelligent systems, and health informatics. He has recently led several grant-sponsored projects on computational cultural dynamics, e-mail-based knowledge acquisition, automatic question answering, and dynamic planning. He holds a Ph.D. degree from Universiti Sains Malaysia and has published more than 80 journal and conference proceedings papers.

Soo Beng Khoh is a principal staff engineer and a certified Six Sigma Black-Belt of a MNC (name withheld for confidentiality). His research interests include real-time embedded systems, controller area network, field-buses, knowledge/strategic management, innovation process, and Six Sigma. Soo Beng holds a Ph.D. degree from University of Warwick, England in Electrical and Electronics Engineering, and is a member of IET and IEEE. He has published over 50 publications in IEE, IEEE, SAE transactions, journals and conferences.

DISCLAIMER

The opinions published in this paper are the authors' own, and by no means represent the company's opinion. All trademarks and companies mentioned in this paper do not imply recommendation or endorsement by the authors, nor that the products are necessarily the best available for the purpose. All trademarks remain the properties of the corresponding companies.

Pedagogy: Instructivism to Socio-Constructivism through Virtual Reality

Moses O. Onyesolu¹, Victor C. Nwasor², Obiajulu E. Ositanwosu³ and Obinna N. Iwegbuna⁴

^{1,3,4}Department of Computer Science

²Department of Educational Foundations

Nnamdi Azikiwe University, Awka Anambra State, Nigeria

Abstract—learning theories evolved with time, beginning with instructivism, constructivism, to social constructivism. These theories no doubt were applied in education and they had their effects on learners. Technology advanced, created a paradigm shift by creating new ways of teaching and learning as found in virtual reality (VR). VR provided creative ways in which students learn, provides opportunity to achieve learning goals by presenting artificial environments. We developed and simulated a virtual reality system on a desktop by deploying Visual Basic.NET, Java and Macromedia Flash. This simulated environment enhanced students' understanding by providing a degree of reality unattainable in a traditional two-dimensional interface, creating a sensory-rich interactive learning environment.

Keywords—learning theory; virtual reality; simulated environment; education; pedagogy

I. INTRODUCTION

Education in its general sense is a form of learning in which knowledge, skills and habits of a group of people are transferred through teaching, training, and research or auto didacticism [1]. It has been a means of transmitting one's culture from one generation to another. It is the process of bringing about a relatively permanent change in human behavior [2]. In its narrow, technical sense, education is the formal process by which society deliberately transmits its accumulated knowledge, skills, customs and values from one generation to another, e.g., instruction in schools [3]. Being the oldest industry, it is the main instrument used by society to preserve, maintain and upgrade its social equilibrium. A society's future depends largely on the quality of its citizen's education. Educators around the world today have leaning towards different theories of learning and as such some believe that these theories do not have to dictate how they teach or have to influence the way learner learn [4].

Instructivism, constructivism and socio-constructivism are the notable theories and practices on pedagogy today which still hinge on the works of earlier philosophy on teaching and learning, though earlier theorists in education are evenly divided on varying conception on teaching and learning. Technology has had effect on these theories as a result of man's ingenuity. In the fields of education, technology has created a paradigm shift, which has created new ways of teaching, and learning. Technology through intelligence has provided creative ways in which students learn and has provided opportunities to guide classroom instruction to meet

all learners [5] and [6]. Virtual Reality (VR), a technology of the 21st century has had impact on learning and changed the thinking of educators [7]. The remainder of the paper is presented as follows: Section 2 provided the background of some learning theories and Virtual Reality. Section 3 looked into some VR learning environments. Section 4 presented the materials and methods employed in the study. Section 5 presented the results and discussions while section 6 concluded the paper.

II. BACKGROUND INFORMATION

A. Instructivism

The Instructivist approach to learning is otherwise referred to as objectivist or institutionalist. Instructivism places emphasis on the passage of information and knowledge encapsulating activities and other learning events for learning to take place thereby resulting in a change of behaviour, attitude, belief etc. Instructivist helps learner to reproduce a series of facts, knowledge, attitude, belief and behavior [4]. The theory of instructivism maintains that knowledge should be transferred directly into the mind of the learner from the instructor. This knowledge is expected to be wholly accepted and not questioned by the learner or even analyzed by him [8]. This learning theory asserts that while the teachers are primary learning agents, students are passive information absorbers. It is thus a process of cognitive dumping in which all the learner has to do is to memorize unprocessed information. Instructivism is a very old learning approach. The lecture method of teaching is found under this domain, and this is where according to [8], 70% - 90% of university professors fall under. Instructivism/Instructionism involves making associations with stimuli, emphasizes drill and practices and memorization. This conditional learning does not give room for self-discovery. Instructivism, a traditional model of teaching and learning is nor learner-centered.

B. Constructivism

Although constructivism has been argued and often referred to as a new orthodoxy, fad, movement, religion with different sects and purely ideological theory [9] and [10], it cannot be over-emphasized that a learner enters his learning situation with pre-conceived notions or ideas about many phenomena [11]. The constructivist theory is rooted in the idea that the learner actively constructs his knowledge, not that it is passively acquired from without [11]. Learning is something the learner does, not that it is imposed on the learner. The

learner uses individual ideas as tools to understand many phenomena presented to him by the teacher.

Piaget contributed to the constructivist theory in Educational Psychology by delineating stages of intellectual development [12]. Influenced by the Piagetian thought afloat in educational circles, Driver and Easley [13] discovered the concrete ideas brought in as “Entry Behaviour” by children to the learning environment. This helped to change the research agenda in science education to encompass constructivist patterns in which learners construct their own knowledge as pot-pouri of ideas (self-generated from teachers). Educational constructivism has long been associated with advanced pedagogy on the basis that it champions a learner-centered approach to teaching, advocates learning in meaningful contexts, and promotes problem-based activities where learners construct their knowledge through interaction with their peers [14].

In constructivism, the teacher is seen as a facilitator/guide instead of being a director, in that learning allows for creative intercourse with the teacher, instead of teaching based on outcomes [15]. Discovery is seen to direct the learner to construct his own knowledge. Knowles, Holton and Swanson [16] gave impetus to this position by adding that constructivism emphasizes knowledge as being bound by context, giving room for individuals to make meaning out of their respective learning experiences.

C. Social Constructivism

With obvious influences from Piaget’s constructivist idea, social constructivism emphasizes the importance of culture and context in giving meaning to what happens in society and constructing knowledge from this understanding [17]. The social constructivist outgrowth of thought in educational psychology is associated chiefly with Lev Vygotsky, Jerome Bruner, and Albert Bandura [11].

The basic assumptions that underlie social constructivism are Reality, Knowledge, Learning and Intersubjectivity of Social Meanings. Social constructivists opine that human activity constructs reality. Reality is thus invented by society together [18]. Knowledge is also believed to be a product of human input and is constructed socio-culturally [19]; hence interactions with one another and their physical and social environments aids individuals in creating meanings.

Learning is seen as a social activity, not a passive behavioural development shaped by external, unempathetic forces. McMahon [17] has it that learning takes place as individuals involves in social activities. Describing intersubjectivity, Rogoff [20] refers to it as a shared understanding among individuals who interact based on common interest and assumptions that gives room for their communication. Social meanings are thus created based on intersubjectivity among individuals and social negotiation within communicating groups of individuals [21]. Vygotsky suggests that intersubjectivity helps individuals to extend their understanding of information to other group members. Such learning is collaborative in nature. In social constructivism, both the learning and social contexts are crucial to understanding.

D. Virtual Reality(VR)

VR is used to describe a range of computer-based systems in which a user can explore hardware and software generated „microworld’ that bears some resemblance to reality [22]. Virtual reality (VR) is a class of computer-controlled multisensory communication technologies that allow more intuitive interactions with data and involve human senses in new ways. VR is basically a way of simulating or replicating an environment and giving the user a sense of being there, taking control, and personally interacting with that environment with his/her own body [23] and [24]. VR is a technology which allows a user to interact with a computer-simulated environment, be it a real or imagined one [25]. VR is potentially a tool for experiential learning. The virtual world is interactive; it responds to the user’s actions. Virtual reality evokes a feeling of immersion, a perceptual and psychological sense of being in the digital environment presented to the senses. The sense of presence or immersion is a critical feature distinguishing virtual reality from other types of computer applications. Virtual Reality is an artificial environment created and maintained by a computer and that is at least partly shaped and determined by the user. A VR system allows the user to depart the real world and step into a world whose sensory inputs (sights, sounds, smells, etc.) are provided not by natural objects but by computer-created means. The objects and processes in the virtual world can then be manipulated to a large extent by the user. The use of VR techniques in the development of educational applications brings new perspectives to the teaching of subjects. An understanding of the technology and techniques is imperative.

VR is classified into three major types: (a) Non-Immersive VR Systems, (b) Semi-Immersive VR Systems and (c) Immersive (Fully Immersive) VR systems [23]. Non-Immersive VR Systems are the least implementation of VR techniques. It involves implementing VR on a desktop computer. This class is also known as Window on World (WoW) [26]. Using the desktop system, the virtual environment is viewed through a portal or window by utilizing a standard high resolution monitor. Interaction with the virtual environment can occur by conventional means such as keyboard, mouse or trackball. Semi-Immersive VR Systems comprise of a relatively high performance graphics computing system which can be coupled with either a large screen monitor; a large screen projection system or multiple television projection system. Using a wide field of view, these systems increase the feeling of immersion or presence experienced by the user and stereographic imaging can be achieved using some type of shutter glasses. Immersive (Fully Immersive) VR Systems are most direct experience of virtual environments.

Here the user either wears an head mounted display (HMD) or uses some form of head-coupled display such as a Binocular Omni-Orientation Monitor (BOOM) to view the virtual environment, in addition to some tracking devices and haptic devices. An HMD or BOOM uses small monitors placed in front of each eye which provide stereo, binocular or monocular images. More successful and popular is the Cave Automatic Virtual Environment (CAVE), In CAVE environments, the illusion of immersion is created by

projecting stereo images on the walls and floor of a room-size cube. Participants wearing lightweight stereo glasses enter and walk freely within the CAVE room, while a head-tracking computer system continuously adjusts the stereo projection to the current position of the viewer.

III. REVIEW OF RELATED WORKS

Simulations play a major role in education not only because they provide realistic models with which students can interact to acquire real world experiences, but also because they constitute safe environments in which students can repeat processes without any risk in order to perceive easier concepts and theories. VR is widely recognized as a significant technological advance that can facilitate learning process through the development of highly realistic 3D simulations supporting immersive and interactive features [27]. Simulated environments have opened new realms in teaching and learning and have found their way in all areas of human endeavors. VR technology has offered strong benefits in science and education. VR technology has not only facilitated constructivist and socio-constructivist learning activities but also supported different types of learners such as those who are visually oriented and disabled [28].

Sampaio and Henriques [29] demonstrated how the technology of VR can be used in the elaboration of teaching material of educational interest in the area of construction processes. They generated models that represented building in two standard situations. Students can interact with the virtual models in such a way that they can set in motion the construction sequence demanded by actual construction work, observe the methodology applied, analyze in detail every component of the work and the equipment needed to support the construction process and observe how the different pieces of a construction element mesh with each other and become incorporated into the model. Their models were used in disciplines involving construction in courses in Civil Engineering and Architecture administered by the Higher Technical Institute of the University of Lisbon.

Georgiou, Dimitropoulos and Manitsaris [27] presented a novel Web-based virtual learning environment for the simulation of volumetric analysis experiments. Their work took advantage of advances on Web and VR technologies to reproduce conditions of a real learning process in a chemical laboratory and enhance learning through a real-time interactive simulation of volumetric analysis experiments. The virtual laboratory presented in their paper was a cost-effective solution for both schools and universities without appropriate infrastructure and a valuable tool for distance learning and life-long education in chemistry.

Zhang [30] described a simulated environment called Second Life. It is a Web-based multi-user 3D virtual environment developed by Linden Lab, a San Francisco-based company. Second Life is one of the most popular virtual reality tools, attracting educators from all over the world. It offers a variety of opportunities for interaction, sense of community, and users' self-building capabilities. Statistics showed that there are over 100 educational institutes that had established their virtual campus in Second Life and are actively working in the virtual world [31].

Bell [32] proposed the need to introduce VR into engineering curriculum and in 1998, Bell and Fogler [33] developed Vicher (Virtual Chemical Reactors) at the University of Michigan to teach students catalyst decay, non-isothermal effects in kinetics, reactor design and chemical plant safety since they believe that humans retain up to 90% of what they learn through active participation [28]. Kim et al [34] developed a computer-based virtual reality simulation that helps students to learn physics concepts at the Kongju National University in Korea. This virtual laboratory has helped students' gain laboratory experience and thus improved students' performance [34]. In training and simulation, battlefield simulations have been developed using real data from Desert Storm [35]. US Navy uses flight simulators to help train pilots for general navigation as well as special assignments. US military uses VR simulations to treat phobia in war veterans.

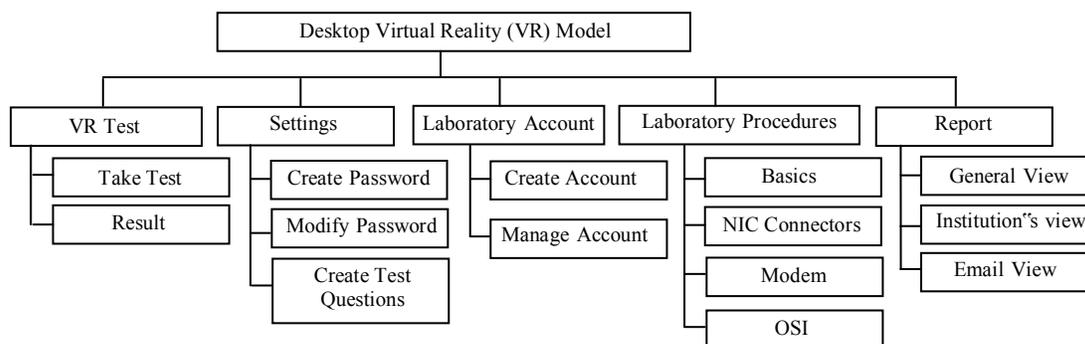


Fig. 1. The High Level Model of the Desktop Virtual Reality System

IV. MATERIALS AND MEHODS

We adopted a hybrid methodology derived from the combination of the Structured System Analysis and Design Methodology (SSADM), Object Oriented Analysis and Design Methodology (OOADM) and the prototyping methodology.

Using this methodology, we developed and implemented a desktop virtual reality model. The investigative phase of the SSADM was deployed as the paradigm for systematic study in order to obtain information on the current trends in the research area of simulated environments. The information obtained necessitated the definition of a high-level model

(HLM) for a simulated environment (Fig. 1). Using this HLM, a simulated desktop VR model was implemented. The simulated environment demonstrated Open System Interconnection (OSI) model and network devices (modem and network interface card). Macromedia Flash and Java were used to create the various objects in the simulated environment. Fig. 2 is the user interface of the model. It was designed to be user friendly. The interfaces were required to access information in the database and objects in the virtual world. Microsoft VB.NET and Java were used to implement the graphical user interface (GUI). The GUI was used for interactive querying, data capture, information display, and viewing of objects in the virtual world.

The desktop Virtual Reality Model was used in a class selected that offers fundamentals of networking in the Department of Computer Science at the Anambra State University to assess the impact of constructivism and socio-constructivism on the students. The selected class was divided into three groups. The subject of the assessment was the RJ45 network interface card (NIC) and connector for connecting computers in a single segment. The students were supposed to learn how an RJ45 NIC is identified and installed in a computer system; how to make an RJ45 connector and fitted to the NIC. Each group consisted of two classes of around 25. After presenting these, the students' knowledge improvements were tested and the students' attitudes about the simulated environment were surveyed using a questionnaire and the testing system in the simulated environment. The test consists of 20 numbers of multiple choice items on the subject. The different groups were tested about general knowledge on networking before the simulated environment was used on them.

The first group called "the instructivist group (IG)" did not use the simulated environment at all. In the second group, the simulated environment was used in the class only by a teacher as a supplement to lecture. This group was called "the constructivist group (CG)". The third group called "the socio-constructivist group (SG)" used the simulated environment in the PC room where the students were actively engaged by themselves without any lecture. They had only the study guide papers and were expected to learn by themselves and interact with one another to share their experiences.

V. RESULTS AND DISCUSSIONS

The result of the simulated environment is presented in Fig. 2 to Fig. 8 while the impact of this simulated environment on students is presented in Fig. 9 to Fig.10 and Tables 1 to 2.

A. The Simulated Environment

VR is a powerful tool for education since people comprehend images much faster than they grasp lines of text or columns of numbers. Participation is critical to learning and VR offers multisensory immersive environments that engage students and allow them visualize information.

User Interface: The graphical user interface (GUI) is shown in Fig. 2. The interfaces are required to access information in the database and objects in the virtual world. Microsoft Visualbasic.NET was used to implement the GUI. The GUI

was used for interactive querying, data capture, information display, and viewing of objects in the virtual world.



Fig. 2. GUI of the Desktop Virtual Reality System

a) *Login Interface*: Fig. 3 shows a screen shot of the login user interface. A user who has created an account in the environment uses the login interface to gain access to the simulated environment.

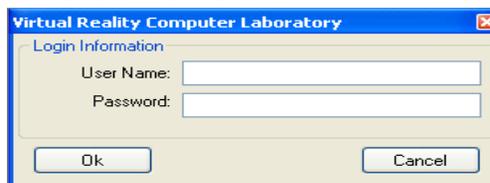


Fig. 3. Login user interface

b) *Create Account*: Fig. 4a and Fig. 4b show screen shots of create account. A new user provides information which is stored in the database. This information is needed so as to monitor users as regards the use of the model and subsequently in taking tests. A new user is first prompted to key in his/her email address and registration number using Fig. 4b. The database is searched to see if these exist. This is so to avoid duplication of record. If such exists, the user is denied registration thus account is not created for such user, else the user is prompted using Fig. 4a to provide other details.

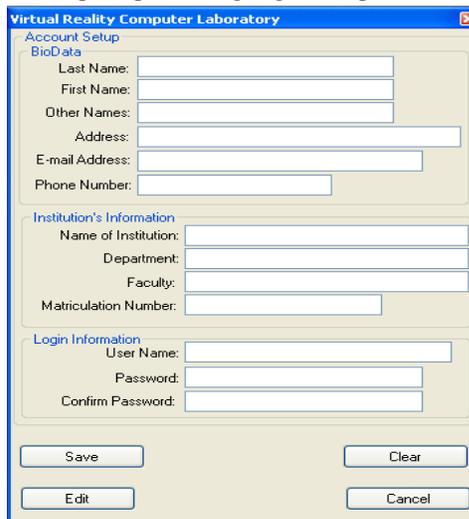


Fig. 4. a: Create account graphical user interface

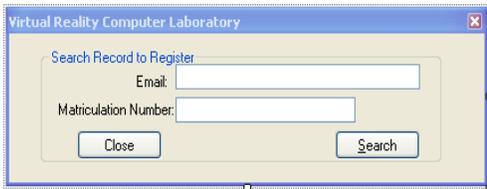


Fig. 4. b:Search database interface

c) *Record View Functionality:* Fig. 5 shows a screen shot of the view records user interface. This interface is connected to the database using a data source VRLDbase.Dataset.xsd. The field names were chosen to be as close to the real field names to avoid a situation where users become confused as to what each field represents. For identification of any record of choice once it is clicked, it will be highlighted. The application (model) successfully allowed the use of the user interface to present an integrated view of the database information in a transparent manner. The different fields not needed are filtered out leaving only the necessary fields needed.

Last Name	First Name	Address	Email	Name of Institution	Department	Registration Number	Lab User Name
Omyesolu	Moses	UNIZIK, Awka, Anambra State	omyesolu_mos@yahoo.com	Nnamdi Azikiwe University	Computer Science	2005517002P	omyesolu_mos
Omyesolu	Isaac	UNN, Nsukka, Enugu State	isaac_omyesolu@yahoo.com	University of Nigeria	Pharmacy	2001020870	isaac_omyesolu
Osiolowo	Raphael	ESUT, Enugu, Enugu State	osioloworaph@yahoo.com	Enugu State University of Technology	Computer Engin...	1999342123	osioloworaph
Imeh	Mary Rose	UNIZIK, Awka, Anambra State	imemary@yahoo.com	Nnamdi Azikiwe University	Computer Science	2005517002P	imemary
Ogbede	Kudinat	UNILAG, Akoka, Yaba, Lagos	kudinatogbede@yahoo.com	University of Lagos	Mass Communic...	1999342123	kudinatogbede
Ogbede	Isaac	No 6, Amobi Street, Awka	isacogbede@yahoo.com	University of Oba...	Informatic	200720810	isacogbede
Egboika	Chidiebere	No 34, Baba Olucho Road, (Ifecha, La...	egboika@hotmail.com	Anambra State University	Civil Engineering	2007071214	egboika
Chidiebere	Henry	No 21, Ogburn Road, Kaduna	henrychidiebere@yahoo.com	Abia State University	Psychology	2005517054	henrychidiebere
Kafung	Shehu	No 34, Kaduna Road, Sabon Gari, Kano	kafungshehu@hotmail.com	Usman Dan Fodio University	Strategic Studies	2006342156	shehu
Awake	Isamaka	No 34, Okonkwo Street, Benin City	isamakaawake@yahoo.com	University of Benin	Electrical Engineering	1999342123	isamakaawake
Osiolowo	Osiolowo	ESUT, Enugu, Enugu State	osiolowoo@yahoo.com	Enugu State University	Computer Engin...	2005517002P	osiolowoo
Appaolu	Ogbonna	UNIZIK, Awka, Nigeria	ogbonnaappalolu@yahoo.com	Nnamdi Azikiwe University	Computer Science	2005514032	ogbonna
Omyesolu	Osiolowo	No 25, Fagge, Onitsha	omyesolu@yahoo.com	University of Benin	Pharmacy	1999342123	omyesolu

Fig. 5. Record View User Interface

d) *Query Record Functionality:* Once the model was executed and loaded, from the laboratory account menu, the database can be queried either if a new account is being created or an existing account is to be managed by editing one or more fields not desirable. The screen shots in Fig. 4a and Fig. 6 are one out of two possible ways to query a record in the database. In Fig. 4a, the query criteria are email and registration number, while in Fig. 6, the query criterion is email. The application was successfully used through the user interface to select condition used to query the records in the database.

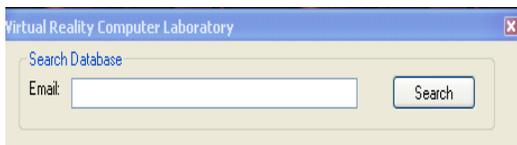


Fig. 6. Query Records User Interface

Laboratory Procedure Functionality: The two laboratory procedures presented were modem – a networking device, and OSI – a networking concept. These two procedures were presented in the form of movies accompanied by a voice

output to describe the concept or the device. The movie can be paused, played, and stopped. It can be made to be viewed in full screen when desired. These objects can be manipulated (turned in 3D) to study its essential parts. The screen shots are presented in Fig.7 and Fig. 8. The laboratory procedures demonstrated clearly showed how objects in the virtual world interact, and how these objects are used. The OSI model concept demonstrated in graphical forms what happens to data travel as it originates from the application layer, travels through the layers to the physical layer and vice versa. Because these objects were presented in 3D, comprehension of the concept and the working of the devices were made easy to a learner or user. At the end of a laboratory procedure, a user may want to take a test to determine how far he/she has grasped the concept learnt or the device studied. This is done by taking a test by selecting the test option on the menu. The test provided series of multiple choice questions and each question has four options to choose from. A user can stop the test anytime desired without completing the test. At the end of the test, the test takers name and score obtained are stored in a database. These can be accessed later if desired, or if the user wished to continue with the test.

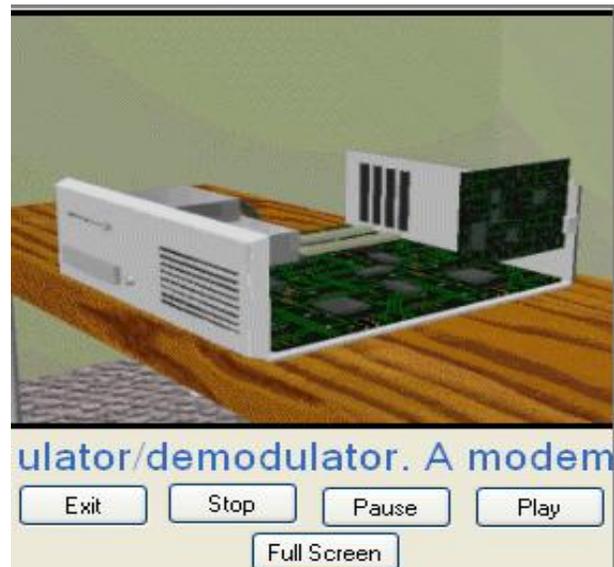


Fig. 7. Connecting modem laboratory procedure

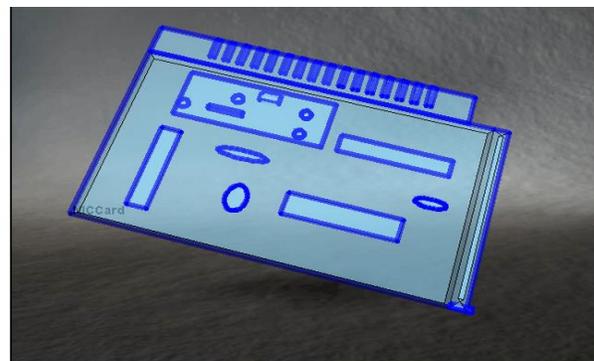


Fig. 8. Viewing and manipulating an object (modem) procedure

B. The Impact of the Environment on Students

TABLE I. PERFORMANCE OF THE DIFFERENT GROUP BEFORE THE TEST

Academic Performance			
Group	IG	CG	SG
Highest Score	18	18	17
Lowest Score	10	08	09
Range of Score			
01-04	Nil	Nil	Nil
05-08	Nil	04	Nil
09-12	12	14	16
13-16	08	04	05
17-20	05	03	04

Table I and Fig. 9 show the academic performance of each group before they were exposed to the simulation program and taught the network components. The IG group's performance was slightly superior to other groups. While the highest score in this group was 18, 5 students scored between 17-20 marks. In contrast to CG group and SG group, 3 students and 4 students scored in that range.

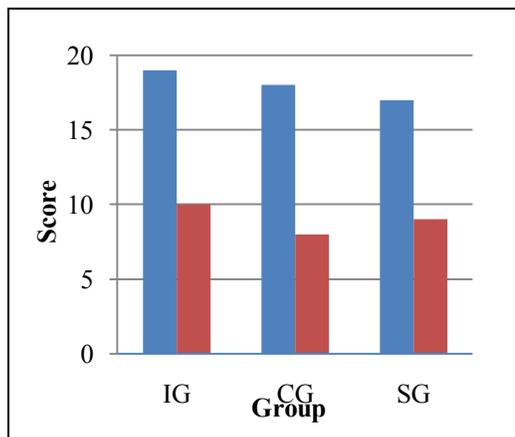


Fig. 9. The academic performance of groups before the test

Table II and Fig. 10 show the academic performance of each group after they were exposed to the simulation program and taught the network components. The SG group's performance was superior to other groups. While the highest score in this group was 19, 9 students scored between 17-20 marks.

In contrast to CG group and SG group, no student scored within that range. The SG group's performance was superior to other groups because it has been demonstrated by [28] that human beings retain up to 90% of what they learn through active participation.

TABLE II. PERFORMANCE OF THE DIFFERENT GROUP AFTER THE TEST

Academic Performance			
Group	IG	CG	SG
Highest Score	13	16	19
Lowest Score	08	10	14
Range of Score			
01-04	Nil	Nil	Nil
05-08	06	Nil	Nil
09-12	12	13	Nil
13-16	07	12	16
17-20	Nil	Nil	09

This facilitated constructivist and socio-constructivist learning. Therefore simulated environments provided the tools with which the students used, visualized and manipulated objects; this made it easier for them to understand how the objects work. It offered tools for increased student participation and a learning experience that they found interesting, this gave them the motivation to learn, interact and shared their experiences with one another.

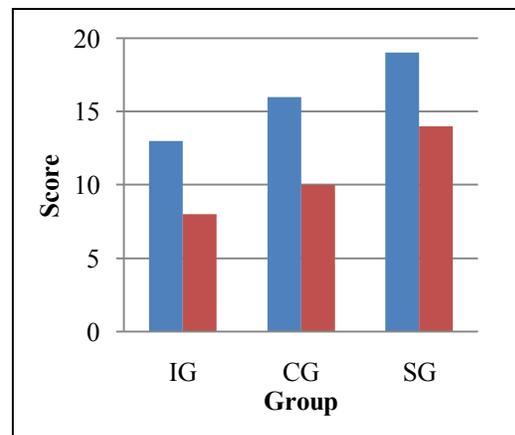


Fig. 10. The academic performance of groups after the test

VI. CONCLUSION

Researchers in the field of learning theory and VR have generally agreed that VR technology is exciting and provides a unique and effective way for students to learn when it is appropriately designed and applied, and that VR projects are highly motivating to learners [36]. From various researches, several specific situations have emerged in which VR has strong benefits or advantages (as shown in this work). For example, VR has great value in situations where exploration of environments or interactions with objects or people is impossible or inconvenient, or where an environment can only exist in computer-generated form.

VR is also valuable when the experience of actually creating a simulated environment is important to learning. Creating their own virtual worlds has been shown to enable some students to master content and to project their understanding of what they have learned. One of the beneficial uses of VR occurs when visualization, manipulation, and interaction with information that is critical for its understanding. This is socio-constructivist approach of VR to learning. It is because of VR's capacity for allowing learners to display and interact with information and environment that some believed is its greatest advantage.

Finally, VR is a very valuable instructional and practice alternative when the real thing is hazardous to learners, instructors, equipment, or the environment. This advantage of the technology has been cited by developers and researchers from such diverse fields as firefighting, anti-terrorism training, nuclear decommissioning, crane driving and safety, aircraft inspection and maintenance, automotive spray painting and pedestrian safety for children [36].

In the fields of education, technology has created a paradigm shift. Technology has created new ways of teaching and learning. VR is a new technology that has provided creative ways in which students learn opportunities to guide classroom instruction to meet all learners.

The use of VR techniques in the development of educational applications brings new perspectives to the teaching of subjects. VR is potentially a tool for experiential learning. VR has opened new realms in the teaching, learning, and practice of medicine, physical sciences and engineering among others. Virtual reality learning environments (VLEs) provide students with the opportunity to achieve learning goals.

VLE-based applications have emerged in mainstream education in schools and universities as successful tools to replace traditional teaching methods (instructivist approach). VR learning environments have been discovered to have greater pedagogical effectiveness on learners because it encompasses instructivism, constructivism and culminated in socio-constructivism.

REFERENCES

- [1] M.O. Onyesolu, "Extent of computer education acquisition in Nigeria: Conflict between policy pronouncements and policy implementations," *Journal of Applied Sciences*, vol. 12, no. 1. pp. 8212-8221, March 2009.
- [2] A. Esu, and A. Junaid, "Educational development: Traditional and contemporary." Retrieved February 30, 2013. Source: <http://www.onlinenigeria.com/education/>
- [3] M.O Onyesolu, O.F Oladipo, and I.M. Ezeani, "Simulating desktop virtual reality for higher education," *International Journal of Ubiquitous Sciences*, vol. 1, issue 1, pp. 1-6. June, 2012.
- [4] O. Olufemi, "Pedagogical approaches and technical subject teaching through internet media," *The Electronic Journal of e-Learning*, vol. 6, no.1, pp.53-66, 2008.
- [5] R. Ginberg and V. McCornick, "Computer use in effective schools," *Journal of Staff Development*, vol. 19 no. 1, pp. 22-5, 1998.
- [6] E. Hyslop-Margison, "Technology, human agency and Dewey's constructionism: Opening democratic spaces in virtual classrooms," *Austrarian Journal of Educational Technology*. vol. 20, no. 2, pp. 137-148, 2004.
- [7] M.O. Onyesolu, "Virtual Reality: An emerging computer technology of the 21st century," *International Journal of Electrical and Telecommunication System Research*. vol.1, no. 1, pp. 36-40, 2006.
- [8] D.P. Diaz and K.F. Bonteball, *Pedagogy-based Technology Training*. Amsterdam: IOS Press, 2000.
- [9] M. Matthews, "Discontent with constructivism," *Studies in Science Education*, vol. 24, pp. 165-172, 1994.
- [10] G. Erikson, Research programmes and the student science learning literature. In R. Millar, J. Leach, J. Osborne (Eds.): *Improving Science education: The contributions of research*. Buckingham: open University Press, 2000.
- [11] K.S. Taber, "Beyond Constructivism: The progressive research programme into learning science," *Studies in Science and Education*, vol. 42, pp. 125-184, 2006.
- [12] S. Sjoberg, In E. Baker, E., B. McGaw, and P. Peterson (Eds.), *International Encyclopaedia of Education*, 3rd Ed., Oxford: Elsevier, 2007.
- [13] R. Driver and J. Easley, "Pupils and Paradigms: A review of literature related to concept development in adolescent science students," *Studies in Science Education*, vol. 5, pp. 61-84, 1978.
- [14] H. Nikitina, "Addressing pedagogical dilemmas in a constructivist language," *Journal of the Scholarship of Teaching and Learning*. vol. 10, no. 2, pp. 90-106, 2010.
- [15] H. Huang, "Towards constructivism for adult learners in online learning environments," *British Journal of educational Technology*, vol. 33, no. 1, pp. 27-37, 2002.
- [16] M.S. Knowles, E.F. Holton and R.A. Swanson, *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 7th Edition, London: Elsevier, 2011.
- [17] M. McMahon, Social constructivism and thw world wide web-A paradigm for learning. *Proceedings ASCILITE97 Conference*, Curtin University, Perth, 1997. Retrieved March 13, 2013. Source <http://www.ascilite.org.au/conferences/perth97/papers/Mcmahon/Mcmahon.html>
- [18] A. Kukla, *Social Constructivism and the Philosophy of Science*, New York: Routledge, 2000.
- [19] R.S. Prawat and R.E. Floden, "Philosophical perspectives on constructivist views of learning," *Educational Psychologist*, vol. 29, no. 1, pp. 37-48, 1994.
- [20] B. Rogoff, *Apprenticeship in thinking: Cognitive development in social context*. New York, NY: Oxford University Press, 1990.
- [21] M.E. Gredler, *Learning and instruction: Theory into practice*, (3rd Ed.). Upper Saddle River, New Jersey: Prentice-Hall, 1997.
- [22] C. Macpherson and M. Keppell, "Virtual reality: What is the state of play in education?" *Australian Journal of Educational Technology*. vol. 14, no. 1, pp. 60-74, 1998.
- [23] M.O. Onyesolu, "Modeling and Simulation of Desktop Virtual Reality," Unpublished Doctoral Dissertation, Department of Computer Science, Nnamdi Azikiwe University, Awka, Nigeria, 2011.
- [24] M. Slater, M. Usoh and A. Steed, "Depth of presence in virtual environments," *Teleoperators and Virtual Environments*, vol. 3, no. 2, pp. 130-144, 1994.
- [25] Wikipedia the free encyclopedia (2013). Virtual Reality. Retrieved March 13, 2013. Source: http://en.wikipedia.org/wiki/Virtual_reality
- [26] M.O. Onyesolu and F.U.Eze, *Understanding Virtual Reality Technology: Advances and Applications*, in M. Schmidt, (Ed.) *Advances in Computer Science and Engineering*, Rijeka, Croatia: InTech Open Access Publishers, 2011.
- [27] J. Georgiou, K. Dimitropoulos and A. Manitsaris, "Virtual reality laboratory for distance education in chemistry," *International Journal of Social Sciences*, vol. 2, no. 1, pp. 34-41, 2007.
- [28] D. Edgar, *Audio-Visual Methods in Teaching*. 3rd ed. London: Holt, Rinehart, and Winston, 1969.
- [29] A.Z. Sampaio and P.G. Henriques, "Virtual Reality technology applied in Civil Engineering education," *Current Developments in Technology-Assisted Education*, pp. 1351-1355, 2006.

- [30] J. Zhang, "Second life: Hype or reality?" Higher Education in the Virtual World, 2007. [Online]. Available: <http://deoracle.org/online-pedagogy/emerging-technologies/second-life.html>
- [31] K. Joly, A second life for higher education, 2007. Retrieved March 15, 2013. Source: <http://www.universitybusiness.com/viewarticle.aspx?articleid=797>
- [32] J.T. Bell, Introducing virtual reality into the engineering curriculum, Proc. of University Programs in Computer Aided Engineering and Design Manufacturing, Charlottesville, VA., 1996. [Online], Available: <http://www.vrupl.evl.uic.edu/vrichel/>.
- [33] J.T. Bell and H.S. Fogler, H.S. Virtual Reality in the Chemical Engineering Classroom, Proc. of American Society for Engineering Education Annual Conference, Seattle, WA., 1998.
- [34] J. Kim, S. Park, K. Yuk, H. Lee and H. Lee, Virtual reality simulations in physics education. Interactive Multimedia Electronic Journal of Computer-Enhanced Learning (IMEJ of CEL), 2000. [Online]. Available: <http://imej.wfu.edu/articles/2001/2/02/index.asp>
- [35] B.Schmitz, "Virtual reality: On the brink of greatness," Computer-Aided Engineering, pp. 26-32, 1993.
- [36] L.J. Ausburn and F.B. Ausburn, "Desktop virtual reality: A powerful new technology for teaching and research in industrial teacher education," Journal of Industrial Teacher Education, vol. 41, no. 4, 2004. [Online]. Available <http://scholar.lib.vt.edu/ejournals/JITE/v41n4/ausburn.html>

AUTHORS PROFILE

Moses Okechukwu Onyesolu has a Ph.D (Virtual Reality), M.Sc. (Internet Computing), B.Sc. (Computer Science) from Nnamdi Azikiwe University, Nigeria where he works as a lecturer and researcher. His research interests are mainly in computer modeling and simulation, e-learning/virtual reality technologies, queueing system/theory and its applications. He has published widely in those areas. He is a member of the following learned societies: Nigerian Computer Society (NCS), Computer Professionals (Registration Council of Nigeria) (CPN), and International Association of Engineers (IAENG), International Association of Computer Science and Information Technology (IACSIT) and European Association for Programming Languages and Systems (EAPLS)

Victor Chekume Nwasor is with the Department of Educational Foundations, Nnamdi Azikiwe University, Awka, Anambra State. He is a member of the following learned societies: Nigerian Council for Educational Psychologists and Nigerian Society for Educational Psychology.

Obiajulu Ekene Ositanwosu has B.Sc. (Computer Science) from Nnamdi Azikiwe University, Awka, Anambra State where he works as a teaching assistant. He is a graduate student in the Department of Computer Science, Nnamdi Azikiwe University. His research interests are in the area of e-learning, biometrics/securities and computer modeling and simulation.

Obinna Nnamdi Iwegbuna has B.Sc. (Computer Science) from Nnamdi Azikiwe University, Awka, Anambra State where he works as a technologist. He is a graduate student in the Department of Computer Science, Nnamdi Azikiwe University. His research interests are in the area of e-learning and biometrics/securities.

A Soft Processor MicroBlaze-Based Embedded System for Cardiac Monitoring

El Hassan El Mimouni
University Sidi Mohammed Ben Abdellah
Fès, Morocco

Mohammed Karim
University Sidi Mohammed Ben Abdellah
Fès, Morocco

Abstract—this paper aims to contribute to the efforts of design community to demonstrate the effectiveness of the state of the art Field Programmable Gate Array (FPGA), in the embedded systems development, taking a case study in the biomedical field. With this design approach, we have developed a System on Chip (SoC) for cardiac monitoring based on the soft processor MicroBlaze and the Xikernel Real Time Operating System (RTOS), both from Xilinx. The system permits the acquisition and the digitizing of the Electrocardiogram (ECG) analog signal, displaying heart rate on seven segments module and ECG on Video Graphics Adapter (VGA) screen, tracing the heart rate variability (HRV) tachogram, and communication with a Personal Computer (PC) via the serial port. We have used the MIT_BIH Database records to test and evaluate our implementation performance. In terms of the resources utilization, the implementation occupies around 70% of the used FPGA, namely the Xilinx Spartan 6 XC6SLX16. The accuracy of the QRS detection exceeds 96%.

Keywords—ECG; FPGA; Heart Rate Variability; MicroBlaze; QRS detection; SoC; Xikernel

I. INTRODUCTION

A Field Programmable Gate Array (FPGA) is a high density Programmable Logic Device (PLD) that allows high performance data processing. Its digital signal processing (DSP) performance is derived from the ability to construct naturally parallel structures or modules, achieving a hardware multitasking. This rapid increase of complexity emerged also the new concept of SoC that integrates most of the functions of the end product into a single chip; this is encouraged by the reuse of optimized Intellectual Properties (IPs), particularly soft and hard processors; in fact, FPGAs are nowadays so dense that it is possible to embed a processor (soft or hard) on the single chip and there is still enough room for eventual additional functionalities. Moreover, these FPGAs come with sophisticated software tools for the processor, like C/C++ compilers and a RTOS kernel; this improves the design abstraction and consequently the productivity [1 - 4].

The electrocardiogram signal (ECG) is one of the most commonly used in medical practice thanks to its non-invasive nature, simple acquisition process and the meaningful information it contains; the analysis of such information permits to evaluate the state of the heart. Thus, the cardiac monitoring by means the ECG is a standard practice in intensive care, emergency rooms, ambulatory monitoring, etc. It is worth noting that the cardiac rhythm monitoring by means the heart rate variability (HRV), increasingly studied in the

recent years, and has become also an important way to assert the heart's condition [5 - 9].

II. SYSTEM OVERVIEW

We have designed and implemented a prototype of basic embedded system for cardiac monitoring, whose functional block diagram is shown in the figure 2; it exhibits a modular structure that facilitates the development and debugging. Thus, it includes 2 main modules:

- An analog module intended for acquiring and conditioning of the analog ECG signal to make it appropriate for use by the second digital module ;
- A soft processor-based module digitizes and processes the ECG signal.

A secondary module acts as a way to facilitate testing system with the standard ECG records contained in the standard MIT-BIH database [10].

This implementation is to be considered as a strong improvement and a migrating to advanced new technology, mentioned above, of one of our previous works [11]; it can also be seen as a contribution to developing a more efficient and low-cost biomedical instrumentation.

ECG signal is the measure, via electrodes acquiring the voltage (potential difference) on the body surface, generated by the heart's electrical activity. As illustrated by the figure 1, the ECG is characterized mainly by 5 waves reflecting the activity of the heart during a cardiac cycle (R-R interval); these waves are called P, Q, R, S and T; the Q, R, and S waves are treated as a single composite wave known as the QRS complex. The ECG signal is typically characterized by maximum amplitude of 1 mV and a bandwidth of 0.05 Hz to 100 Hz [9, 12].

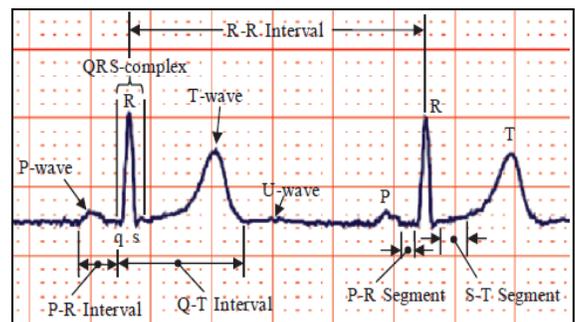


Fig. 1. Typical ECG of a healthy person.

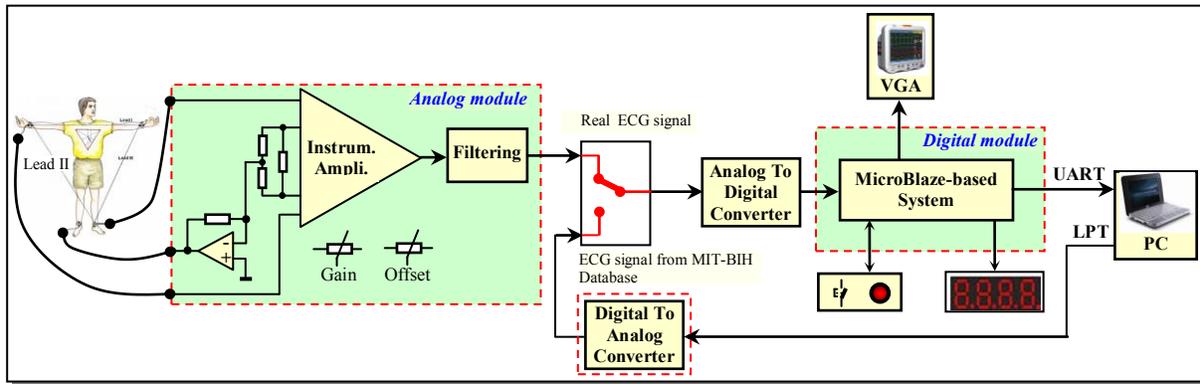


Fig. 2. System functional diagram block.

III. ANALOG MODULE

We used the Lead II that is the most commonly used lead for ECG ambulatory monitoring, because it's a relative high-voltage deflection resulting in P, R and T waves. The reference electrode is connected to the right leg via an amplifier to reduce common mode noise, a principle commonly termed by "Right Leg Drive". The two other electrodes, representing Lead II, attack an instrumentation amplifier (IA) we have achieved with off the shelf operational amplifiers, which present good performances, especially a relative good common mode rejection ratio (CMRR). The different types of noise are reduced by means analog filtering: low pass and anti-aliasing filter of 0-70 Hz, and notch filter for reducing noise effect of 50 Hz of AC power line. Finally, the output signal gain and offset of this stage are properly adapted to the analog to digital converter (ADC) of the digital module.

IV. SOFT PROCESSOR-BASED DIGITAL MODULE

Today's FPGAs integrate existing IP cores achieving functionalities commonly used in the embedded systems world (GPIO, Timers, UART, SPI, VGA, etc.); they can be easily instantiated into a top-level design. Among them, the most important are the soft processors cores; indeed, the availability of such embedded subsystems in FPGAs opens a whole new world of possibilities. A soft or virtual processor is built by combining blocks of optimized HDL code inside an FPGA. In our case, it was the MicroBlaze, which is a reduced instruction set computer (RISC) optimized for implementation in Xilinx FPGAs [13].

The development of this module was guided by the Hardware/Software co-design techniques, which try to exploit the synergy of Hardware and Software with the goal to optimize and satisfy design constraints of a final product [14]. Thus, as development tool, we have used the Xilinx Embedded Development Kit (EDK), which is a suite of tools and Intellectual Property (IP) that permits the design of a complete embedded processor system for implementation in a Xilinx FPGAs [15]. We mention in particular:

- The Xilinx Platform Studio (XPS) that is the development environment used for designing the hardware aspect of an embedded processor system.

- The Software Development Kit (SDK) that is an integrated development environment, complementary to XPS used for C/C++ embedded software application creation and verification.

As development board, we have used the Nexys 3 of Digilent that features Xilinx's Spartan-6 XC6SLX16 FPGA, 48 Mbytes of external memory, and enough I/O devices and ports to host a wide variety of digital systems [16].

A. Hardware

The foundation of the hardware of the design is created using the Base System Builder (BSB) wizard within XPS. This allows the use of pre-developed IPs cores with a series of Buses and Interfaces to connect the various hardware elements of the design. The figure 3 shows the architecture of our implementation; the design consists of the main following IPs cores:

- MicroBlaze soft processor is the main and central element of the entire architecture; in other words, it represents the central and processing unit (CPU) of the system.
- The Digital Clock Manager (DCM) primitive in FPGA provides advanced clocking capabilities; it optionally multiply or divide the incoming clock frequency to synthesize a new clock frequency.
- Bus Local memory Bus (LMB) provides single-cycle access to on-chip RAM.
- On-chip dual-port block RAM (BRAM) stores processor's program instructions and data; as the application program exceeds the BRAM limit (32 Kbytes), the external SRAM is used for this purpose.
- The Processor Local Bus (PLB) provides a connection to both on-and off-chip peripherals and memory.
- Timer 0 is required by Xilkernel RTOS to tick its kernel.
- Timer 1 is used to generate an interrupt every 5 ms (200 Hz) for ECG analog signal sampling, according to the Shannon theorem.

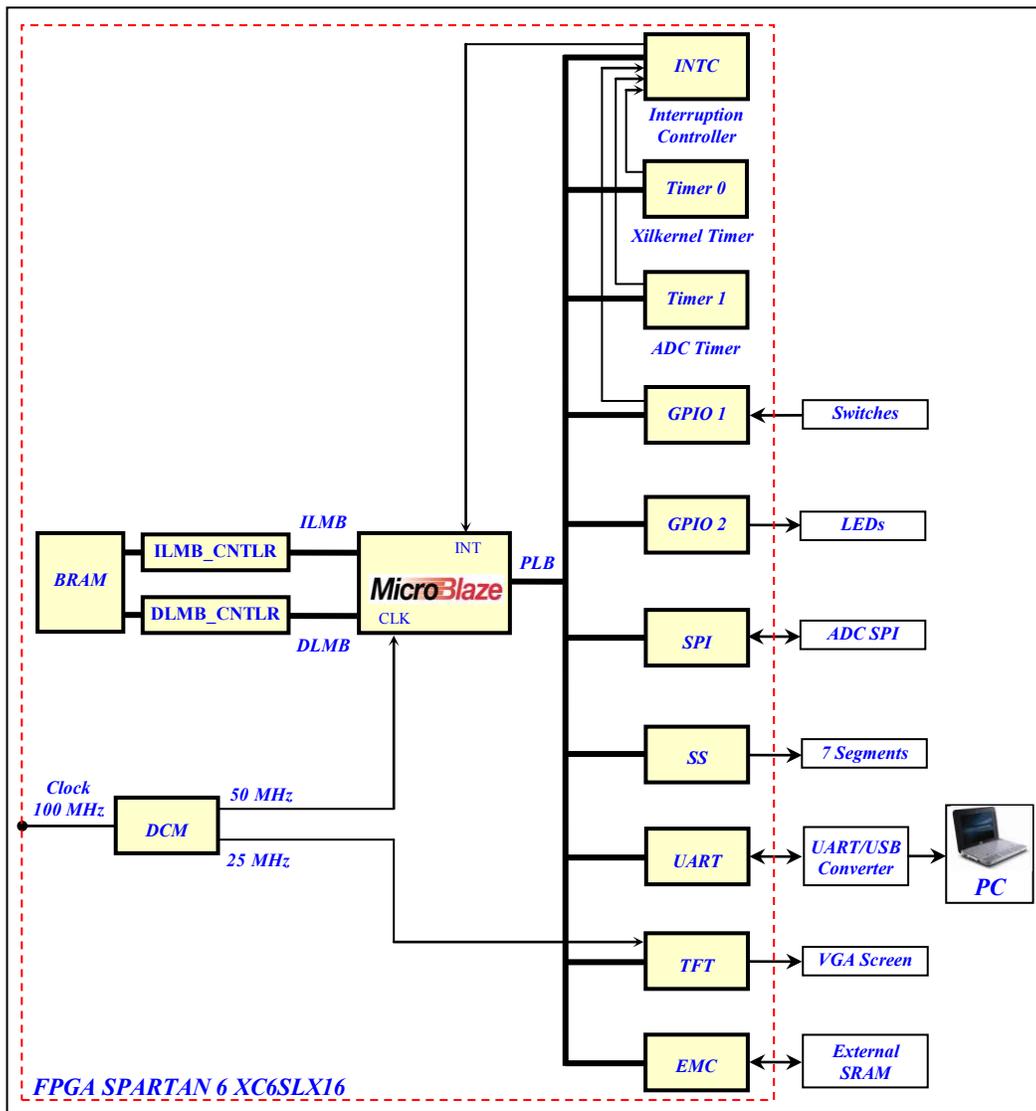


Fig. 3. MicroBlaze-based hardware.

- General Purpose Input Output (GPIO) is used to communicate with simple human machine interface (HMI) that consists of switches and LEDs.
- INTC is an interrupt controller used to concentrate multiple interrupt inputs from peripheral devices to a single interrupt output driving the unique processor interrupt input; hence, in this application, it captures kernel tick interrupt, 5 ms interrupt and so on.
- SPI (Serial Peripheral Interface) is the IP core providing a serial interface to the small board PmodAD1 from Digilent, which features ADCS7476 that is an SPI ADC serving for ECG analog signal digitizing.
- UART (Universal asynchronous Receiver Transmitter) is the IP core allowing serial communication with a PC (COM port), via an UART/USB converter, that is the FT232RQ; indeed, there is no COM port in the most modern PCs.
- SS is a customized IP from Digilent; it controls the seven segments with time multiplexing, intended to heart rate display; that is a good example for Hardware/Software co-design illustration; in fact, this task could have been achieved by a certain C code function executed by the CPU; this customized IP offloads then the processor.
- TFT (Thin Film Transistor) is the IP serving as interface for 18-bit VGA; it supports 25 MHz clock, generated from DCM, for display resolution of 640x480 pixels at 60 Hz refresh rate.
- EMC (External Memory Controller) is the interface for on-board SRAM, which contains the application program and the video memory for VGA display.

B. Software

As the complexity of the embedded systems increases, the development in an RTOS environment is essential; it becomes even a normal and obvious design activity, with the modern

tools available in the new computer aided development (CAD) tools, like EDK. Thus, an application is broken into small pieces (tasks or threads) more easy to manage; each task deals with certain aspect of the application. In our case, the RTOS was Xilkernel from Xilinx. Xilkernel is a small, light-weight easy to use, robust and modular kernel. It provides features like scheduling, threads, inter-process communication (IPC) and synchronization, with a POSIX subset interface. It is integrated with EDK, easy to configure and free software [17].

Thus, the software application consists of Xilkernel threads executing on top of the kernel. In the following text of this section, we describe the RTOS structure of our software written in C language and in a bit more detail the QRS detection algorithm.

1) Software Structure

The figure 4 illustrates the software structure of the application in the Xilkernel RTOS context. We have configured Xilkernel to have round robin scheduling. Once Xilkernel is called and initialized, it calls its scheduler to manage the different application tasks:

- The first task to run is the “main_thread”; it has the highest priority (0) and therefore always runs before any other tasks; this task creates the others threads and achieves necessary initializations and exits.
- “IHM_thread” is the task where the state of a switch is read permitting the choice between acquisition of ECG or reading a small sequence of ECG signal stored in an array for testing purposes; A LED (LD0 in the board) blinks at measured cardiac rhythm.

- “HR_thread” is the task responsible for calculating the heart rate for 10 s with an interpolation over 1 mn.
- “VGA_thread” deals with the display of the ECG in VGA screen with a resolution of 640 x 480.
- “interrupt_handler” is the interrupt service routine (ISR) that can be considered as a thread of high priority; that is the Timer 1 interrupt allowing processing critical events, like real-time acquiring of ECG, QRS detection and so on.

We have used as means of communication between threads:

- Protected global variables, so that there is no conflict in access to these variables.
- Message queue, an available technique in Xilkernel.

2) QRS Detection

The QRS complex of the ECG signal is the reference point for the most ECG applications. In this paper we adapted the algorithm subject of our previous work [18], in the MicroBlaze context. This detector uses a modified Pan and Tompkins algorithm [19], which is the most widely used and highly acknowledged algorithm, for its real-time aspect, robustness and efficiency; its structure with 2 stages is shared by many algorithms in the recent years:

- The pre-processing stage has for purpose enhancing QRS complex with noise suppressing and artifacts.
- The decision stage determines the QRS candidates, based on thresholding technique and a set of rules; this detection provides a logic signal (QRS flag).

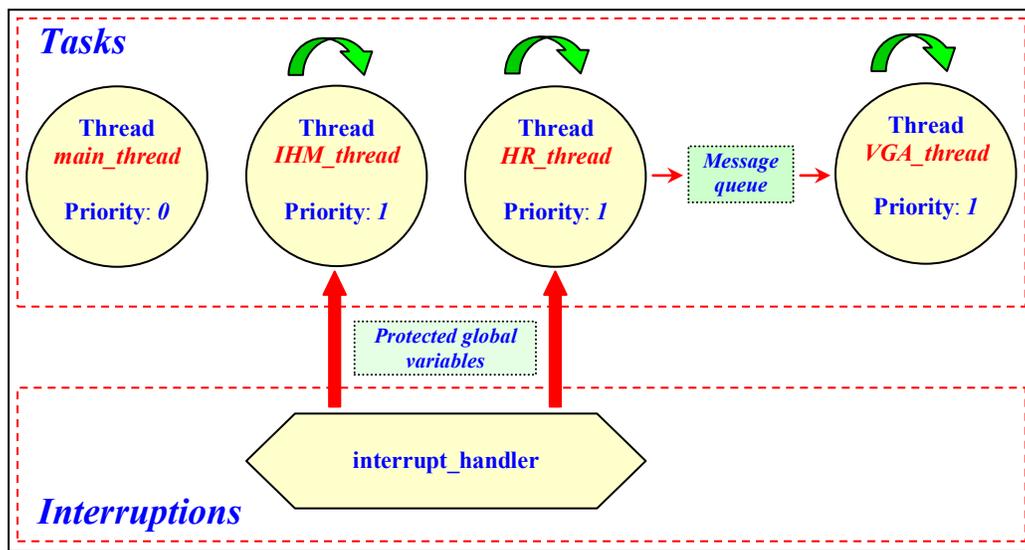


Fig. 4. Application RTOS structure.

V. EVALUATION

As described above, the system is implemented in reconfigurable circuit, that is, FPGA representing the strong actual trend for embedded systems designing. Our system is developed in the Xilinx EDK suite that enables design a complete embedded processor system for implementation in a Xilinx FPGA device. This evaluation focuses on the resources estimation and the QRS detection accuracy:

- Based on the same tests realized in our previous work [18], the total accuracy of the QRS Detector exceeds 96%.
- The implementation occupies around 70% of the used FPGA device, namely the Xilinx Spartan 6 XC6SLX16, which is the core of the development board, Nexys 3 of Digilent.

For testing our work, an excerpt of a MIT-BIH record, with 10000 samples (50 s), is saved in a text file, read from a simple Visual Basic application and presented to Nexys 3 board, via the parallel port, i.e. LPT1 of a PC, driving a simple digital to analog converter (DAC), provided by an R-2R resistor ladder network, at the sampling period mentioned above (5 ms). By convenience, the data is quantized with 8 bits. The figure 5 shows such an experimental setup. On another hand, the PC, via USB port emulating an UART link, receives from the Nexys 3 board the QRS flag and the HRV signals.

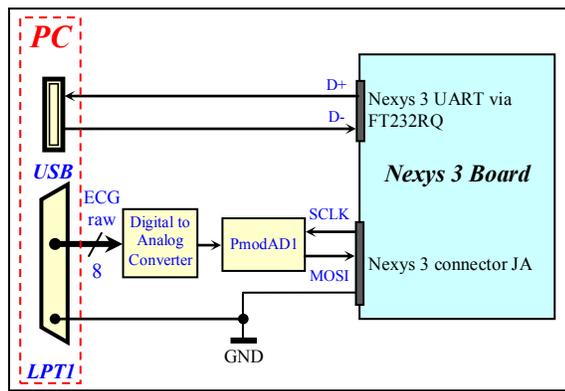


Fig. 5. The Design test via LPT1 of a PC

The figure 6 illustrates the global result of the experience with a screenshot, showing the ECG signal sent to the board (bottom signal), the QRS flag received from the board (top signal) and the HRV tachogram (middle signal).

The figure 7 presents a captured photo of the implementation running the program, with HR equal to 66 bpm (beats per minute) and a blinking LED (LD0 in the board) at measured cardiac rhythm.

The figure 8 illustrates the display result on VGA screen; it presents some resolution problems; it seems that the resolution doesn't exceed (640x100); we think that this problem is related to some dysfunction in TFT IP core; we are trying to find the responsible bug.

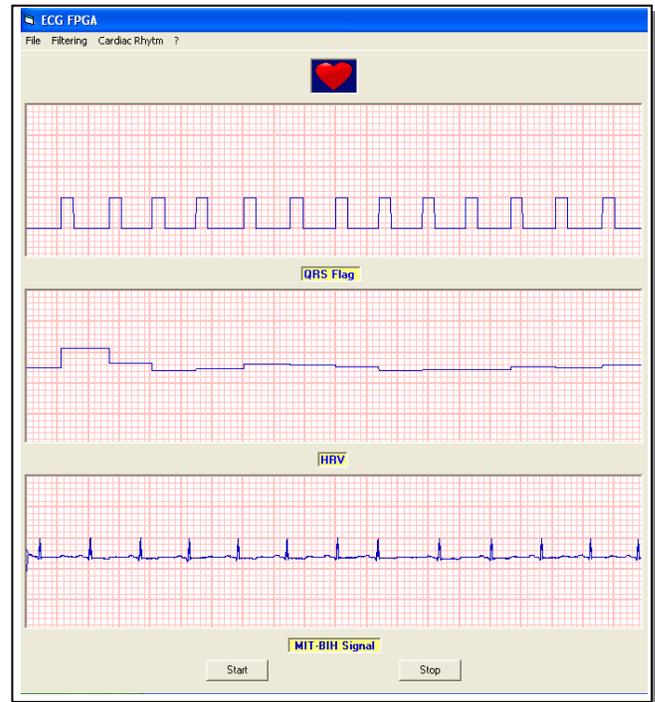


Fig. 6. Screenshot of the test with LPT1 of a PC.

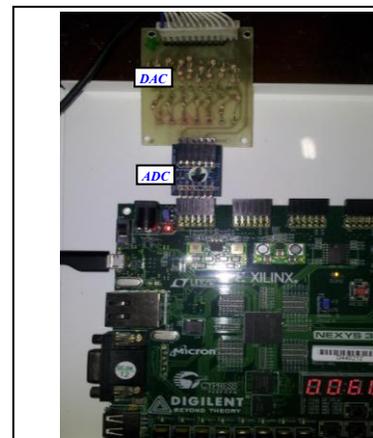


Fig. 7. Photo of the design with running program.



Fig. 8. ECG display on VGA screen.

VI. CONCLUSIONS

We have designed and implemented a soft processor MicroBlaze-based embedded system for cardiac monitoring; it implements the most popular algorithm and widely adopted by the patient monitoring industry, that is, the Pan and Tompkins QRS detector algorithm, which is based on slope, amplitude and width information. It was developed with EDK suite tools with C code and in the Xilkernel RTOS environment. The accuracy of the QRS detection exceeds 96%. In terms of the resources utilization, our implementation occupies around 70% of the used FPGA, namely the Xilinx Spartan 6 XC6SLX16. Thus, there is still space for additional functionalities. The Hardware/Software co-design and RTOS were compatible with the parallelism natively permitted by the FPGA; so, the implementation parts are operated in parallel and executed concurrently, making possible a real-time and multitask processing.

Based on this successful basic implementation, the ongoing work has for purpose to try optimization and implementing more functionalities in such low cost FPGA devices, particularly an algorithm for myocardial Ischemia detection; there is also plan to use multiprocessor design, if needed, for more determinism, which satisfies the real-time constraints.

REFERENCES

- [1] J. O. Hamblen, T. S. Hall, and M. D. Furman, *Rapid Prototyping of Digital Systems*, SOPC edition, Springer, 2008.
- [2] Zainalabedin Navabi, *Digital Design and Implementation with Field Programmable Devices*, Springer, 2005.
- [3] Xilinx System Generator for DSP V12.2, User Guide , July 2010.
- [4] J. cong, B. Liu, J. Noguera, K. Vissers, and Z. Zhang, "High-level synthesis for FPGAs: from prototyping to deployment," in *IEEE Transactions on computer-aided design of integrated circuits and systems*, Vol. 30, No 4, April 2011.
- [5] D. I. Fotiadis, A. Likas, L. Michalis and C. Papaloukas, "Electrocardiogram (ECG): Automated Diagnosis", *Wiley encyclopedia of biomedical engineering*, Vol. 2, pp. 1259-1275, ISBN 978-0-471-24967-2 (set), Wiley, 2006.
- [6] L. Sörnmo and P. Laguna, *Bioelectrical SignalProcessing in Cardiac and Neurological Applications*, Elsevier (Academic Press), 2005.
- [7] L. Sörnmo and P. Laguna, "Electrocardiogram (ECG) Signal Processing", *Wiley encyclopedia of biomedical engineering*, Vol. 2, pp. 1298-1313, ISBN 978-0-471-24967-2 (set), Wiley, 2006.
- [8] W. J. Tompkins, *Biomedical Digital SignalProcessing: C-Language Examples and Laboratory Experiments for the IBM PC*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [9] Rajendra Acharya U, A.E. Spaan, Jasjit Suri, and Shankar M. Krishnan, *Advances in Cardiac Signal Processing*, Springer, 2007.
- [10] <http://www.physionet.org/cgi-bin/atm/ATM>.
- [11] E. El Mimouni, and M. Karim, "Design and implementation of an embedded system for ambulatory cardiac monitoring," in *Journal of Telecommunications and Information Technology (JTIT)*, issue n°1 p. 23-32, Mars 2011.
- [12] J. G. Webster, *Medical Instrumentation Application and Design*, Wiley, 2010.
- [13] http://www.xilinx.com/support/documentation/sw_manuals/mb_ref_guide.pdf.
- [14] J. Teich, "Hardware/Software codesign: the past, the present, and predicting the Future," in *Proceedings of the IEEE*, Vol. 100, May 13th, 2012.
- [15] http://www.xilinx.com/support/documentation/sw_manuals/xilinx13_1/edk_ctt.pdf.
- [16] <http://digilentinc.com/>
- [17] http://www.xilinx.com/ise/embedded/edk91i_docs/xilkernel_v3_00_a.pdf.
- [18] E. El Mimouni, M. Karim, "Novel real-time FPGA-based QRS detector using adaptive threshold with the previous smallest peak of ECG signal," in *Journal of Theoretical and Applied Information Technology (JTIT)*, pp 33-43, Vol. 50 No.1, 10th April 2013.
- [19] J. Pan, and W. J. Tompkins, "A real-time QRS detection algorithm," in *IEEE Trans. Biomed. Eng.*, Vol. BME-32, pp. 230-236, 1985.

Recommender System for Personalised Wellness Therapy

Thean Pheng Lim

School of Computer Sciences
Universiti Sains Malaysia
11800 Penang, Malaysia

Wahidah Husain

School of Computer Sciences
Universiti Sains Malaysia
11800 Penang, Malaysia

Nasriah Zakaria

School of Computer Sciences
Universiti Sains Malaysia
11800 Penang, Malaysia

Abstract—rising costs and risks in health care have shifted the preference of individuals from health treatment to disease prevention. This prevention treatment is known as wellness. In recent years, the Internet has become a popular place for wellness-conscious users to search for wellness-related information and solutions. As the user community becomes more wellness conscious, service improvement is needed to help users find relevant personalised wellness solutions. Due to rapid development in the wellness market, users value convenient access to wellness services. Most wellness websites reflect common health informatics approaches; these amount to more than 70,000 sites worldwide. Thus, the wellness industry should improve its Internet services in order to provide better and more convenient customer service. This paper discusses the development of a wellness recommender system that would help users find and adapt suitable personalised wellness therapy treatments based on their individual needs. This paper introduces new approaches that enhance the convenience and quality of wellness information delivery on the Internet. The wellness recommendation task is performed using an Artificial Intelligence technique of hybrid case-based reasoning (HCBR). HCBR solves users' current wellness problems by applying solutions from similar cases in the past. From the evaluation results for our prototype wellness recommendation system, we conclude that wellness consultants are using consistent wellness knowledge to recommend solutions for sample wellness cases generated through an online consultation form. Thus, the proposed model can be integrated into wellness websites to enable users to search for suitable personalized wellness therapy treatment based on their health condition.

Keywords—recommender system; rule-based reasoning; case-based reasoning; Wellness

I. INTRODUCTION

The concept of health has shifted from treatment of disease to prevention of health problems [1, 2]. As the online community has become more health conscious, improvement is needed to ensure accessibility, reliability and quality of wellness services on the Internet. Many people are turning to wellness websites for work-life balance and healthy lifestyle programmes to fulfil their wellness needs.

Research shows that there are more than 70,000 health informatics websites [3] and one recent study, found more than 54,700,000 results in Google for the term "health assessment tools" [4]. This excess of information has created complications in searching for wellness solutions on wellness websites.

A study on the reliability of health information provided by websites shows inconsistent recommendations, for example in managing fever in children [5]. Inconsistent recommendations negatively affect the quality of health information on the web [6]. It can lead to actual or perceived untrustworthiness of web content, so that users have to navigate and filter to find credible information [7]. The user experiences wellness information overload, which causes uncertainty, confusion and distraction [8]. Furthermore, it is a challenge for users to select a suitable wellness therapy on the Internet since they are lacking in wellness and health understanding. For example, with limited knowledge about health, human anatomy and medical terminology, users may inaccurately describe their symptoms which may cause errors in diagnosis. Most wellness websites do not provide support or consultation for users of their wellness recommendations. Disclaimers on wellness websites frequently indicate that the solutions provided are not for wellness recommendation purposes [9, 10, 11]. Information and guidance provided on such websites are not verified by qualified wellness practitioners, and users may be confused if they follow the suggestions offered. Wellness websites should state clearly what users can get from their web services rather than frustrating them.

The purpose of this study is to design a wellness recommendation model that will suggest personalised wellness therapy or treatments to online users. The objective of the study is to find the best technique for matching users' wellness concerns with appropriate wellness therapy. This model will help to ensure the reliability of wellness recommendations proposed to users based on their current wellness conditions and constraints. At the end of the study, a field evaluation survey was carried out to ensure the reliability and suitability of the wellness solutions proposed by the recommender system.

II. BACKGROUND STUDY

This section reviews two recommender systems that are commonly used in e-commerce transactions: content-based filtering and collaborative filtering. The section also reviews two popular knowledge-based techniques that are currently implemented in the health care environment: case-based reasoning and rule-based reasoning.

A. Recommender Systems

In the modern knowledge era, the extensive number of applications available on the Internet make searching increasingly convenient for users.

This is due to the introduction of recommender systems that filter unseen information, predict the preferences and needs of users [12], and make suggestions to them. Recommender systems are widely used in e-commerce to search for product information and assist customers in deciding what to buy [13]. A recommender system helps users navigate through a large information space to selective descriptions of items that they need [14]. Two common types of recommender system are content-based filtering and collaborative filtering [15].

Content-based filtering is based on similarity of content, such as previous successful transactions [16]. The general principle of the content-based approach is to identify common characteristics of a user's past choices and then recommend new items that share these characteristics [17]. Content-based recommender systems are a type of classifier system, and are related to machine learning research [14]. A user's profile in the system learns from feedback and responses provided by the user [18]. A content-based recommender system compares a user's profile of past selections with the information stored in the database, which is sorted according to similarity and ranked based on the user's known preferences. The content-based approach has several shortcomings. Most importantly, the system can only recommend based on a user's previous ratings; the system cannot recommend new items unless they are similar to previously-liked items [14]. The content-based recommender system is thus not ideal if the user is new to the system, because very little information is available to be compared. In such circumstances, it will negatively impact the effectiveness of the system's recommendations to the user.

On the other hand, collaborative filtering recommends items based on aggregated user preferences, which does not depend on similarity of item descriptions [19]. A collaborative recommender system recommends solutions based on what other similar users have liked. Users with similar preferences are grouped together and are called neighbours [20]. Amazon.com uses collaborative filtering to recommend books to its customers based on books that other similar customers have said that they liked [20]. Collaborative filtering also suffers from a few limitations. There is a significant delay in the rating process because recommendations are made based on preferences of similar users. If there are very few users, it will be harder to find groups of similar ones. Also, if the solution or item has not been rated by similar users, or the solution is very new, the recommender system will not propose the solution because there are not enough ratings to support the recommendation [21].

B. Knowledge-based systems

John McCarthy defined Artificial Intelligence (AI) as "the science and engineering of making intelligent machines." [22]. It is related to the activities of computers in understanding human intelligence. The central problems of AI include reasoning, learning, knowledge and communication, all of which are common and valuable to most industries [23]. Expert systems are a branch of AI that applies reasoning methodologies and domain-specific knowledge to make recommendations, just like a human expert would [24]. To enable this quick and reliable decision-making process, human experts' knowledge is converted to a knowledge-based system

which can be queried for assistance. In the medical context, a physician can diagnose and suggest treatments for an illness despite the ambiguity of symptoms and wide range of medical problem faced by different individuals [25]. Therefore, a physician must use several different types of reasoning. In AI, the most frequently used knowledge-based reasoning techniques are rule-based reasoning (RBR) and case-based reasoning (CBR).

RBR uses "if-then-else" rule statements [26]. Rules are patterns, so the RBR engine searches for patterns in the rules that match patterns in the data. Problem solving becomes more complex if there are too many rules to match the pattern of data in the database [26]. The RBR system uses rule chaining and a combination of data and the system's justification capability to provide a solution to the users [27]. However, RBR lacks the ability to learn due to the difficulty of acquiring new expertise in pattern matching or new rules [28]. RBR also requires a user to take into account all the domain rules; in real life problem solving, the pressures of time restrict problem solvers from looking into large unmanageable rule sets to solve a problem [29]. Thus, RBR is an ideal approach for solving simple problems where not many rules exist.

The CBR approach is similar to human problem-solving behaviour. In CBR, the problem is solved based on experience gained by solving similar problems in the past [30]. CBR is a proven methodology that applies past solutions to solve new but similar problems [31]. The CBR cycle consists of four steps: retrieving previous cases from the case database, reusing a previous case to recommend a solution, revising a previous solution to match the current problem, and retaining new cases after a solution has been successfully found [32].

CBR is a useful alternative to RBR. In complex RBR systems, it is difficult to formulate situations with hard and fast rules due to the complexity of converting problem-solving knowledge due to incomplete problem specifications [30]. However, in CBR, if new knowledge is not available but solutions can be derived from old cases, then on the basis of past experiences, the problem can be solved [30]. CBR is not restricted only to reuse of past cases. It also has good learning capabilities, and its problem solving skills improve as new cases are solved and stored in the database [26]. In our preliminary research, we propose a CBR recommender model to solve users' wellness needs, as illustrated in Fig. 1.

When a new case is entered, the CBR system retrieves similar cases from its case database. The system reuses a solution from a previous case, if necessary adapting it for the new case. The proposed solution is revised to confirm the validity and reliability of the solution. Finally, the revised case will be stored in the case database so it can be used for solving new problems in the future.

III. PROPOSED RECOMMENDER MODEL

This section discusses the characteristics of the proposed hybrid case-based reasoning recommender system for wellness, and walks through a simulation of the recommender system using sample cases, of a paragraph.

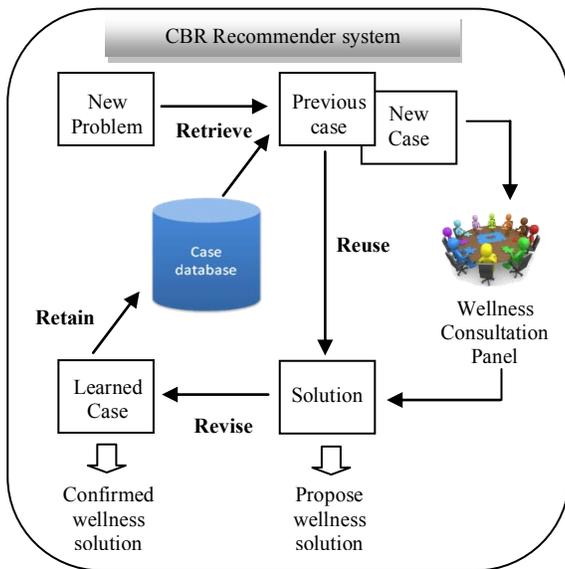


Fig. 1. Wellness recommender model using CBR [33]

Table 1 is a comparison of the advantages and disadvantages of RBR and CBR techniques.

TABLE I. COMPARISON OF RBR AND CBR

	RBR	CBR
Advantages	<ul style="list-style-type: none"> • Each rule can be interpreted as one unit of knowledge. • Knowledge is expressed in the same format. • Rules are a natural format to express knowledge. 	<ul style="list-style-type: none"> • Easy to acquire knowledge from human experts. • Provides model of learning. • Can distinguish different problems and select appropriate cases as solution.
Dis-advantages	<ul style="list-style-type: none"> • Difficult to represent informal knowledge. • Rules obtained from human experts tend to be highly heuristic in nature. • Relies only on the rules and does not learn. • Poor efficacy due to repetition of previous errors. 	<ul style="list-style-type: none"> • Large case database leads to high search cost. • Difficult to determine good criteria in indexing cases and matching similar cases.

A. The hybrid matching technique

In the preceding section, we discussed two types of knowledge-based systems: rule-based reasoning (RBR) and case-based reasoning (CBR). RBR and CBR each have their respective pros and cons.

A combination of CBR and RBR will benefit the system wherein RBR helps to speed the filtering process of CBR [30, 34]. The task of indexing cases in the case database is also handled efficiently by the RBR system [26, 35]. This complements the limitation of CBR to return too many similar cases when there are too many criteria for problem solving in a large case database [26]. Standard rules are acquired from wellness experts and can be applied to solve new problems [36].

Hybrid case-based reasoning (HCBR) is the best approach for a wellness recommender system [27, 29, 36]. HCBR uses a combination of CBR and RBR approaches. HCBR incorporates the advantages of RBR into a sub-system which helps to standardize the format of rules to be used in indexing and searching for similar cases. The panel of wellness experts can contribute to storing standard rules in the RBR sub-system even when there is no similar case in the case database. The RBR sub-system proposes solutions based on standard rules stated in the rule database. The CBR component of HCBR has the ability to learn from previous and new cases, which compensates for the drawback of the RBR system. Therefore, HCBR performs better than RBR alone in terms of accuracy of solution [36]. The matching task is performed using an artificial intelligence (AI) approach employing case-based reasoning [37]. Each case is a module of knowledge that contains structured information about a wellness problem and the appropriate therapy. A case is triggered by matching its relevant components to those in the problem submitted by users. Fig. 2 shows the process flow of the proposed HCBR wellness recommender system.

In making a wellness recommendation, HCBR will consider three alternatives when a new wellness problem is entered. First, the system checks for the same wellness case in the database. If it is present, the solution from that previous case is used to solve the current wellness problem. If no identical wellness case exists in the case database, the system attempts to match users' wellness concerns with similar cases in the database. Similar cases are sorted using rules, as indicated in the RBR sub-system; the system uses these rules to calculate the level of similarity among cases. If the value of the most-similar case is above the acceptance threshold, the solution of that case is proposed to the user. Finally, if there is no case whose similarity level exceeds the acceptance threshold, the system triggers the RBR sub-system to apply standard rules predetermined by the wellness experts.

B. Simulation of Wellness Recommender Prototype

Prototyping begins with the design of the wellness recommendation module. In this study, the purpose of the system prototype is to capture the intended design and simulate the appearance, process and surface texture of the wellness recommendation module. The user requirements for the wellness recommender prototype were gathered from in-depth interview sessions with wellness experts through the Delphi interview technique. Twelve experts from different areas of the wellness industry were interviewed, including spa and relaxation, reflexology, beauty and slimming, skin and body care, and fitness. After attaining consensus on the user requirements, the wellness recommender prototype was developed.

Three sample cases are stored in the recommender system case database for simulation and evaluation purposes. The three sample cases are shown in Fig. 3.

To complete the simulation process, we created a new case to represent a current problem faced by a user. This new case is shown in Fig. 4.

NEW CASE		SAMPLE 1
Age: 27	← 0.8 →	Age: 36
Sex: Male	← 1.0 →	Sex: Male
Health Status: High blood pressure, Depression	← 0.9 →	Health Status: High blood pressure
Personal lifestyle: Ability to relax (poor), Smoke (often), Stress level (Work 10)	← 0.9 →	Personal lifestyle: Ability to relax (poor), Stress level (Work 10)
Wellness concerns: Stress, Anxiety	← 0.9 →	Wellness concerns: Stress

NEW CASE		SAMPLE CASE 2
Age: 27	← 1.0 →	Age: 27
Sex: Male	← 0 →	Sex: Female
Health Status: High blood pressure, Depression	← 0.7 →	Health Status: Insomnia
Personal lifestyle: Ability to relax (poor), Smoke (often), Stress level (Work 10)	← 0.6 →	Personal lifestyle: Average sleeping hours (3), Alcohol (often)
Wellness concerns: Stress, Anxiety	← 0.7 →	Wellness concerns: Fatigue

NEW CASE		SAMPLE CASE 3
Age: 27	← 0.8 →	Age: 35
Sex: Male	← 0 →	Sex: Female
Health Status: High blood pressure, Depression	← 0.9 →	Health Status: High Blood Pressure
Personal lifestyle: Ability to relax (poor), Smoke (often), Stress level (Work 10)	← 0.2 →	Personal lifestyle: Skin Type (Dry), Exercise (Never)
Wellness concerns: Stress, Anxiety	← 0 →	Wellness concerns: Dry Body (Arm, Back, Chest)

← - - - → Low importance ↔ High importance

Fig. 5. Comparison of new case with Sample Cases 1, 2 and 3

Similarity (New Case, Sample Case 1)

$$= 1/17 * [(1*0.8) + (1*1.0) + (5*0.9) + (5*0.9) + (5*0.9)]$$

$$= 1/17 * (0.6 + 1.0 + 4.5 + 4.5 + 4.5)$$

$$= \mathbf{0.89}$$

Similarity (New Case, Sample Case 2)

$$= 1/17 * [(1*1.0) + (1*0) + (5*0.7) + (5*0.6) + (5*0.7)]$$

$$= 1/17 * (1.0 + 0 + 3.5 + 3.0 + 3.5)$$

$$= \mathbf{0.65}$$

Similarity (New Case, Sample Case 3)

$$= 1/17 * [(1*0.8) + (1*0) + (5*0.9) + (5*0.2) + (5*0)]$$

$$= 1/17 * (0.8 + 0 + 4.5 + 1.0 + 0)$$

$$= \mathbf{0.37}$$

The results of the similarity computation determine the most-similar case, whose solution will be adopted as the solution to the new case. Based on the similarity calculations above, the system will choose Sample 1 as the solution for the new case (0.89 > 0.65 > 0.37). Note that in this simulation, only selection of attributes is considered; in the actual system, other features such as previous wellness record would also be considered. Fig. 6 shows the wellness therapy proposed by the recommender system after analysing the new case. The system has returned a solution with a match of 89%, as computed above.

The recommendation shows that the user may choose from six types of wellness therapy treatment, which have been grouped into three categories of massage, reflexology and fitness therapies.

To match your current wellness concern and needs, I-Wellness recommends you the following therapies:

Major wellness concern: **Stress and Fatigue**

Value of case acceptance level: **89 %**

Wellness Therapy Recommendation :

Massage therapy	
Swedish Massage	Swedish massage uses firm but gentle pressure to promote relaxation, ease muscle tension and create other health benefits. Massage therapists use massage oils to facilitate smooth, gliding strokes called effleurage. Classic Swedish massage moves include kneading, friction, stretching and (sometimes) tapping
Aromatherapy Massage	Aromatherapy is the ancient science of healing, relaxing and energizing by the use of plants and their parts. Essential oils form the basis of aromatherapy. These oils can be inhaled, massaged onto your body, added to the bath or shower or sprayed in the room. This helps is alleviating stress and speeding up healing.
Reflexology therapy	
Neck & Shoulder / Back & Shoulder Therapeutic Massage	Therapeutic Massages concentrating on neck, shoulder and back areas help to release the tension and relieve pain, discomfort, or stiffness, giving you a new lease on life
Hand and Foot Reflexology	Hand and foot reflexology is the practice of stimulating specific nerves on the hands or feet, to encourage positive effects on the corresponding parts of the body, and to improve general health and well-being. The treatment helps in relieving stress & reduce tiredness, relax muscles & stimulate nerve connections and induce deep relaxation
Fitness therapy	
Body Balance	Body balance is a fitness course that helps in strengthening your body, improving health and shaping your body. This class will leave you feeling refreshed and renewed with a sense of wellbeing and calmness.
PRABHU Yoga Therapy	This class will improve your strength towards the prevention and cure of body imbalance, headache, neck and shoulder pain, migraine, stress and stomach disorders.

Fig. 6. Results computed from the wellness recommender system

C. Evaluation of wellness recommender system

At the end of the study an evaluation survey was carried out to verify the accuracy of the recommendations proposed for the new case in the simulation. In the evaluation phase, a Likert scale was used to measure the responses from wellness consultants and therapists regarding the prototype's wellness therapy recommendations. A sample size of 40 respondents,

10 from each of four wellness categories (namely, spa and relaxation, reflexology, beauty and slimming and fitness), were randomly selected to answer a simple questionnaire. Completed questionnaires were collected for further analysis.

In the questionnaire, wellness consultants and therapists were asked to express their agreement or disagreement with the wellness recommendations proposed by the prototype for the new case. Rowe and Wright [39] stated that “sensible questions are only sensible if they relate to the domain of knowledge of the specific experts”. Therefore, wellness consultants and therapists are the right people to verify the reliability of the proposed wellness recommender system. First, solutions for Sample Cases 1, 2 and 3 were provided by the wellness experts.

These solutions were then compared and applied to those proposed by the prototype for the new case problem. Table 2 shows the descriptive analysis of the evaluation results for Sample Case 1 solutions adapted for the new problem while Table 3 and Table 4 show the analysis of evaluation results for Sample Case 2 and 3 solutions respectively adapted for the new problem.

TABLE II. DESCRIPTIVE ANALYSIS OF EVALUATION RESULTS (SOLUTION FOR NEW CASE BASED ON SOLUTIONS IN SAMPLE CASE 1)

Wellness therapy	Mean	Standard Deviation
Swedish massage	4.1	0.568
Aromatherapy Massage	4.0	0.667
Hand and foot reflexology	4.0	0.816
Body Balance	4.0	0.471
Neck & shoulder / Back & shoulder Therapeutic Massage	3.7	0.675
PRABHU Yoga	3.7	0.675
Body scrub & massage	3.1	0.316
Swedish Delight	2.8	0.422

TABLE III. DESCRIPTIVE ANALYSIS OF EVALUATION RESULTS (SOLUTION FOR NEW CASE BASED ON SOLUTIONS IN SAMPLE CASE 2)

Wellness therapy	Mean	Standard Deviation
Hand & foot reflexology	4.0	0.816
Foot reflexology	3.5	0.527
Head, shoulder & back massage	2.7	0.483
Body scrub and massage	2.6	0.516
Aerobic	2.6	0.516
Basic yoga	2.6	0.516
Malay traditional massage	2.4	0.516

TABLE IV. DESCRIPTIVE ANALYSIS OF EVALUATION RESULTS (SOLUTION FOR NEW CASE BASED ON SOLUTIONS IN SAMPLE CASE 3)

Wellness therapy	Mean	Standard Deviation
Traditional Massage	2.0	0.471
Traditional papaya facial	1.8	0.422
Silky body treatment	1.3	0.483
Body Nourishing	1.3	0.483
Steam bath	1.2	0.422

From the descriptive analysis, we can see that most of the mean in Table 2 is above 3.0. The rating result from evaluation

is higher than the Likert five-point average (average=3.0). By referring to Fig. 6 above, we see that the level of acceptance in comparing the new case with sample case 1 is 89%. This means that there is an 89% chance that the solution in sample case 1 can be adapted to solve the problem posed in the new case. However, wellness consultants from beauty and slimming disagreed with Swedish Delight therapy as a solution in the new case. This indicates that a higher level of acceptance is needed in order to propose accurate solutions for problems in the beauty and slimming category.

On the other hand, most of the mean in Table 3 is below 3.0. The rating result from evaluation here is lower than the Likert five-point average (average=3.0). Referring to the similarity calculation, we can see that the level of acceptance in comparing the new case with sample case 2 is 65%. This means that the solution for sample case 2 is not appropriate for solving the problem posed in the new case. However, wellness consultants in the area of reflexology agreed that similar cases can be used even though the level of acceptance is 65%.

In Table 4, most of the mean is below 2.0. This indicates that wellness consultants agree that the solution in sample case 3 cannot be used to solve the problem posed in the new case. This is confirmed by the low similarity calculation of the new case with sample case 2, which is only 37%.

IV. CONCLUSION

This study proposed that online wellness consultation can be carried out using HCBR in a wellness recommender system. The purpose of employing HCBR in wellness recommendation is to capture experts' knowledge and experience together with previously resolved wellness cases in a case database. Users present their wellness concerns, which are solved by locating an identical or similar case in the case database (CBR). If the case database contains no sufficiently similar cases, the system will recommend a suitable wellness solution by using pre-determined standard rules (RBR).

The present study concentrated on non-medical wellness therapies that fall into five disciplines, namely: spa and relaxation, reflexology, beauty and slimming, skin and body care, and fitness. Future work should focus on the study of medical wellness recommendation, such as acupuncture, anti-aging, medical massage and other therapies, in order to provide a holistic solution in the wellness market. In conclusion, HCBR can be effectively applied in an online wellness recommender system to match a user's current wellness concerns with suitable wellness therapies.

ACKNOWLEDGMENT

The authors would like to thank the Universiti Sains Malaysia, Penang for supporting this study.

REFERENCES

- [1] Omar, A., Wahlqvist, M. L., Kouris-Blazos, A., & Vicziany, M., "Wellness management through Web-based programmes," (2005) Journal of Telemedicine and Telecare, vol. 11 (1), pp. 8-11.
- [2] Stahl, V., "Wellness and sustainability," (2010) [internet] Available at http://www.cognis.com/NR/rdonlyres/3BE09E03-8E04-49FD-B56F-E1E8EB6D7A85/0/05_CC_Worldwidetrends_e_Sept07_ffcf.pdf [Accessed 21 August 2012].

- [3] Grandinetti, D. A., "Doctors and the web: Help your patients surf the Net safely," (2000) *Medical Economics*, vol. 77 (5), pp. 186.
- [4] Susan, D., "Ramifications of web-based health assessment tools on consumer behaviours and health outcomes," (2010) Available at: http://susandorfman.com/uploads/DHA_Dissertation_susan_dorfman_fe_b_27_2010_Final.pdf [Accessed 21 August 2012].
- [5] Impicciatore, P., Pandolfini, C., Casella, N., & Bonati, M., "Reliability of health information for the public on the World Wide Web: systematic survey of advice on managing fever in children at home," (1997) *BMJ*, 314 (1875).
- [6] Purcell, G. P., Wilson, P., & Delamothe, T., "The Quality of health information on the Internet," (2002) *BMJ*, 324.
- [7] Eysenbach, G., & Thomson, M., "The FA4CT Algorithm: A New Model and Tool for Consumers to Assess and Filter Health information on the Internet," (2007) Kuhn K (ed.) *Medinfo 2007 Proceedings* (in press) [Internet] Available at: <https://tspace.library.utoronto.ca/handle/1807/9795> [Accessed 10 March 2013].
- [8] Beller, S., "Information overload and Health Decision-Making (Part 1)," (2006) [Internet] Available at: <http://curinghealthcare.blogspot.com/2006/09/information-overload-and-health.html> [Accessed 11 March 2013].
- [9] Fitness Wellness Guide, "Fitness Wellness Disclaimer," (2008) [Internet] Available at: <http://www.fitness-wellness-guide.com/disclaimer.html> [Accessed 10 March 2013].
- [10] Whitaker Wellness Institute, "Disclaimer," (2009) [Internet]. Available at: <http://www.whitakerwellness.com/about-us/disclaimer/> [Accessed 10 March 2013].
- [11] Wellness Layer, "DietWatch Personal Terms of Use," (2009) [Internet] Available at: http://www.dietwatch.com/dietwatch/shared/about_us.asp?view=Terms%20of%20Use [Accessed 10 March 2013].
- [12] Long, F., Zhang, Y., & Hu, F., "A Personalised Recommender System Framework based on Ontology and Bayesian Network in E-commerce," (2011) *Advanced Materials Research*, vol. 143-144, pp. 961-965.
- [13] Chen, S., "An Electronic Commerce Recommender System Based on Product Character," (2011) *Advanced Materials Research*, vol. 267, pp. 909-912.
- [14] Burke, R., "Knowledge-based recommender systems," (2000) *Encyclopaedia of Library and Information Science*.
- [15] Kim, H-N., et al., "Collaborative user modelling for enhanced content filtering in recommender systems," (2011) *Decision Support Systems*, doi:10.1016/j.dss.2011.01.012.
- [16] Maidel, V., Shoval, P., Shapira, B., & Maimon, M., "Ontological content-based filtering for personalised newspapers," (2010) *Online Information Review*, vol. 34, No. 5, pp. 729-756.
- [17] Desrosiers, C., & Karypis, G., "A Comprehensive Survey of Neighbourhood-based Recommendation Methods," (2011) *Recommender Systems Handbook Part 1*, pp. 107-144.
- [18] Meteren, R. V., & Someren, M. V., "Using content-based filtering for recommendation," (2000) [Internet] Available at: http://www.ics.forth.gr/~potamias/mlnia/paper_6.pdf [Accessed 13 March 2013].
- [19] Cai, X., et al., "Collaborative filtering for people to people recommendation in social networks," (2010) *Advances in Artificial Intelligence*, vol. 6464, pp. 476-485.
- [20] Sarwar, B. M., Karypis, G., Konstan, J. A., & Riedl, J. T., "Application of Dimensionality Reduction in Recommender System – A Case Study," (2000) [Internet] Available at: <http://robotics.stanford.edu/~ronnyk/WEBKDD2000/papers/sarwar.pdf> [Accessed 13 February 2013].
- [21] Smyth, B., & Cotter, P., "A Personalized Television Listings Service," (2000) *Communication of the ACM*, vol. 43 (8).
- [22] McCarthy, J., "What is Artificial Intelligence?" (2007) [Online] Stanford University Available at: <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html> [Accessed 11 February 2013].
- [23] Luger, G. F., "Artificial Intelligence," (2002) Fourth Edition. Addison-Wesley, Pearson Education.
- [24] Abu-Naser, S.S., El-Hissi, H., & El-khozondar, N., "An expert system for endocrine diagnosis and treatments using JESS," (2010) *Journal of Artificial Intelligence*, vol. 3(4) pp. 239-251.
- [25] Eshach, H., & Bitterman, H., "From Case-based Reasoning to Problem-based Learning," (2003) *Academic Medicine*, vol. 78 (5).
- [26] Wang, W. M., Cheung, C. F., Lee, W. B., & Kwok, S. K., "Knowledge-based treatment planning for adolescent early intervention of mental healthcare: a hybrid case-based reasoning approach," (2007) *The Journal of Knowledge Engineering*, vol. 24 (4), pp. 232-251.
- [27] Lim, T. P., "I-Wellness: Application of hybrid case-based reasoning for wellness therapy recommendation in wellness community portal," (2010) *Universiti Sains Malaysia*.
- [28] Lee, M. R., "An exception handling of rule-based reasoning using case-based reasoning," (2002) *Journal of intelligent and Robotic System*, Vol. 35, pp. 327-338.
- [29] Lim, T.P., & Husain, W., "I-Wellness: A hybrid Case-based Framework for personalised wellness therapy," (2010) *Proceeding of International Symposium on Information Technology 2010 (ITSIM'10)*, vol. 3, pp. 1193-1198.
- [30] Jaglan, V., Dalal, S., & Srinivasan, S., "Improving performance of business intelligence through case based reasoning," (2011), *International Journal of Engineering Science and Technology (IJEST)*, vol. 3, No. 4.
- [31] Branden, M., Wiratunga, N., Burton, D., & Craw, S., "Integrating case-based reasoning with an electronic patient record system," (2011), *Artificial Intelligence in Medicine*, vol. 51 (2), pp. 117-123.
- [32] Schmidt, R., & Gierl, L., "Case-based reasoning for antibiotics therapy advice: an investigation of retrieval algorithms and prototypes," (2001) *Artificial Intelligence in Medicine*, Vol. 23 (2), pp. 171-186.
- [33] Lim, T.P., & Husain, W., "Integrating knowledge-based system in Wellness Community Portal," (2010) *Proceeding of International Conference on Scientific & Social Science Research (CSSR 2010)*. [Internet] Available at: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5773798 [Accessed 21 August 2013].
- [34] Golding, A. R., & Rosenbloom, P. S., "Improving Rule-Based Systems through Case-Based Reasoning," (1991) In *Proceedings of AAAI*, Anaheim, CA, pp. 22-27.
- [35] Chan, F. T. S., "Application of a hybrid case-based reasoning approach in electroplating industry," (2005) *Expert Systems with Application*, vol. 29, pp. 121-130.
- [36] Lim, T.P., & Husain, W., "The Development of Personalized Wellness Therapy Recommender System Using Hybrid Case-Based Reasoning," (2010) *Proceeding of 2nd International Conference on Computer Technology and Development (ICCTD 2010)*. Available at: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5646071 [Accessed 22 August 2013].
- [37] Ravana, S. D., Abdul Rahman, S. S., & Chan, H. Y., "Web-Based Diet Information System with Case-Based Reasoning Capabilities," (2006) *International Journal of Web information Systems*, vol 2 (3/4), pp. 154-163.
- [38] Burke, R., "Hybrid Recommender Systems: Survey and Experiments," (2002) *User Modelling and User-Adapted Interaction*, vol. 12, pp. 331-370.
- [39] Rowe, G., & Wright, G., "The Delphi techniques as a forecasting tool: issues and analysis," (1999) *International Journal of Forecasting*, 15, pp. 353-375

A Survey of Network-On-Chip Tools

Ahmed Ben Achballah

Dept. of Electrical Engineering
National Institute of Applied Sciences and Technology
Polytechnic School of Tunisia, LSA laboratory
Centre Urbain Nord, BP 676- 1080 Tunis Cedex, Tunisia.

Slim Ben Saoud

Dept. of Electrical Engineering
National Institute of Applied Sciences and Technology
Polytechnic School of Tunisia, LSA laboratory
Centre Urbain Nord, BP 676- 1080 Tunis Cedex, Tunisia.

Abstract—Nowadays System-On-Chips (SoCs) have evolved considerably in term of performances, reliability and integration capacity. The last advantage has induced the growth of the number of cores or Intellectual Properties (IPs) in a same chip. Unfortunately, this important number of IPs has caused a new issue which is the intra-communication between the elements of a same chip. To resolve this problem, a new paradigm has been introduced which is the Network-On-Chip (NoC). Since the introduction of the NoC paradigm in the last decade, new methodologies and approaches have been presented by research community and many of them have been adopted by industrials. The literature contains many relevant studies and surveys discussing NoC proposals and contributions. However, few of them have discussed or proposed a comparative study of NoC tools. The objective of this work is to establish a reliable survey about available design, simulation or implementation NoC tools. We collected an important amount of information and characteristics about NoC dedicated tools that we will present throughout this survey. This study is built around a respectable amount of references and we hope it will help scientists.

Keywords—*Embedded Systems; Network-On-Chip; CAD Tools; Performance Analysis; Verification and Measurement*

I. INTRODUCTION

The insatiable market demands for more innovative technologies have induced a considerable evolution of the integration capacities in recent platforms. In fact, Semiconductor industrials have offered, are offering and will continue to offer many powerful hardware chips. Gates scaling continue to fall down (40nm, 35nm and recently 28nm), also power consumption is decreasing and GHz working frequencies are increasing [1-2]. A chip with the last cited advantages will enlarge the intervention domain of engineers and many design issues can be solved because the computing power and the chip flexibility are enormous.

Like the hardware side of technology, the software side which is represented by the Computer Aided Design (CAD) tools was dramatically innovated. This includes modeling, simulation, synthesis and implementation tools. Also, new design flows have emerged like the CoDesign concept. Moreover many techniques were proposed by research community and some of them were adopted by industrials like High Level synthesis (HLS) or Model Based Design (MBD). All of these enhancements and novelist techniques share common purposes i) To increase the abstraction level of a desired design flow ii) To furnish preventive estimations for engineers at earlier stages of the design to ensure low cost fixes iii) To accelerate the design flow.

Traditionally, a SoC is composed by some processing elements (processors, dedicated Intellectual Properties (IPs), etc), few memory blocks and In/Out communication modules. Nowadays, the number of these On-Chip elements is extremely growing: This is a direct result of the innovations and advancements cited earlier [1, 3]. Recent platforms are often Multi-Processors SoC (MPSoC) with multiple functionalities and a lot of options. For example we can cite recent personal computers, video games, smart phones and tablets. However, the growth of the On-Chip elements has provoked new issues like the communication between internal elements. In fact, classical buses could not assure a reliable connection between them. A new solution has to be found to face this problem. In 2002, the NoC paradigm has been introduced by Luca Bennini and Giovanni De Micheli [4]. This proposal has resolved the intra-communication problem and data exchange.

The NoC paradigm was important because it allowed design engineers to follow technology advancements and so, integrating many cores at the same chip by overcoming the intra-communication problems. For this purpose, many studies were conducted and many NoC architectures were proposed and later some of them were enhanced. We can find in the literature some relevant surveys and comparative studies between NoC proposals. Reference [5] details the NoC concept and discusses some examples. Other references also discusses this subject in many aspects by proposing a detailed comparison between NoC architectures and performances or by exposing the future of NoC related researches [6-8]. In our case, we will present the NoC concept to give lecturers an overview about it, this will be the subject of the second section. As we said before, this case study is focused on establishing a study about NoC dedicated tools. We will present our findings in this subject respectively in section 3 and 4. Finally, we review related works in section 5 and we conclude the work and expose perspectives in section 6.

II. THE NOC CONCEPT

In this section we will introduce the NoC concept and later we will present some of their principal characteristics. At the end of this section we will show the research axes and the problems facing the research community when developing NoCs followed by some common NoC architecture proposals. In purpose to show the importance of NoC in recent SoCs, we decided to begin this section by a comparison between classical buses and Network-On-Chips. Table 1 gives a qualitative comparison between conventional buses and NoCs. As we can see this table demonstrates the usefulness of NoC

based systems especially in case of the SoC contains many processing elements.

TABLE I. QUALITATIVE COMPARISON BETWEEN BUSES AND NOCS

Buses			NoCs
Each item will add a parasitic capacitance, hence the degradation of electrical performance increase with the number.	-	+	The elements are connected by point to point interconnections for all sizes of networks, local performance is not degraded.
The time management of the bus is difficult.	-	+	The data transfer can be accomplished by delayed transitions because the connections are point to point.
Delays caused by arbitration at the bus can cause blockages, especially if the number of masters is important.	-	+	Routing decisions are distributed.
The bandwidth is limited and shared by all elements of the bus.	-	+	The bandwidth increases with the size of the network.
The bus tests are long and problematic.	-	+	The dedicated BIST (Built In Self Test) are locals, complete and fast.
The bus latency is the speed of a connection control circuit if the bus is granted.	+	-	Internal decisions making can add delays.
The concept is simple and easy to understand.	+	-	Designers need upgrades in order to exploit new concepts.

A. NoC Architecture

The NoCs consist typically of routers, network adapter (network interface) and connections [6, 9].

1) Router: directs the data according to the protocol selected. It contains the routing strategy.

2) Network Adapters: provide a bridge between the router and the element attached to them. Their main task is to separate calculation (IPs) of the communication (network). This consists of two operations which are protocol conversion and packages construction.

3) Connections: are the channels of transmission of data between the various circuit elements to the network.

B. Topology

The topology of a network is the way in which routers, network adapters and connections are organized. There are several topologies that we can call regular or irregular [10-11]. This classification is based on the distribution of routers in the network. Figure 1 shows some regular topologies we can find a) mesh b) mesh torus c) ring d) fat-tree.

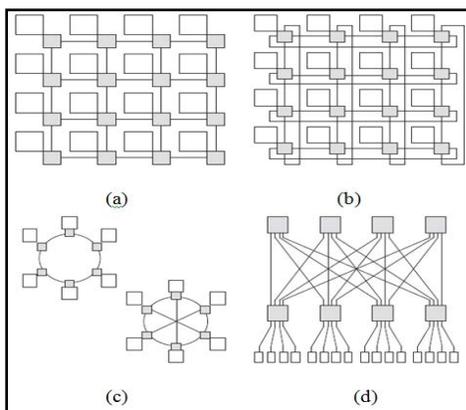


Fig. 1. Examples of regular NoC topologies

In the other side, irregular topologies are composed of two or three regular topologies such as a mesh topology and ring simultaneously.

C. Routing

Routing is to transfer data from source to destination with a clearly defined strategy. In the literature, researchers have classified the routing algorithms according to different criteria:

1) The type routing is called source routing if only the sender provides the path by which the data will flow, it is called distributed if the transit decision is taken locally at each node. We can also find a classification similar to the previous one except that it defines a more general routing strategy regardless of source, in effect if the routing decisions are identically distributed across the network, routing is called centralized. If these decisions are taken locally, routing is still called distributed [5]. As we can see that decision does not take into account the sender as the one before.

2) The routing is deterministic if the transit path is determined by the sender and the receiver only. The path between the same network correspondents is invariable. However, if the transit of data between two network elements can be achieved through multiple paths, routing is then called adaptive. This is possible thanks to the decisions taken locally at the nodes. The implementation of adaptive routing algorithms can generate complicated nodes but can ensure a better flow of data within a NoC.

3) The routing is called circuit switching when a circuit (a path) between the transmitter and receiver is reserved for the duration required to transfer data. It is called packet routing when the data to be transmitted is divided into packets containing a portion of the data and routing information. The packets may follow different paths to reach their destination.

A routing algorithm usually has one or more of the characteristics mentioned a little earlier. For example, an adaptive routing is generally a packet switching routing.

D. Switching techniques

The primary function of switches is to determine when and how the inputs of a router will be connected to its outputs [12]. There are several switching techniques among them store-and-forward, virtual cut-through and wormhole.

1) *Store-and-forward*: the transferred data is split into packets and each packet contains routing information. When a packet reaches a node, it is entirely saved in a buffer and routing information is extracted to determine the appropriate output port.

2) *Virtual cut-through*: the routing information is contained in the first bytes of the packet. Instead of saving the entire package like store-and-forward, the packages are sent as soon as the output port is determined. In case this port is used, the package will be saved in a buffer [6, 9].

3) *Wormhole*: the packets are split into sub-packets called flits (Flow Control Unit). The control data are contained in the header flit. As a result, a single packet can be transmitted by different nodes. This will reduce latency, but may cause many bottlenecks in the network.

E. Related research axes and issues

Scientific Research for NoCs is conducted on several axes. They can be classified into three broad categories or levels: networks, interconnection or system. In what follows we will describe these different orientations.

1) Network Level

The research at the network level is the most solicited level between scientific community. This is because NoC were in developing phases. The most discussed topics are the following:

- a) *Topology*: Regular, non-regular or mixed.
- b) *Protocol*: routing, switching
- c) *Flow control of data*: Anti-blockage mechanism, virtual channels, buffering.
- d) *Quality of Service (QoS)*: throughput, latency

2) Connection Level

Interconnection level research can be considered as a direct consequence of that done at the network level. Since the interconnections are used to join the network adapters to routers and also routers between them, a non-optimized communication can affect network performances. So it is also important that interconnections have to be studied. The items concerned are:

- a) *Synchronization*
- b) *Parallel vs. serial*
- c) *Reliability*
- d) *Pipeline*

3) System level

Given the advances in research for NoC and especially at the architectural level (network), the current proposed NoCs are increasingly complicated. Not only conceptually but also at simulation and testing phases. Currently, we can find in the literature more than sixty proposals with very different configurations [6, 9]. This includes the topology, routing,

switch mode, the implementation technology and even the method of simulation and evaluation. This mixture, added to an exponentially growing complexity of architecture, has prompted researchers to turn to new design methodologies. Methodologies where the level of abstraction is raised to the system level in order to facilitate the designers work. The research for NoCs at the system level is summarized in the following:

- a) *Design methodology*: modeling, co-design.
- b) *Evaluation and assessment of performance*: latency, throughput, power consumption and space.
- c) *Architectures*: system-level composition, reconfigurable NoCs.

The advancements obtained from the researches at the system level have conducted to the development of many CAD tools dedicated for NoCs. These tools will allow the management of very complex NoCs architectures throughout the design flow (modeling, simulation and implementation). A second advantage came from the fact they are especially designed for NoCs and not for general use, thereby ensuring more relevant results. Before we develop this topic in details later in paragraph II and IV, we will present some NoCs proposals in what follows.

F. Some NoCs proposals

In this section we present examples of NoC architectures from the literature. We will restrict ourselves to some examples because the objective behind this research is not the architecture proposals but NoCs tools used on their development. They are many works that have focused on the collection, classification and characterization of different architectures and implementations of NoCs to date. For a more complete list of NoCs proposals, we advice lecturers to consult these references [5-9].

1) *ÆTHREAL*

This Network-On-Chip was developed by Philips Research Laboratories. This NoC offers QoS for data transfer within a SoC, such as a) No loss b) No corruption c) Organized Transfer Order and so the transfer rates are guaranteed and the latency is predictable [13-14].

2) *SPIN*

The SPIN architecture has been developed by the University "Pierre et Marie Curie" [10]. The main characteristics of this NoC are a) expanded tree topology b) Routing packet.

3) *QnoC*

This NoC is developed by the Israeli Institute of Technology [15]. It is based on a mesh topology that can be irregular and the wormhole as switching technique.

III. NOC DEDICATED TOOLS

Recent SoCs typically contain a relatively complicated architecture with a large number of computing elements [1, 16]. This requires a NoC based design to ensure an optimum management of the transit of internal data [3, 9]. To facilitate the development of embedded systems containing a network on chip, several dedicated tools have been proposed. These

initiatives are often presented by the scientific community through research teams. Nevertheless, there are some other proposals from the industry.

The dedicated NoC tools vary depending on the purpose for which they are developed. We can distinguish two main classes: synthesizers and simulators. Regarding synthesizers, points often discussed are the quality of generated architectures (space, energy consumption) and the level of abstraction for modeling NoCs, the higher the level of abstraction is, the higher the design and its correction are fast. Recent compilers are becoming more powerful and it exists some commercial versions like FlexNoc from Arteries [17], INOC [18] and The Tool Suite Works CHAIN from Silistix [19-20]. In the other side, two criteria are often addressed for simulators: the estimation of power dissipation and performance computing (throughput, latency, and reliability). These two criteria are crucial since NoC are an integral component of an embedded system. Because these systems are often subject to hard constraints of space, energy and execution time, a relevant estimation of the NoCs characteristics could be very helpful to the designer. We will show in the following NoCs synthesis or simulation tools we collected from the literature. We recall that this list is not exhaustive.

A. NS-2

NS-2 was first developed for prototyping and simulating ordinary computer networks. However, since NoCs shares many characteristics with classic networks, NS-2 was widely used by many NoC researchers to simulate NoCs [21-22]. Many NoC studies have used NS-2 as a simulation tool making it a reliable reference especially when comparing the performances of two different architectures [23-25]. Finally, NS-2 is an open source, discrete event driven simulator and developed in C++ and OTcl. These modularity and availability have facilitated its spreading between researchers.

B. Noxim

This tool has been proposed by the Computer Architecture team at the University of Catania [26]. It is developed in SystemC language. It allows the user to define a 2D mesh NoC architecture with various parameters including: 1) Network size 2) Buffers size 3) Packet size 4) Routing algorithm 5) Injection rate of packets. Noxim allows the evaluation of NoCs in terms of throughput, latency and power consumption.

C. DARSIM

DARSIM is a NoC simulator which was developed at the Massachusetts Institute of Technology (MIT). This tool allows the simulation of mesh NoC architectures of 2 and 3 dimensions. It offers a multitude of NoC simulation configurations with various parameters. This includes two generation modes of data generation:

a) *Trace-driven injection which involves the injection of packets into the network and monitors their spatial and temporal evolution. Each injection contains the tracing parameters that are time constraints, the identifier of the stream, the packet size and possibly the injection frequency.*

b) *MIPS Simulation mode: each node can be configured as a MIPS (Microprocessor Without Interlocked Pipeline Stages) with its own memory. These are connected to the NoC MIPS and receiving/sending data from/to the network can be simulated cycle by cycle. Using these methods and a more detailed explanation of the simulator are presented in this reference [27].*

D. SunFloor - 3D SunFloor

SunFloor is a support tool for NoC design. It can be used at earlier design phases to synthesize the most appropriate topology with these constraints as input (Model, Energy and Space, Design Objectives). From these data, SunFloor generates a system specification ready to be translated into comprehensive architecture, usually in SystemC language and by the intervention of a second tool which is xpipesCompiler [28-29].

SunFloor 3D is an extension of the later version. The main feature added is the generation of specifications for the future 3D Wafers [30]. Both versions were developed by the team of Prof. Giovanni De Micheli, a pioneer of research for NoCs with many publications in the NoC subject like a particular article [4] with over 1900 citations (Google Scholar statistics) and several books on this subject [31].

E. ORION 2.0

ORION 2.0 is the successor of the version proposed by a team from Princeton University in 2003 [32]. It is a simulator dedicated primarily to the estimation of power and space for NoCs architectures. Among the improvements compared to the first version we find the support for new semiconductor technology through models of transistors and capacitances upgraded from industry [33-34].

F. NoC Emulation techniques

There are other techniques for simulation and the On-Chip verification of NoCs like the emulation technique proposed by [35]. This technique allows the emulation of NoC architectures such *ÆTHREAL* or those generated by xpipesCompiler through a standard platform. This involves interfacing IPs capable of injecting or retrieving data to and from the emulated NoC [36].

G. Other available NoC tools

Several other tools dedicated to NoCs are also available. The following table summarizes the tools we collected in the literature. We stress at the fact that list is not exhaustive.

TABLE II. NOC TOOLS PROPOSALS

#	Tool	Year	Team	References
1	NS-2	1995	DARPA and later Contributors	[22]
2	Noxim	2010	Catagne University	[26]
3	DARSIM	2009	MIT	[27]
4	SunFloor – 3D	2006 - 09	EPFL (switzerland)	[28-30]
5	ORION 1 et 2	2003 - 09	Princeton University	[32-34]
6	INSEE	2005	Basque University (Spain)	[37]
7	ATLAS	2005	Federal University of Brazil	[38]
8	NOCIC	2004	Massachussetts University	[39]
9	Pestanna Environment	2004	Phillips Research Laboratories	[40]
10	PIRATE	2004	Polytechnique School of Milan	[41]
11	SUNMAP	2004	Stanford University	[42]
12	xpipesCompiler	2004	Bologne University – Stanford University	[43]
13	μSpider	2004	Bretagne Sud University	[44]
14	OCCN	2004	ST microelectronics	[45]
15	NoCGEN	2004	University of New South Wales	[46]
16	FlexNoC	-	ARTERIS	[17]
17	iNoC	-	-	[18]
18	The CHAIN works tool suite	-	Silistix	[19]

IV. NOC EVALUATION METHODS

NoC Evaluation is an important step to classify the proposed architecture among the others. There are 3 general criteria or metrics that are considered by the majority of the research community: i) area consumption ii) power consumption and iii) latency. There is other metrics which were reported in some other works like packet loss or wire length. The authors of reference [6] reported that in average 3 metrics are discussed in any NoC proposal, which is inadequate says the author. In our study the NoC tools we found are globally focused on these three metrics cited above but some of them may have extra auxiliary options and configurations. Table 3 lists the NoC tool we collected by their characteristics. It includes modeling, simulation, hardware synthesis and availability. Besides, the modeling process includes many options which are: the network size, the buffers size, the packets distribution, the routing algorithm, the packet injection ration, the selection strategy and finally the traffic distribution. The simulation process contains: the area consumption, the power consumption, the network throughput and the latency.

V. RELATED WORKS

Many works were done in the NoC area since their appearance. The major recent proposals are essentially based on more sophisticated architectures offering diverse advantages among them the quality of service (QoS) [12] or

globally asynchronous locally synchronous (GALS) architectures which resolve the clocking difference problems inside a SoC [47-48].

Technology advancements have also pushed researchers to reconsider their point of view about NoCs. Besides, some works have focused on developing 3D NoC architectures [49] and as we have seen in section 3, NoC tools developers have also anticipated these advancements by proposing tools that are dedicated for 3D NoC design and simulation [30]. However, other studies proposed different approach by adding the NoC concept to the bus one and so, keeping some data transfer to classical buses. The objective is often to reduce coasts in term of area and power consumption and of course without degrading the system performances in terms of throughput and latency [50].

Other researchers have applied an existing concept which is basically developed for SoCs to the Network-On-Chip one like the reconfigurability. The term of ReNoCs which means Reconfigurable NoCs is more and more developed inside the scientific community and as a result some initiatives were elaborated on this subject [51].

VI. FINAL REMARKS AND CONCLUSION

In this survey we tried to focus on a subject concerning NoCs that somehow was not deeply studied in the literature. In this paper we presented the NoC concept and its importance in recent SoCs.

Then we presented the tools dedicated to their development which includes the modeling, simulation and implementation processes. We stress again at the fact that this list is not exhaustive but can represent an important number of

nowadays available NoC tools. Meanwhile, when we are writing this manuscript some other tools emerged and that we don't hesitate to include like MCoreSim [52] or a flexible parallel simulator with error control [53].

TABLE III. NOC TOOLS CHARACTERISTICS

#	Tool	Specification												
		Modeling							Simulation				Hardware Synthesis	Availability
		NS	BS	PD	RA	SS	PIR	TD	AC	PC	T	L		
1	NS-2	+	+	+	+	+	+	+	+	+	+	+	-	+
2	Noxim	+	+	+	+	+	+	+	-	+	+	+	-	+
3	DARSIM	+	+	+	+	+	+	+	-	-	+	+	-	-
4	SunFloor – 3D	+	-	-	-	-	-	-	-	+	+	+	+	-
5	ORION 2.0	-	+	-	-	-	-	-	+	+	-	-	-	-
6	ATLAS	+	-	+	+	-	+	-	-	-	+	+	+	+
7	PIRATE	+	+	-	-	-	-	-	+	+	-	-	-	-
8	SUNMAP	+	+	-	-	-	-	-	-	+	+	+	+	-
9	µSpider	+	+	-	-	-	-	-	-	-	+	+	+	-
10	NoCGEN	-	+	-	-	-	-	-	-	-	+	+	+	-
11	FlexNoC	+	+	+	+	+	+	+	+	+	+	+	+	commercial
12	iNoC	+	+	+	+	+	+	+	+	+	+	+	+	commercial
13	The CHAIN works tool suite	+	+	+	+	+	+	+	+	+	+	+	+	commercial

TABLE IV. NOMENCLATURE

Modeling	Simulation
NS: Network Size	AC: Area Consumption
BS: Buffers Size	PC: Power Consumption
PD: Packets Distribution	T: throughput
RA: Routing Algorithm	L: Latency
PIR: Packets Injection Ratio	
SS: Selection Strategy	
TD: Traffic Distribution	

REFERENCES

[1] S. Vangal, et al., "An 80-Tile Sub-100-W TeraFLOPS Processor in 65-nm CMOS," IEEE Journal of Solid-State Circuits, pp. 29 - 41, 2008.

[2] S. Chen, et al. (2010), Xilinx Next Generation 28 nm FPGA Technology Overview, available at www.xilinx.com/support/documentation/white_papers/wp312_Next_Gen_28_nm_Overview.pdf.

[3] S. Borkar, "Thousand core chips: a technology perspective," in Design Automation Conference, 2007, pp. 746 - 749.

[4] L. Benini and G. D. Micheli, "Networks on Chips: A New SoC Paradigm," Computer, pp. 70 - 78, 2002.

[5] T. Bjerregaard and S. Mahadevan, "A Survey of Research and Practices of Network-on-Chip," ACM Computing Surveys, vol. 38, 2006.

[6] E. Salminen, et al., "Survey of Network-on-chip Proposals," OCP-IP White paper, 2008.

[7] E. Salminen, et al., "On network-on-chip comparison," in 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools, 2007, pp. 503 - 510.

[8] A. Agarwal, et al., "Survey of Network on Chip (NoC) Architectures & Contributions," Journal of Engineering, Computing and Architecture, 2009.

[9] G. D. Micheli, et al., "Networks on Chips: from Research to Products," in Design Automation Conference, 2010, pp. 300 - 305.

[10] P. Guerrier and A. Greiner, "A Generic Architecture for On-chip Packet-switched Interconnections," in Proceedings of the conference on Design, Automation and Test in Europe, 2000, pp. 250 - 256.

[11] A. S. Vaidya, et al., "Impact of Virtual Channels and Adaptive Routing on Application Performance," IEEE Transactions on Parallel and Distributed Systems, pp. 223 - 237, 2001.

[12] M. Moadeli, "Quarc: An Architecture For Efficient On-Chip Communication," PhD Thesis, University of Glasgow, 2010.

[13] K. Goossens and A. Hansson, "The Æthereal Network on Chip after Ten Years: Goals, Evolution, Lessons, and Future," in Design Automation Conference, 2010.

[14] K. Goossens, et al., "Æthereal Network on Chip: Concepts, Architectures, and Implementations," IEEE Design & Test of Computers, pp. 414 - 421, 2005.

[15] E. Bolotin, et al., "QNoC: QoS architecture and design process for network on chip," Journal of Systems Architecture, pp. 105-128, 2004.

[16] Tilera website. Available: www.tilera.com/products/processors

[17] Arteris website. Available: www.arteris.com/flex_noc.php

[18] iNoCs website. Available: www.inocs.com

[19] Silistix website. Available: www.silistix.com

[20] F. Gebali, et al., Networks-on-Chips: Theory and Practice, CRC Press ed., 2009.

[21] M. Ali, et al., "Using the NS-2 Network Simulator for Evaluating Network on Chips (NoC)," in International Conference on Emerging Technologies, 2006, pp. 506 - 512.

[22] NS-2 website. Available: nsnam.isi.edu/nsnam/index.php/Main_Page

[23] R. Kourdy, et al., "Using the NS-2 Network Simulator for Evaluating Multi Protocol Label Switching in Network-on-Chip," in 2nd

- International Conference on Computer Research and Development, 2010, pp. 795 - 799.
- [24] S. Vangalt, et al., "A 5.1 GHz 0.34 mm² Router for Network-on-Chip Applications," in IEEE Symposium on VLSI Circuits, 2007, pp. 42 - 43.
- [25] R. Lemaire, et al., "Performance evaluation of a NoC-based design for MC-CDMA telecommunications using NS-2," in The 16th IEEE International Workshop on Rapid System Prototyping, 2005, pp. 24 - 30.
- [26] Noxim website. Available: www.noxim.org
- [27] M. Lis, et al., "DARSIM: a parallel cycle-level NoC simulator," in 6th Annual Workshop on Modeling, Benchmarking and Simulation, 2010.
- [28] A. Pullini, et al., "NoC Design and Implementation in 65nm Technology," in ACM/IEEE International Symposium on Networks-on-Chip, 2007, pp. 273-282.
- [29] S. Murali, et al., "Designing Application-Specific Networks on Chips with Floorplan Information," in IEEE/ACM International Conference on Computer Aided Design, 2006, pp. 355 - 362.
- [30] C. Seiculescu, et al., "SunFloor 3D: A Tool for Networks on Chip Topology Synthesis for 3D Systems on Chips," in Proceedings of the conference on Design, Automation and Test in Europe, 2009, pp. 9 - 14.
- [31] G. D. Micheli and L. Benini, Networks on Chips: Technology and Tools, Morgan Kaufmann ed., 2006.
- [32] H. Wang, et al., "Orion: A Power-Performance Simulator for Interconnection Networks," in 35th Annual IEEE/ACM International Symposium on Microarchitecture, 2002, pp. 294 - 305.
- [33] A. Kahng, et al., "ORION 2.0: A Fast and Accurate NoC Power and Area Model for Early-Stage Design Space Exploration " in Proceedings of Design Automation and Test in Europe, 2009.
- [34] Orion tool website. Available: www.princeton.edu/~peh/orion.html
- [35] N. Genko, et al., "A Novel Approach for Network on Chip Emulation," in IEEE International Symposium on Circuits and Systems, 2005, pp. 2365 - 2368.
- [36] N. Genko, et al., "NoC Emulation: A Tool and Design Flow for MPSOC," IEEE Circuits and Systems Magazine, pp. 42 - 51, 2007.
- [37] F. J. R. Perez and J. Miguel-Alonso, "INSEE: An Interconnection Network Simulation and Evaluation Environment " in 11th International Euro-Par Conference, 2005, pp. 1014 - 1023.
- [38] ATLAS website. Available: <https://corfu.pucrs.br/redmine/projects/atlas/wiki>
- [39] V. Venkatraman, et al., "NoCIC: A Spice-based Interconnect Planning Tool Emphasizing Aggressive OnChip Interconnect Circuit Methods," in Proceedings of the 2004 international workshop on System level interconnect prediction, 2004, pp. 69 - 75.
- [40] S. G. Pestana, et al., "Cost-Performance Trade-offs in Networks on Chip: A Simulation-Based Approach," in Proceedings of the conference on Design, Automation and Test in Europe, 2004.
- [41] G. Palermo and C. Silvano, "PIRATE: A Framework for Power/Performance Exploration of Network-on-Chip Architectures," in 14th International Workshop Power and Timing Modeling, Optimization and Simulation, 2004, pp. 521-531.
- [42] S. Murali and G. D. Micheli, "SUNMAP: A Tool for Automatic Topology Selection and Generation for NoCs," in ACM/IEEE Design Automation Conference, 2004.
- [43] A. Jalabert, et al., "×pipesCompiler: A tool for instantiating application specific Networks on Chip," in Proceedings of the conference on Design, Automation and Test in Europe, 2004, pp. 1530-1591.
- [44] S. Evain, et al., "µSpider: a CAD Tool for Efficient NoC Design," in Proceedings of the Norchip Conference, 2004, pp. 218 - 221.
- [45] M. Coppola, et al., "OCCN: A Network-On-Chip Modeling and Simulation Framework," in Proceedings of the conference on Design, Automation and Test in Europe, 2004.
- [46] J. Chan and S. Parameswaran, "NoCGEN: A Template Based Reuse Methodology for Networks on Chip Architecture," in IEEE 17th International Conference on VLSI Design, 2004, pp. 717 - 720.
- [47] A. Rahimi, et al., "Energy/Throughput Trade-off in a Fully Asynchronous NoC for GALS-Based MPSoC Architectures," in IEEE International Conference on Design & Technology of Integrated Systems in Nanoscale Era, 2010.
- [48] G. Campobello, et al., "GALS Networks on Chip: A New Solution for Asynchronous Delay-Insensitive Links," in Proceedings of the conference on Design, automation and test in Europe, 2006, pp. 160 - 165.
- [49] Y. Xie, et al., "Three-Dimensional Network-on-Chip Architecture," in Three Dimensional Integrated Circuit Design, Y. Xie, et al., Eds., Springer US ed, 2009, pp. 189-217.
- [50] P. Zarkesh-Ha, et al., "Hybrid network on chip (HNoC): local buses with a global mesh architecture," in Proceedings of the ACM/IEEE international workshop on System level interconnect prediction, 2010, pp. 9-14.
- [51] Y. E. Krasteva, et al., "Reconfigurable Networks on Chip: DRNoC architecture," Journal of Systems Architecture, pp. 293-302, 2010.
- [52] S. Kumar, et al., "A Latency Simulator for Many-core Systems," in 44th Annual Simulation Symposium, 2011.
- [53] Q. Yu and P. Ampadu, "A Flexible Parallel Simulator for Networks-on-Chip with Error Control," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, pp. 103 - 116, 2010.

Construction of Neural Networks that Do Not Have Critical Points Based on Hierarchical Structure

Tohru Nitta

National Institute of Advanced Industrial Science and Technology (AIST)
Tsukuba, Japan

Abstract—a critical point is a point at which the derivatives of an error function are all zero. It has been shown in the literature that critical points caused by the hierarchical structure of a real-valued neural network (NN) can be local minima or saddle points, although most critical points caused by the hierarchical structure are saddle points in the case of complex-valued neural networks. Several studies have demonstrated that singularity of those kinds has a negative effect on learning dynamics in neural networks. As described in this paper, the decomposition of high-dimensional neural networks into low-dimensional neural networks equivalent to the original neural networks yields neural networks that have no critical point based on the hierarchical structure. Concretely, the following three cases are shown: (a) A 2-2-2 real-valued NN is constructed from a 1-1-1 complex-valued NN. (b) A 4-4-4 real-valued NN is constructed from a 1-1-1 quaternionic NN. (c) A 2-2-2 complex-valued NN is constructed from a 1-1-1 quaternionic NN. Those NNs described above do not suffer from a negative effect by singular points during learning comparatively because they have no critical point based on a hierarchical structure.

Keywords—critical point; singular point; redundancy; complex number; quaternion

I. INTRODUCTION

A neural network is a network composed of neurons, and can be trained to find nonlinear relationships in data. NNs have been studied for many years in the hope of achieving human-like flexibility to process information. The common objective of training of a neural network is to determine the global minimum of an error function. However, learning algorithms for NN such as the back-propagation learning algorithm take a very long time to find the global minimum due to the standstill of learning generally.

If ω_* is the global minimum of error function $E(\omega)$, then $\partial E(\omega_*) / \partial \omega = 0$. Nevertheless, even when $\partial E(\omega_*) / \partial \omega = 0$, ω_* is not necessarily a global minimum. ω_* , a point satisfying $\partial E(\omega_*) / \partial \omega = 0$, is designated as the *critical point* of the error function E . A critical point can be a local minimum, a local maximum, or a saddle point.

Fukumizu et al. mathematically proved the existence of a local minimum resulting from a hierarchical structure in a real-valued NN (ordinary NN handling real-valued signals).

They demonstrated that critical points in a three-layer real-valued NN with $H-1$ hidden neurons behave as critical points in a three-layer real-valued NN with H hidden neurons, and that they are local minima or saddle points. This kind of critical point turns into singular points of a real-valued NN to stagnate training.

A complex-valued NN extends (real-valued) parameters such as weight and threshold values in an ordinary NN to complex numbers. It is suitable for information processing of complex-valued data and two-dimensional data. Moreover, it is applicable to communications, image-processing, biologic information processing, land-mine detection, wind prediction, independent component analysis (ICA), etc. Reportedly, a critical point in a three-layer complex-valued NN also behaves in the same manner as that in a three-layer real-valued NN [1]: critical points in a three-layer complex-valued NN with $H-1$ hidden neurons turn into critical points in a three-layer complex-valued NN with H neurons, which are saddle points (except for cases meeting rare conditions).

Such singular points have been emerging lately as objects of study. Learning models with a hierarchical structure or symmetry of exchange of weights, such as a hierarchical NN and Gaussian mixture model, usually have a singular point. It has been revealed that a singular point affects the training dynamics of a learning model and that it engenders stagnation of training.

This paper presents an attempt to implement an NN having no critical point based on a hierarchical structure.

II. ANALYSIS

In this section, it is demonstrated that NNs having no critical point based on a hierarchical structure can be constructed by decomposing a high-dimensional NN into equivalent lower-dimensional NNs.

A. Construction of a 2-2-2 real-valued NN

A 2-2-2 real-valued NN having no critical point based on a hierarchical structure is constructed from a 1-1-1 complex-valued NN.

Consider a 1-1-1 complex-valued NN (called *NET 1* here). We will use $a+ib \in \mathbf{C}$ for the weight between the input neuron and the hidden neuron, $v+iw \in \mathbf{C}$ for the weight

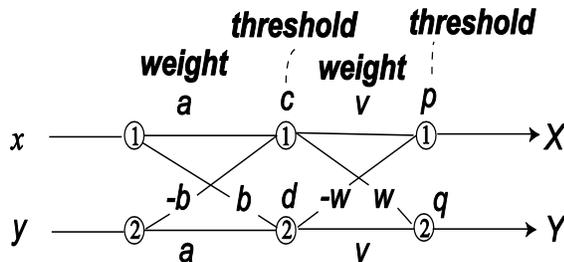


Fig. 1. NET 2, a 2-2-2 real-valued NN equivalent to the NET 1. Restriction applies between weights. $x, y, a, b, c, d, v, w, p, q, X$ and Y are all real numbers.

between the hidden neuron and the output neuron, $c + id \in \mathbf{C}$ for the threshold of the hidden neuron, and $p + iq \in \mathbf{C}$ for the threshold of the output neuron, where i denotes $\sqrt{-1}$ and \mathbf{C} denotes the set of complex numbers. We assume that $a + ib \neq 0$ and $v + iw \neq 0$. Let $x + iy \in \mathbf{C}$ denote the input signal, and let $X + iY \in \mathbf{C}$ denote the output signal. We will use activation functions defined by the following equations:

$$f_c(z) = \tanh(z^R) + i \tanh(z^I), \quad z = z^R + iz^I \in \mathbf{C} \quad (1)$$

for the hidden neuron, and

$$g_o(z) = z, \quad z \in \mathbf{C} \quad (2)$$

for the output neuron. This 1-1-1 complex-valued NN is apparently equivalent to a 2-2-2 real-valued NN (called NET 2 here) shown in Fig. 1.

Proposition 1 NET 2 has no critical point based on a hierarchical structure.

(Proof) Assume a 2-1-2 real-valued NN obtained by removing the hidden neuron 1 from the NET 2 (called NET 3 here) (Fig.2). Also assume that the learning parameter of NET 3 is a critical point that implements mapping $F_1(x, y)$. It is necessary to realize any one of the following three conditions for implementing the same mapping F_1 by appending once-removed hidden neuron 1 to NET 3 again.

1) A weight vector between hidden neuron 1 appended and the two output neurons is $\mathbf{0}$.

$v = w = \mathbf{0}$ must hold in this case, but this violates the assumption $v + iw \neq \mathbf{0}$.

2) A weight vector between hidden neuron 1 appended and the two input neurons is $\mathbf{0}$.

$a = b = \mathbf{0}$ Must hold in this case, which violates the assumption $a + ib \neq \mathbf{0}$.

3) For weight vector \mathbf{w}_1 between hidden neuron 1 appended and the two input neurons, and the weight vector \mathbf{w}_2 between hidden neuron 2 and the two input neurons, $\mathbf{w}_1 = \mathbf{w}_2$ or $\mathbf{w}_1 = -\mathbf{w}_2$.

$a = b = \mathbf{0}$ Must hold in this case, but this violates the assumption $a + ib \neq \mathbf{0}$.

Therefore, mapping F_1 cannot be implemented by NET 3

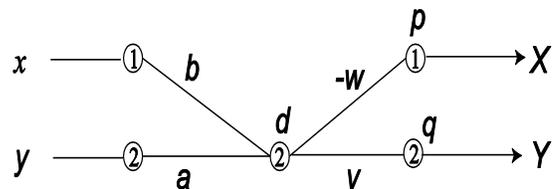


Fig. 2. NET 3, a 2-1-2 real-valued NN obtained by removing hidden neuron 1 from NET 2.

with the original hidden neuron 1 appended and having the weight structure of NET 2.

The description above illustrates a case in which hidden neuron 1 is removed, but removal of the hidden neuron 2 engenders the same conclusion. Consequently, NET 2 has no critical point based on a hierarchical structure. (QED)

See appendix A for the practical implementation process for 2-2-2 real-valued NN having no critical points based on a hierarchical structure.

B. Construction of 4-4-4 real-valued NN

A 4-4-4 real-valued NN having no critical point based on a hierarchical structure is constructed from a 1-1-1 quaternionic NN. The quaternionic NN is an extension of the classical real-valued neural network to quaternions, of which the weights, threshold values, input and output signals are all quaternions, where a quaternion is a four-dimensional number invented by W. R. Hamilton in 1843.

Consider a 1-1-1 quaternionic NN (called NET 4 here). Let the weight between the input neuron and hidden neuron be $A = a + ib + jc + kd \in \mathbf{Q}$, and the weight between a hidden neuron and an output neuron be $B = \alpha + i\beta + j\gamma + k\delta \in \mathbf{Q}$, where \mathbf{Q} represents a set of quaternions. We assume that $A \neq \mathbf{0}$ and $B \neq \mathbf{0}$. Let $C = p + iq + jr + ks \in \mathbf{Q}$ denote the threshold of the hidden neuron, $D = \mu + i\nu + j\rho + k\sigma \in \mathbf{Q}$ represent the threshold of the output neuron, $I = v + iw + jx + ky \in \mathbf{Q}$ be the input signal, and $O = V + iW + jX + kY \in \mathbf{Q}$ be the output signal. We can use the activation functions defined by the following equations:

$$f_o(u) = \tanh(u_1) + i \tanh(u_2) + j \tanh(u_3) + k \tanh(u_4), \quad (3)$$

$$u = u_1 + iu_2 + ju_3 + ku_4 \in \mathbf{Q} \quad (4)$$

For the hidden neuron, and

$$g_o(u) = u, \quad u \in \mathbf{Q} \quad (5)$$

For the output neuron. Because a quaternion is non-commutative for multiplication, the computational result varies with the multiplication sequence of an input value and weight: $IA \neq AI$. Accordingly, quaternion neurons of two kinds exist: a normal quaternary neuron (computing AI) and

an inverse quaternary neuron (computing IA). This paper specifically addresses a quaternionic NN that comprises only inverse quaternary neurons as an example.

In fact, NET 4 is apparently equivalent to a 4-4-4 real-valued NN (called NET 5 here) shown in Fig.3.

Proposition 2 NET 5 has no critical point based on a hierarchical structure.

(Proof) Assume a 4-3-4 real-valued NN obtained by removing the hidden neuron 1 from the NET 5 (called NET 6 here). Also assume that the learning parameter of the NET 6 is a critical point that implements mapping $F_2(v, w, x, y)$. It is necessary to realize any one of the following three conditions for implementation of the same mapping F_2 by appending once-removed hidden neuron 1 to the NET 6 again.

1) A weight vector between hidden neuron 1 appended and the four output neurons is $\mathbf{0}$.

$B = 0$ must hold in this case, but this violates the assumption $B \neq 0$.

2) A weight vector between hidden neuron 1 appended and the four input neurons is $\mathbf{0}$.

$A = 0$ must hold in this case, but this violates the assumption of $A \neq 0$.

3) Letting \mathbf{w}_j denote the weight vector between hidden neuron j and the four input neurons for any $1 \leq j \leq 4$ where the hidden neuron 1 is the appended one, then there exist some $2 \leq j \leq 4$ such that $\mathbf{w}_1 = \mathbf{w}_j$ or $\mathbf{w}_1 = -\mathbf{w}_j$

In this case, $A = 0$ must hold, which violates the assumption of $A \neq 0$.

Therefore, mapping F_2 cannot be implemented by the NET 6 with the original hidden neuron 1 appended and having the weight structure of the NET 5.

The description above presents a case in which hidden neuron 1 is removed, but removal of the hidden neuron j engenders the same conclusion ($2 \leq j \leq 4$). Consequently, NET 5 has no critical point based on a hierarchical structure. (QED)

See appendix B for the practical implementation process used for the 4-4-4 real-valued NN having no critical points based on a hierarchical structure.

C. Construction of 2-2-2 complex-valued NN

A 2-2-2 complex-valued NN having no critical point based on a hierarchical structure is constructed from a 1-1-1 quaternionic NN.

Next we consider a 1-1-1 quaternionic NN (NET 4) defined in Section II-B, assuming that

$$a + ib \neq \pm i(c + id), \quad (6)$$

$$C = 0. \quad (7)$$

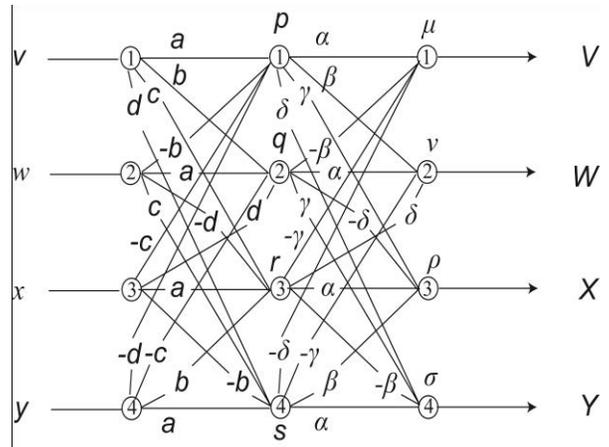


Fig. 3. NET 5, a 4-4-4 real-valued NN equivalent to the NET 4. Restriction applies between weights. $v, w, x, y, a, b, c, d, \alpha, \beta, \gamma, \delta, p, q, r, s, \mu, \nu, \rho, \sigma, V, W, X, Y$ are all real numbers.

We designate this 1-1-1 quaternionic NN as NET 7. Equation (6) has the meaning described below. The weight A between the input neuron and the hidden neuron of the NET 7 can be written using Cayley-Dickson notation as follows.

$$A = a + ib + jc + kd = x_1 + x_2 j \quad (8)$$

Where $x_1 \square a + ib \in \mathbf{C}$ and $x_2 \square c + id \in \mathbf{C}$. Equation (6) can be rewritten as $x_1 \neq \pm ix_2$ from (8). Therefore, if we regard x_1 and x_2 respectively as two vectors, x_1 and x_2 do not intersect orthogonally. Furthermore, if $x_1 = \mp ix_2$, then $A = x_1 \pm x_2 j = x_1(1 \pm k)$ holds. That is, the weight A has information related only to x_1 . Consequently, (6) means the exclusion of such a special case.

In addition, (7) means that the threshold value of the hidden neuron is 0, which is necessary for application of the condition for complex-valued NN being reducible.

Cayley-Dickson notation reveals that the NET 7 is equivalent to a 2-2-2 complex-valued NN (called NET 8 here) shown in Fig. 4, where $v' = v + iw, x' = x + iy, a' = a + ib, c' = c + id, \alpha' = \alpha + i\beta, \gamma' = \gamma + i\delta,$

$\mu' = \mu + iv, \rho' = \rho + i\sigma, V' = V + iW, X' = X + iY$. The activation functions are given as (1) and (2).

Proposition 3 NET 8 has no critical point based on a hierarchical structure (as a complex-valued NN).

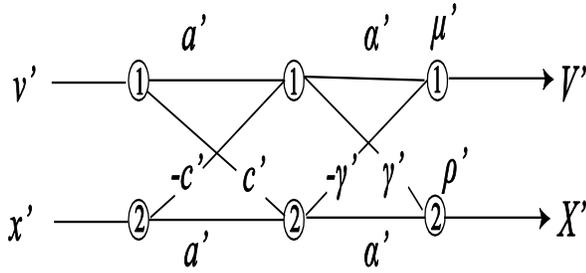


Fig. 4. NET 8, a 2-2-2 complex-valued NN equivalent to a 1-1-1 quaternionic NN (NET 7). Restriction applies between weights. $v', x', a', c', a', \gamma', \mu', \rho', V', X'$ are all complex numbers. The threshold values of the hidden neurons are all omitted because they are 0.

(Proof) Assume a 2-1-2 complex-valued NN obtained by removing the hidden neuron 1 from NET 8 (called NET 9 here). Also assume that the learning parameter of the NET 9 is a critical point that implements complex mapping $G(v', x')$. It is necessary to realize any one of the following three conditions for implementing the same complex mapping G by appending once-removed hidden neuron 1 to the NET 9 again.

1) A weight vector between hidden neuron 1 appended and the two output neurons is $\mathbf{0}$.

$\alpha' = \gamma' = 0$ Must hold in this case, but this violates the assumption $B = \alpha' + \gamma'j \neq 0$.

2) A weight vector between hidden neuron 1 appended and the two input neurons is $\mathbf{0}$.

$a' = c' = 0$ must hold in this case, but this violates the assumption $A = a' + c'j \neq 0$.

3) Let ϕ_1 be the net input to the hidden neuron 1 appended, and ϕ_2 the net input to the hidden neuron 2. Then, ϕ_1 and ϕ_2 are rotation-equivalent: $\phi_1 \equiv \phi_2$ or $\phi_1 \equiv -\phi_2$ or $\phi_1 \equiv i\phi_2$ or $\phi_1 \equiv -i\phi_2$.

Because $\phi_1(v', x') = v'a' - x'c'$ and $\phi_2(v', x') = v'c' + x'a'$,

a) Case of $\phi_1 \equiv \pm\phi_2$

$a' = c' = 0$ must hold in this case, but this violates the assumption $A = a' + c'j \neq 0$.

b) Case of $\phi_1 \equiv \pm i\phi_2$

$a' = \pm ic'$ Must hold in this case, but this violates the assumption $a + ib \neq \pm i(c + id)$ ((6)).

Therefore, the complex mapping G cannot be implemented by the NET 9 with the original hidden neuron 1 appended and having the weight structure of NET 8.

The description above presents case in which hidden neuron 1 is removed, but removal of the hidden neuron 2 engenders the same conclusion. Consequently, NET 8 has no

critical point based on a hierarchical structure (as a complex-valued NN). (QED)

This paper assumes the threshold value of the hidden neuron of a 1-1-1 quaternionic NN to be 0 ((7)). This threshold value is necessary to apply the 'three conditions for a complex-valued NN to be reducible' as described in the proof of Proposition 3. As a result, all threshold values of the hidden neuron of the obtained 2-2-2 complex-valued NN are 0. Considering a 1-1-1 quaternionic NN with possibly non-zero threshold value of a hidden neuron might yield a 2-2-2 complex-valued NN with the possibly non-zero threshold value of a hidden neuron. For a three-layer complex-valued NN with a possibly non-zero threshold value of a hidden neuron to be reducible, exceptional reducibility is necessary in addition to the three conditions presented above [2].

See the appendix C for the practical implementation process of a 2-2-2 complex-valued NN having no critical points based on a hierarchical structure.

III. DISCUSSION

Fukumizu and Amari proved that a critical point of the three-layered real-valued NN with $H-1$ hidden neurons always gives many critical points of the three-layered real-valued NN with H hidden neurons. These critical points can be local minima or saddle points.

Local minima cause plateaus, which have a strong negative influence on learning. Recently, it was proven that most of the local minima that Fukumizu et al. discovered are resolved by extending the real-valued NN to complex numbers; most of the critical points attributable to the hierarchical structure of the complex-valued NN are saddle points, which is a prominent property of the complex-valued NN [1]. That is, there exist many critical points based on a hierarchical structure both in the real-valued NN and the complex-valued NN.

Such critical points can be local minima or saddle points in the real-valued NN, although most critical points of the complex-valued NN are saddle points. However, in both cases, critical points do exist in the networks. As described in this paper, an attempt is made to remove critical points themselves from NNs based on a hierarchical structure.

IV. CONCLUSION

This paper presented a proposal for an implementation process of a NN having no critical point based on a hierarchical structure. Results demonstrate that real-valued and complex-valued NNs having no critical point based on a hierarchical structure can be constructed by decomposing a high-dimensional NN into equivalent real-valued or complex-valued NNs. Concretely, the following three cases are shown:

(a) A 2-2-2 real-valued NN is constructed from a 1-1-1 complex-valued NN.

(b) A 4-4-4 real-valued NN is constructed from a 1-1-1 quaternionic NN.

(c) A 2-2-2 complex-valued NN is constructed from a 1-1-1 quaternionic NN. Those NNs described above do not suffer

from a negative effect by singular points during learning comparatively because they have no critical point based on a hierarchical structure.

The author expects to address the following issues in future studies.

1) Although quaternionic NN that comprise only inverse quaternary neurons are used for this study, the case with normal quaternary neurons shall be considered.

2) General complex-valued NNs with possibly non-zero threshold values of a hidden neuron shall be analyzed, which requires consideration of exceptional reducibility [2].

3) A 2^s -dimensional Clifford NN having no critical point based on a hierarchical structure shall be produced by decomposing a general 2^n -dimensional Clifford NN [3] into equivalent Clifford NNs of 2^s dimensions ($s < n$).

V. APPENDICES

A. Implementation of the 2-2-2 real-valued NN

The practical implementation of the 2-2-2 real-valued NN having no critical points based on a hierarchical structure is described below.

1) Consider NET 1 (1-1-1 complex-valued NN) defined in Section II-A.

2) Create NET 2 shown in Fig.1 by decomposing NET 1 where the complex numbers are decomposed into two real numbers. That is, the complex number $a + ib \in \mathbf{C}$ representing the complex-valued weight between the input neuron and the hidden neuron is decomposed into the two real numbers $a \in \mathbf{R}$ and $b \in \mathbf{R}$. The complex number $v + iw \in \mathbf{C}$ representing the complex-valued weight between the hidden neuron and the output neuron is decomposed into the two real numbers $v \in \mathbf{R}$ and $w \in \mathbf{R}$. The complex number $c + id \in \mathbf{C}$ representing the complex-valued threshold of the hidden neuron is decomposed into the two real numbers $c \in \mathbf{R}$ and $d \in \mathbf{R}$. The complex number $p + iq \in \mathbf{C}$ representing the complex-valued threshold of the output neuron is decomposed into $p \in \mathbf{R}$ and $q \in \mathbf{R}$.

3) The activation functions of NET 2 are as follows:

$$f_R(u) = \tanh(u), u \in \mathbf{R} \quad (9)$$

for the hidden neurons, and

$$g_R(u) = u, u \in \mathbf{R} \quad (10)$$

for the output neurons. The following conditions are imposed on NET 2 for the assumption that $a + ib \neq 0$ and $v + iw \neq 0$ for NET 1: ($a \neq 0$ or $b \neq 0$) and ($v \neq 0$ or $w \neq 0$).

B. Implementation of the 4-4-4 real-valued NN

Practical implementation of the 4-4-4 real-valued NN having no critical points based on a hierarchical structure is the following.

1) Consider NET 4 (1-1-1 quaternionic NN) defined in Section II-B.

2) Create NET 5 shown in Fig.3 by decomposing NET 4 where the quaternions are decomposed into four real numbers. That is, quaternion $A = a + ib + jc + kd \in \mathbf{Q}$ representing the quaternionic weight between the input neuron and the hidden neuron is decomposed into the four real numbers $a \in \mathbf{R}$, $b \in \mathbf{R}$, $c \in \mathbf{R}$, and $d \in \mathbf{R}$. The quaternion $B = \alpha + i\beta + j\gamma + k\delta \in \mathbf{Q}$ representing the quaternionic weight between a hidden neuron and an output neuron is decomposed into the four real numbers $\alpha \in \mathbf{R}$, $\beta \in \mathbf{R}$, $\gamma \in \mathbf{R}$, and $\delta \in \mathbf{R}$. The quaternion $C = p + iq + jr + ks \in \mathbf{Q}$ representing the quaternionic threshold of the hidden neuron is decomposed into four real numbers $p \in \mathbf{R}$, $q \in \mathbf{R}$, $r \in \mathbf{R}$, and $s \in \mathbf{R}$. The quaternion $D = \mu + iv + j\rho + k\sigma \in \mathbf{Q}$ representing the quaternionic threshold of the output neuron is decomposed into four real numbers $\mu \in \mathbf{R}$, $v \in \mathbf{R}$, $\rho \in \mathbf{R}$, and $\sigma \in \mathbf{R}$. The activation functions of NET 5 are as follows:

$$f_R(u) = \tanh(u), u \in \mathbf{R} \quad (11)$$

for the hidden neurons, and

$$g_R(u) = u, u \in \mathbf{R} \quad (12)$$

for the output neurons.

The following conditions are imposed on NET 5 for the assumption that $A = a + ib + jc + kd \neq 0$ and $B = \alpha + i\beta + j\gamma + k\delta \neq 0$ for NET 4: ($a \neq 0$ or $b \neq 0$ or $c \neq 0$ or $d \neq 0$) and ($\alpha \neq 0$ or $\beta \neq 0$ or $\gamma \neq 0$ or $\delta \neq 0$).

C. Implementation of the 2-2-2 complex-valued NN

The practical implementation of the 2-2-2 complex-valued NN having no critical points based on a hierarchical structure is the following.

1) Consider NET 7 (1-1-1 quaternionic NN) defined in Section II-C.

Create NET 8 shown in Fig. 4 by decomposing NET 7 where the quaternions are decomposed into the two complex numbers. That is, the quaternion $A = a + ib + jc + kd \in \mathbf{Q}$ representing the quaternionic weight between the input neuron and the hidden neuron is decomposed into the two complex numbers $a' = a + ib \in \mathbf{C}$ and $c' = c + id \in \mathbf{C}$. The quaternion $B = \alpha + i\beta + j\gamma + k\delta \in \mathbf{Q}$ representing the quaternionic weight between a hidden neuron and an output neuron is decomposed into the two complex numbers $\alpha' = \alpha + i\beta \in \mathbf{C}$ and $\gamma' = \gamma + i\delta \in \mathbf{C}$.

The quaternion $D = \mu + iv + j\rho + k\sigma \in \mathbf{Q}$ representing the quaternionic threshold of the output neuron is decomposed into two complex numbers $\mu' = \mu + iv \in \mathbf{C}$ and $\rho' = \rho + i\sigma \in \mathbf{C}$. The activation functions of NET 8 are the following: (1) for the hidden neurons and (2) for the output neurons.

The following conditions are imposed on NET 8 for the assumption that $a + ib \neq \pm i(c + id)$ ((6)) and $C = 0$ ((7)) for NET 7: $a' \neq \pm ic'$, and the thresholds of the hidden neurons are all equal to zero.

ACKNOWLEDGMENT

The author would like to give special thanks to the anonymous reviewers for valuable comments.

REFERENCES

- [1] T. Nitta, "Local minima in hierarchical structures of complex-valued neural networks," *Neural Networks*, vol. 43, pp.1-7, 2013
- [2] M. Kobayashi, "Exceptional reducibility of complex-valued neural networks," *IEEE Trans. Neural Networks*, vol.21, no.7, pp.1060-1072, 2010.
- [3] E. Hitzer., T. Nitta, and Y. Kuroe, "Applications of clifford's geometric algebra," *Advances in Applied Clifford Algebras*, vol.23, no.2, pp.377-404, DOI: 10.1007/s00006-013-0378-4, 2013.

Acoustic Strength of Green Turtle and Fish Based on FFT Analysis

Azrul Mahfurdz, Sunardi, Hamzah Ahmad
Faculty of Electrical and Electronics Engineering
University Malaysia Pahang (UMP)
26600 Pekan, Malaysia

Syed Abdullah Syed Abdul Kadir, Nazuki Sulong
Turtle and Marine Ecosystem Center (TUMEC)
Malaysia Fisheries Department
Rantau Abang, 23050 Dungun, Malaysia

Abstract—the acoustic power at difference angle and distance were measure for four different ages of Green Turtles and three species of fish using modified echo sounder V1082. The echo signal from TVG output was digitized at a sampling rate 1MHz using analog to digital converter (Measurement Computing USB1208HS). Animals were tied with wood frame to ensure it can't move away from the sound beam. The scatter value for fish demonstrates echo strength is different and depends on the angle of measurement. The lowest acoustic power of fish was recorded from their tail. The finding show that, there is significant difference between fish and turtles aged 12 to 18 years at 4.5 meter and 5 meter. The carapace and plastron of sea turtle gives high backscattering strength compare to other side. The high value obtained probably because of the hard surface of the carapace and plastron. This result is considered important in determining the best method of separating sea turtle and fish. Through this result, revealed that size, surface and animal angle play important role in determining acoustic strength value.

Keywords—Echosounder; Green Turtle; acoustic power; TED

I. INTRODUCTION

Sea turtles are marine reptiles that can be found live throughout the world tropical and subtropical seas. Sea turtles can be threatened by several factors, some natural and others caused by human activities.

Over the last few centuries, sea turtle populations have declined dramatically due to various activities such as shore development, oil exploration, commercial fishing, marine recreation and pollution.

By catch in fisheries activities has been determined to be a major factor of death for juvenile and adult sea turtles [1][2]. For several years now we have heard large numbers of turtle taken in fisheries net. Report from Japanese fisheries an estimated 40,000 sea turtles of three species were caught and 16,000 dead in the Japanese tuna long line fleet in the pacific [3].

Interview with six islander drift net vessel owner and operators in Malaysia reported 140 turtle were caught annually in 2005 to 2006. Green and Hawksbill species were report to be most frequently caught [4].

By catch of sea turtles in shrimp fisheries in tropical areas attracted more public concern especially problems related to by catch in trawl fisheries. This issue has had wide political and economic impacts on global fisheries and trade [5].

Although there are method avoiding turtle trap in fisheries net like gear modification, material and fishing method [4], but that method need cooperation between the fisherman to modify their vessel and change fishing method.

To address this problem, National Marine Fisheries Service (NMFS) suggested every shrimp trawler larger than 25 feet to use a turtle excluder device (TED) [6]. The process of reducing the incidental capture of sea turtles in regional shrimp fisheries through the use of TED has been extremely important [7].

A solution technique to separate turtles from shrimps in trawl was available by the early 1980 [8]. A traditional TED generally consists of metal grids that have been installed in a trawling net to enable endangered sea turtles to pass safely out of the net through a trapdoor [9]. Although metal grid TED method gives the solution to protect sea turtle but this method may reduce number of fish and shrimp catches [2]. The reports using TED with and without accelerator funnels were cause shrimp loss rates of 3.6 and 13.6 percent respectively [9].

The observation vessel equipped with TED found that highest reduction in prawn catch occurred during tows through areas with large amounts of star fish, sponges, sea urchins, sea cucumber and benthic debris. Catch loss occurred as a result starfish blocking the grid or tangling the guiding flap, causing inefficient operation [10].

The traditional TED consist metal trap door in trawling net seen not efficient solution for reducing turtle by catch, because it would exclude the larger commercial specimens [11].

Therefore the improvement of TED is needed in order to ensure the device use to separate endangered species can be used effectively. One of the alternative solution is suggested using sound technique [12][13].

Although the using of ultrasound is capable to prevent turtles from entering the fishing net, the alert sound will be emitted all the time. This situation will lead to wastage of power and disturb other marine live that also sensitive to the sound

To overcome this problem, the device should be able detect the presence of sea turtle. By this method sounds are emitted only after the device was able to identified sea turtle. One of the best ways in detecting underwater object is using acoustic technique. This technique has been choosing because of the sound's ability to propagate long distance in water [14].

The knowledge on turtle identification using sound is very limited in previous study. So the acoustic strength of the sea turtle in this study becomes important on designing electronic turtle excluder device.

II. MARINE LIFE DETECTION USING SOUND

Sound technique has been widely used in various fields to identify objects in the water. These applications include tracking underwater vehicle, aquatic vegetation detection and fisheries research.

The device used to observe marine animal called echo sounder. These tools are widely used to detect the distribution of fish. Basically, acoustic echo sounders have operated at frequencies in the tens to hundreds of kilohertz [15].

Aquatic organisms are complicated scatter by nature through shape, size, orientation, swim bladder and so on [16]. Smaller animals have lower echo strengths and larger animals have higher echo strengths [17][18].

The presence of swim bladder in fish body is the primary biological factor influencing the amount and variability of backscattering sound from fish [19]. Natural variations in swim bladder volume and shape may cause variation in fish echo. Understanding scattering strength of the swim bladder is important factor to study fish school. The echo waveform is different depends on orientation of fish and strongest echo occurred when the incident signal was perpendicular to the swimbladder [20].

Acoustic strength for mammals like whale may be depending on their lung and blubber layer [21]. Other than that, Dolphins have a combination of unique scattering characteristics that makes it possible to separate them from other animal. Stronger echoes, expected from their lung [22].

There are many approaches to modeling the scattering of sound by objects. The particular approach depends upon the shape and material properties of the body [23]. Study on acoustic strength of the shelled animal quite challenge because involved a variety of body shapes and biology properties, so their acoustic scattering characteristic is sometimes very complicated [24].

The scattering process of the animals was observed to be quite complex as the echoes were strongly dependent upon both frequency and angle of orientation [25][26].

Scattering from elastic shelled animal like periwinkles is characterized by a very strong echo reflected by their hard shell and also angle of orientation [27]. Moreover, study on acoustic scattering by a shell covered seafloor discovered that shellfish played an important role scattering seafloor [25].

Studies on the sea turtle are very limited because there are no acoustic characteristic of that animal recorded. However, study on fresh water turtle discovered that it can be detecting using echo sounder and the high backscattering strength is from their shell [28].

This finding has paved a way to differentiate echo strength between sea turtle and fish, which is sharing same habitat in the sea.

III. EXPERIMENT PROCEDURE

The acoustic data collection were conducted in Turtle and Marine Ecosystem Center hatchery, Rantau Abang Terengganu, Malaysia. The marine animals that have been involved in this study were four Green Turtles (*Chelonia Mydas*) and three species of fish as listed in Table I. The experiment conducted in a 13m x 2.4m rectangular tank contained 8235.68 gallons saline water. Animals were tied with wood frame as depicted in Fig. 1, to ensure it can't move away from the sound beam. The frame has been designed to make sure it is able to measure turtle at the different angle.

TABLE I. ANIMAL INVOLVED IN EXPERIMENT

Animal	Age	Weight	Carapace size
Green Turtle	1 year	1.8kg	25cmx23cm
	5 years	10kg	43cmx41cm
	12 years	27kg	61cmx56cm
	18 years	60kg	71cmx61cm
Fish	Species		Body Size
	Indian Mackerel		19cmx4.5cm
	Indian Scad		18cmx4cm
	Bigeye Scad		21.5cmx5.2cm

There are five different angles of measurement for green turtle (head, tail, side, carapace & plastron) and three angles for fish (head, lateral and tail). Reflected signal from animals were measure at 1 meter to 5 meter distances. Sea turtles will be lifted to the surface every 15 minutes to breathe, it is important to avoid drowning in the water.

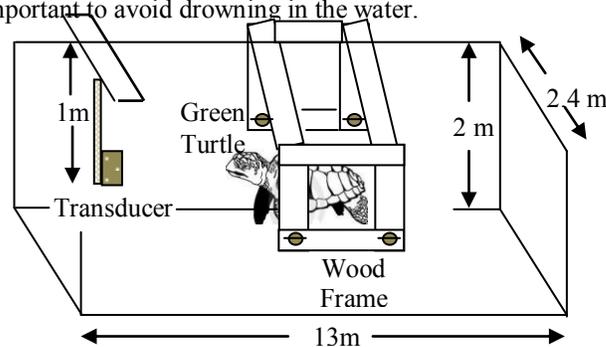


Fig. 1. Sea turtle attach with wood frame to ensure the turtle can't move away from the sound beam.

Modified dual frequency echo sounder V1082 used in this research. The signal is taken from TVG output and connected to the high speed analog digital converter (Measurement Computing USB-1208HS). The envelope of the echo was digitized at a sampling rate 1MHz using echo recording program created in matlab software. The total of 6525 random sample of sound has been recorded and store in laptop. There are several things to be considered before experiment was carried out such as echo reflection from tanks, operating frequency and sound speed. Therefore, device should be handling and setting properly. The echosounder parameters setting involved in this study are shown in Table II.

TABLE II. ECHOSOUNDER PARAMETER SETTING

Parameter	Setting
Frequency	200kHz
Gain	35dB
Noise Reduction	Low
Echo Level	7
STC	4
Dynamic Ring	5dB

IV. DATA ANALYSIS AND RESULT

The acoustic characteristics of the empty frame were measure before it has been attach to green turtle and fish. Visual observation of the echogram revealed that maximum detection without interference is five meter. Therefore, the measurement carries out in this research limited to five meter distance.

The primary objectives of this study were to compare acoustic power backscatter from sea turtle and fish. Sound sample was analysed in time and frequency domain using matlab program. Despite the sounds of 6525 samples were recorded, only 580 of them can be used in the analysis. The example echo waveform is depicted in Fig. 2.

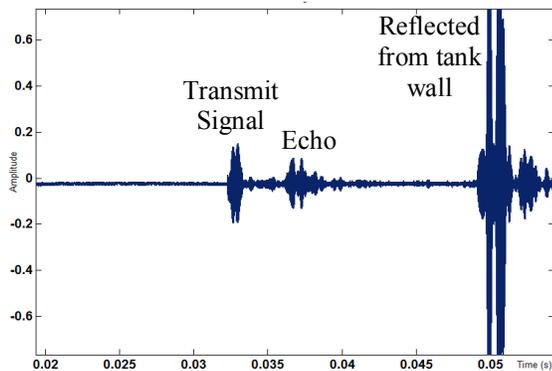


Fig. 2. Envelope echo signal in time domain.

A sound wave received from TVG output contains a transmitter and echo signal. To investigate the power spectrum in the frequency domain, each transmitted signal must be removed first, to ensure that the values obtained contains only the reflected signal from the object. The total number FFT calculated is 8192 point and power value is taken from the highest magnitude peak starts at frequency 450 KHz to 460 KHz. The power spectrum and magnitude point involved in calculating echo power of the green turtle and fish is shown in Fig. 3.

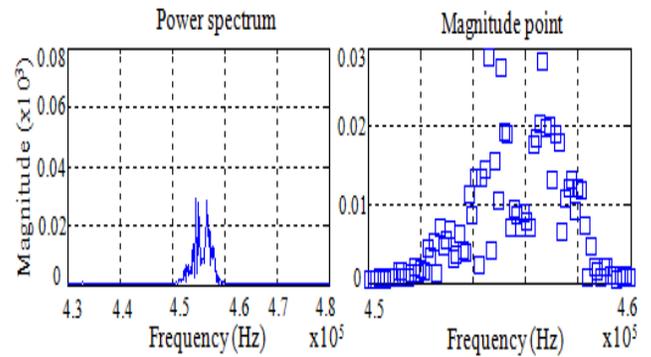


Fig. 3. Power spectrum and magnitude point of the echo signal

The results are presented in scatter plot based on echo power value. The fish echo power strength represented by circle symbols and green turtle represented by triangle symbols (Fig.4). The power value of sea turtle is obtained from head, side and tail angle, while fish is taken at head, lateral and tail from 1meter to 3meter distance. From the graph, it is obviously show that echo power are overlapped each other. However, there is some value of turtle located above the fish echo level.

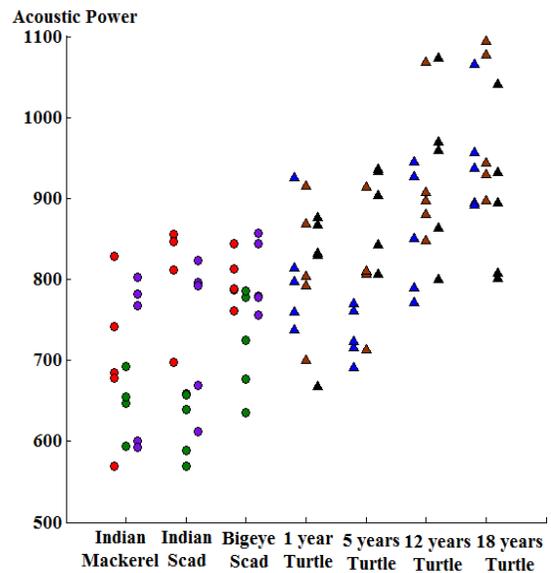


Fig. 4. Echo power comparison at 1 meter to 3 meter distance.

Based on the Fig. 5, the result revealed that there are significant different between green turtle and fish mainly for 12 years and 18 years turtle (in the dashed line circle). The acoustic power has been plotted for 3meter to 5meter distance. The difference value obtained at 4.5 meter and 5 meter distance.

Most of the value located above fish echo level in range 900 to 1300. From the graph, we also found that acoustic power range of fish is between 500 to 900.

Scatter value for fish demonstrate echo strength are different and also depend on angle of measurement. Observation on Indian Mackerel species show lateral side contributed high value compare to other side. Moreover, the lowest value of fish obtained is from their tail side, which is as we know have small shape.

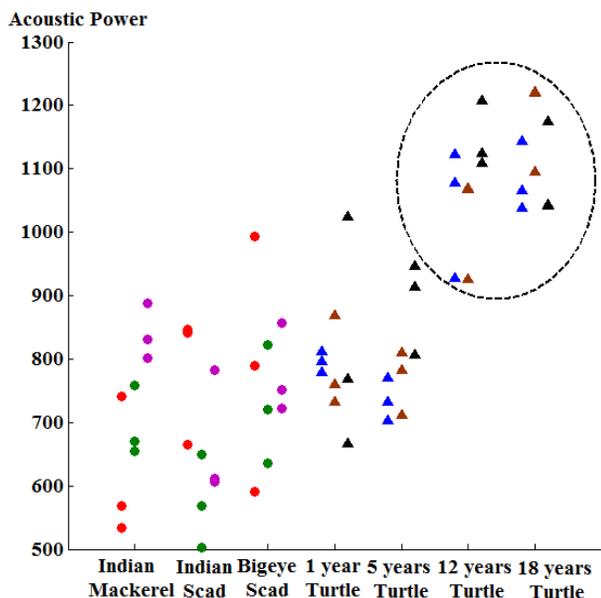


Fig. 5. Echo power comparison at 3 meter to 5 meter distance.

The comparison between fish and turtle at carapace and plastron angle are shown in Fig. 6. The carapace echo power value represented by blue circle and plastron in brown circle. Although, result for head, tail and side angle not show different for all turtles, but comparison the echo power value for the carapace and plastron shows the results otherwise. The highest power recorded is from 18 years turtle plastron, which is contributed 2207.23. Most of the value obtained from this angle located above 1000, which is higher compared to fish. These finding indicated that side gives high backscattering strength compare to other side.

One of the reasons behind of this event is a sound covering a broad surface of the carapace and plastron. Other than that, the high value obtained may also because of the hard surface of the carapace and plastron.

V. CONCLUSIONS

The results demonstrate the capability of modified echosounder to detect sea turtle and some species of fish. Measurements were made from all different angles from 1 meter to 5 meter distance, in order to find the difference echo range between that marine animal. The experimental results show that, there is significant difference between fish and turtles aged 12 to 18 years at 4.5 meter and 5 meter distance and not significant for age 1 to 5 years.

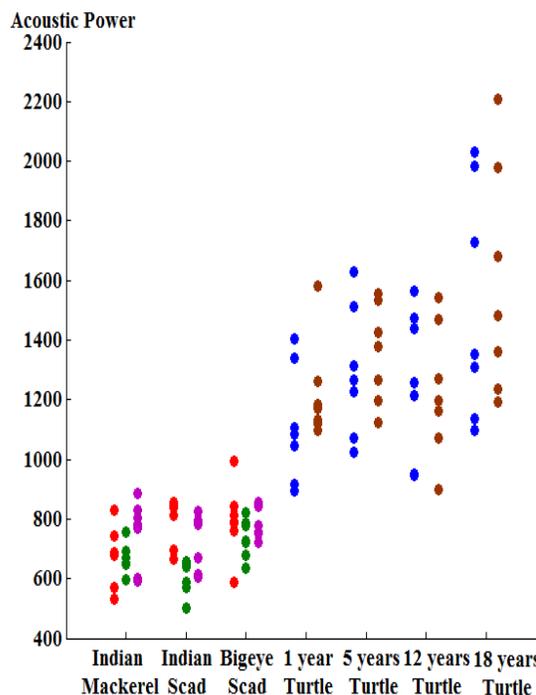


Fig. 6. Echo power comparison between fish and turtle at carapace and plastron angle

Besides that, results also show overlap happened at 1 meter to 3 meter distance for both animals. This condition may occur due to the sound beam aperture is too small at that distance and cause sound energy is focused on the same surface. However, this finding is considered important in determining the best method of separating them from other animals. Other than that, through this result, it revealed that size, surface and animal angle play important role to determine acoustic strength value. Based on the result, we can conclude that by using this method we can separate adult turtle and fish, especially for adult turtles that are frequently victims by fishing vessel. Although the study show significant result, but further research must be conduct for different species of fish, in order to ensure there are no overlap value between sea turtle and fish. Additionally, experiments are carried out on the dead fish, so another experiment on live fish is proposed to ensure that the values obtained accurate. This study only focused measurement in the fiber tank, which is limited space and distance. Although, we get the clean signal without interference from tank wall, but the maximum range of experiments that can be performed is limited to 5 meter position. Therefore, further research suggested conducting in large space such as pool and sea.

ACKNOWLEDGMENT

We thank TUMEC staff for expert sharing especially in handling turtle and providing facilities during experiment. We also want to thank people who involved in the experiment directly or indirectly to make sure the success of this study and not forgetting to Electrical & Electronic Engineering Faculty, Universiti Malaysia Pahang because providing a good learning environment and encourage a culture of research and innovation among students.

REFERENCES

- [1] C.J. McDaniel, L.B. Crowder, and J.A. Priddy, "A spatial analysis of sea turtle abundance and shrimp intensity in the U.S Gulf of Mexico," Proceedings of The Nineteenth Annual Symposium On Sea Turtle Conservation and Biology, 1999.
- [2] R. Arauz, "Implementation of the turtle excluder device (TED) by the shrimp fleet of Pacific Central America," Proceedings of The Eighteenth International Sea Turtle Symposium, 2000.
- [3] D.T. Crause, "After TEDs : What's next," Proceedings of The Eighteenth International Sea Turtle Symposium, 2000.
- [4] E. Gilman, J. Gearhart, B. Price, S. Eckert S, H. Miliken, J. Wang, Y. Swimmer, D. Shiode, O. Abe, S.H. Peckham, M. Chaloupka, M. Hall, J. Mangel, J.A. Shigueto, P. Palzell, and A. Ishizaki, Fish and Fisheries, 2010, Vol. 11, pp57-88.
- [5] J.W. Valdemarsen, and P. Suuronen, "Modifying fishing gear to achieve ecosystem objectives," Conference on RESPONSIBLE FISHERIES IN THE MARINE ECOSYSTEM, 2001, Iceland.
- [6] R.L. Lewisson, L.B. Crowder, and D.J. Sharer, "The impact of turtle excluder devices and fisheries closures on Loggerhead and Kemp's Ridley stranding in the Western Gulf of Mexico," Conservation Biology, vol. 17 n. 4, 2003, pp. 1089 - 1097.
- [7] A.Yudhana, J. Din, S. Abdullah, Sunardi, "Sound stimulus for detection turtle hearing Threshold," International Conference Information & Communication Technology and System, August 19 - 21, 2008, Bandung.
- [8] S.P Epperly, "Fisheries related mortality and turtle excluder devices (TEDs)," The Biology of Sea Turtles, vol. 11, 2003, pp. 339 - 353.
- [9] B.J Gallaway, J.G. Cole, J.M. Nance, R.A. Hart, G.L. Graham, "Shrimp loss associated with turtle excluder devices : Are the historical estimates statistically biased," North American Journal of Fisheries Management, vol. 28, 2008, pp. 203-211.
- [10] J.G. McGilvray, R.P. Mounsey, MacCartie, "The AusTED II, an improved trawl efficiency device 1.Design theories", Fisheries Research 40 (1999) 17-27.
- [11] P. Casale, L. Laurent, G. De Metrio, "Incidental capture of marine turtles by the Italian trawl fishery in the north Adriatic Sea," Biological Conservation 119 (2004) 287-295.
- [12] A. Yudhana, Sunardi, J. Din, S. Abdullah, and R.B.R. Hassan, "Turtle hearing capability based on ABR signal assessment," TELKOMNIKA : Indonesian Journal of Electrical Engineering, 2010, Vol. 8(2), pp. 187-194.
- [13] M. Lenhardt, "Marine turtle acoustic repellent/alerting apparatus and method," Reviews of Acoustical Patents Journal Acoustical Society of America, vol. 112 n.5, 2002.
- [14] S. Sarangapani, J.H. Miller, G.R. Potty, D.B. Reeder, T.K. Stanton, D. Chu, "Measurements and modeling of the target strength of divers," Oceans 2005 - Europe, vol. 952, 2005, pp. 952 - 956.
- [15] T.K. Stanton, D. Chu, M. Jech, D.I. James, " New broadband methods for resonance classification and high resolution imagery of fish with swimbladders using a modified commercial broadband echosounder," ICES Journal of Marine Sciences, 2010 pp. 365 - 378.
- [16] K. Abe, K. Sadayasu, K. Sawad, K. Ishii, Y. Takao, "Precise target strength measurement and morphological observation of juvenile Walleye Pollock (Theragra Chalcogramma)," MTTs/IEEE TECHNO - OCEAN '04, vol. 1, 2004, pp. 370 - 374.
- [17] K.J. Benoit Bird, W.W.L Au, "Echo strength and density structure of Hawaiian Mesopelagic boundary patches," J. Acoust. Soc. Am., vol. 114, 2003, pp. 1888 - 1897.
- [18] J. Frouzova, J. Kubecka, "Changes of acoustic target strength during juvenile Perch development," Fisheries Research, vol. 66, 2004, pp. 355 - 361.
- [19] Sunardi, A.Yudhana, J. Din, R.B.R. Hassan, "Fish species identification based on its acoustic target strength using in situ measurement," Aquaculture, Aquarium, Conservation & Legislation International Journal of The Bioflux Society, vol. 3 n. 3, 2010, pp. 195-209.
- [20] W.W.L. Au, and K.J. Benoit Bird, "Acoustic backscattering by Hawaiian lutjanidsnappers," J. Acoust. Soc. Am. Vol. 114(5), 2003.
- [21] J.H. Miller, D.C. Potter, "Active high frequency phased array sonar for whale shipstrikeavoidance: target strength measurements," MTTs/IEEE Conference And Exhibition, vol. 4, 2001, pp. 2104 -2107.
- [22] K.J. Benoit Bird, and W.W.L Au, "Prey dynamics affect foraging by a pelagic predator (StenellaLongirostris) over a range of spatial and temporal scales," Behavior Ecology Sociobiology, vol.53, 2003b, pp. 2104 - 2107.
- [23] T.K. Stanton, P.H. Wiebe, D. Chu, "Difference between sound scattering by weakly scattering spheres and finite length cylinders with applications to sound scattering by zooplankton," J. Acoust. Soc. Am., vol. 103 n.1, 1998.
- [24] T. Mukai, K. Lida, Y. Ando, H. Mikami, Y. Maki, R. Matsukura, "Measurements of swimming angles, density and sound speed of the krill Euphausia Pacifica for target strength estimation," MTTs/IEEE TECHNO - OCEAN '04, 2004, pp. 383 - 388.
- [25] T.K. Stanton, "On acoustic scattering by a shell covered seafloor," J. Acoust. Soc. Am. Vol. 108 n. 2, 2000.
- [26] T.K. Stanton, D. Chu, P.H. Wiebe, R.L. Eastword, J.D. Warren, "Acoustic scattering by benthic and planktonic shelled animals," J. Acoust. Soc. Am., 108(2), 2008.
- [27] J.D. Warren, T.K. Stanton, D.E. McGehee, D. Chu, "Effect of animal orientation on acoustic estimates of zooplankton properties," IEEE Journal of Oceanic Engineering, vol. 27, 2002, pp. 130 - 138.
- [28] A. mahfurdz, Sunardi, H. Ahmad, Target strength of underwater objects using echo voltage reference, International Journal of Computer and Communication Engineering, Vol. 1, No. 4, 2012.

A Distributed Key Based Security Framework for Private Clouds

Ali Shahbazi

Dept. of Computer Science
East Carolina University
Greenville, NC 27858 USA

Julian Brinkley

Dept. of Computer Science
East Carolina University
Greenville, NC 27858 USA

Ali Karahroudy

Dept. of Computer Science
East Carolina University
Greenville, NC 27858 USA

Nasseh Tabrizi

Dept. of Computer Science
East Carolina University
Greenville, NC 27858 USA

Abstract—Cloud computing in its various forms continues to grow in popularity as organizations of all sizes seek to capitalize on the cloud’s scalability, externalization of infrastructure and administration and generally reduced application deployment costs. But while the attractiveness of these public cloud services is obvious, the ability to capitalize on these benefits is significantly limited for those organization requiring high levels of data security. It is often difficult if not impossible from a legal or regulatory perspective for government agencies or health services organizations for instance to use these cloud services given their many documented data security issues. As a middle ground between the benefits and security concerns of public clouds, hybrid clouds have emerged as an attractive alternative; limiting access, conceptually, to users within an organization or within a specific subset of users within an organization. Private clouds being significant options in hybrid clouds, however, are still susceptible to security vulnerabilities, a fact which points to the necessity of security frameworks capable of addressing these issues. In this paper we introduce the Treasure Island Security Framework (TISF), a conceptual security framework designed to specifically address the security needs of private clouds. We have based our framework on a Distributed Key and Sequentially Addressing Distributed file system (DKASA); itself borrowing heavily from the Google File System and Hadoop. Our approach utilizes a distributed key methodology combined with sequential chunk addressing and dynamic reconstruction of metadata to produce a more secure private cloud. The goal of this work is not to evaluate framework from an operational perspective but to instead provide the conceptual underpinning for the TISF. Experimental findings from our evaluation of the framework within a pilot project will be provided in a subsequent work.

Keywords—private cloud security framework; distributed key; dynamic metadata reconstruction; cloud security

I. INTRODUCTION

Cloud computing, in its varying incarnations, continues to emerge as an attractive deployment option for enterprises and organizations seeking ways to reduce and better manage the costs associated with application deployment. Commercial cloud services allow organizations to consume computing resources in a manner similar to traditional utilities like electricity or water; paying for computing resources in a matter

commensurate with their use. This Platform as a Service (PaaS) model additionally externalizes the costs associated with infrastructure and systems administration while providing a potentially more scalable and reliable deployment environment [1]. These significant benefits have created an impression in the minds of many consumers and organizational decision makers that “the cloud” is the answer to any number of software dilemmas.

But while the aforementioned benefits are undoubtedly attractive, these public cloud services are not without significant drawbacks within certain usage scenarios. In circumstances involving highly confidential, sensitive or secret data, security issues inherent to public clouds render their use inadvisable, impractical or even impossible depending upon legal and regulatory requirements. Government entities and health care organizations for instance often face legally mandated data security requirements that nearly all cloud services are incapable of satisfying due to a host of real and perceived security related issues [2-5]. The perception of the security and confidentiality vulnerabilities of public clouds has been reinforced by a number of data breaches reported in the media [6]. While governmental entities, regulatory bodies and medical organizations may benefit from the cloud given the large volumes of data generally involved with their respective activities, the risk of a single data breach often outweighs the potential benefits. Although some cloud providers continue to address these security and regulatory issues, as Microsoft has with the addition of Health Insurance Portability and Accountability Act (HIPAA) compliance features added to its Windows Azure cloud service [7], public clouds still possess too many security unknowns for many organizations.

As a result of these issues hybrid clouds have emerged as a middle ground between the aforementioned benefits of cloud computing and the identified security issues. Private clouds, significant aspects of private clouds, are developed and administered by an organization’s internal IT department for the exclusive use by specific users or user groups within the organization [1, 8]. It is presumed that this greater degree of control *guarantees* an elimination of the security and regulatory issues posed by public clouds. But these private clouds may

also suffer from security issues, leading to a number of proposals designed to address these issues. In this paper we introduce the Treasure Island Security Framework (TISF) which builds upon existing thinking to provide a scalable security framework for private clouds.

II. RELATED WORK

There is a significant body of work which documents the challenges and proposed solutions to the issue of cloud security; both public and private [9-12]. Many of these works take substantially different approaches to the issue of cloud security given the broad topic that is cloud computing. The approach taken within this works revolves primarily around an overlapping use of encryption, distributed key methodology, sequential chunk addressing and dynamic metadata reconstruction to improve system security.

Distributed key methodology is not a new concept having significant support with the literature albeit in significantly different conceptualizations and implementations [13-14]. A form of distributed key methodology serves as the backbone of the security effort proposed within this work whereby the key necessary to decrypt individual file chunks and reconstruct a stored file is distributed within our proposed system. This methodology stands in contrast to the some of the more common methods of cloud authentication such as those based primarily on password protection and Private Key Infrastructure (PKI). While these techniques are relatively easy to implement, they have a number of deficiencies which have been documented in the scientific literature and the media. Password protection for instance is dependent upon the user's ability to maintain confidential information against social engineering attacks wherein information enabling the reconstruction of one or more passwords may be divulged inadvertently by a user [15]. Password methodology is additionally troublesome given the average user's penchant for password reuse [16]. While the reuse of an existing password minimizes the cognitive load that memorizing a number of different passwords for different system creates, reuse effectively means that a password compromised for one system allows malicious users to access a host of other user accounts.

This is especially troublesome for email accounts which often serve as key link in the user verification process for many systems. Ticket based authentication using the Kerberos protocol is an improvement over pure password based protection however this methodology still possesses security risks [17]. A breach of a system's authentication server will result in the exposure of all user accounts due to the centralization of authentication management [18]. Public Key Infrastructure (PKI) addresses some of these issues through the use of digital certificates for entity identity verification [19]. This approach however also has a number of flaws [20]. The distributed key approach utilized within our proposed system addresses many of these issues by using a decentralized form of authentication that eliminates the single point of failure found in password protection scheme given the use of a segmented, dispersed key. This methodology is further bolstered by our use of a form of dynamic metadata reconstruction, which protects information about the stored data, and chunk encryption [21].

III. THE TREASURE ISLAND SECURITY FRAMEWORK (TISF)

While the Google File System (GFS) [22-23] and the Hadoop Distributed File System (HDFS) [24-26], a GFS derivative, are commonly utilized with private clouds, we have proposed an alternative cloud architecture upon which the Treasure Island Security Framework is based. The Distributed Key and Sequentially Addressing Distributed file system (DKASA), which builds upon aspects of both GFS and HDFS, improves the security of data storage and file distribution in a private cloud primarily through the introduction of dynamic metadata reconstruction, sequential addressing and distributed key methodology.

A. Distributed Key and Sequentially Addressing File System

The Distributed Key and Sequentially Addressing Distributed file system (DKASA), as illustrated abstractly in Fig. 1, has a number of characteristics borrowed from the GFS including a single master configuration, use of fixed chunk sizes and chunk replication. We provide assumptions with respect to the configuration of DKASA within our proposed framework to contextual the security risk model discussed in Section 3 C.

1) Single Master Server

Both Single Master (SM) and Dual Main Server (DMS) configurations are possible within the DKASA file system although our proposal is based on the use of the former; mirroring the GFS. The DMS configuration, which involves the use of a management server and a file retrieval server, is potentially less robust than the SM configuration given its relatively poor performance under stress. It is likely, with moderate to high levels of network traffic and significant numbers of large files, that the retrieval server in the DMS configuration becomes a bottleneck for the entire system.

2) Fixed Chunk Size and Multiple Replicas

We anticipate the use of a fixed chunk size, which enables the use of the GFS mutation and lease method to reduce network traffic. Unlike the GFS which uses a fixed size of 64MB we have consciously chosen to leave the size of the chunk ambiguous as both large and small chunk sizes have advantages and disadvantages. A large chunk size for instance will reduce the number of chunk servers needed for each client while also reducing the client's interaction with the master for reading metadata and namespaces. This large chunk size however is also incompatible with smaller files. A smaller chunk size is compatible with smaller files however it may result in greater data fragmentation. In an actual implementation of our framework a system architect would determine the appropriate chunk size for the specific usage scenario.

Each chunk within the system will have a number of replicas (k) to ensure data availability; each replica chunk exists in isolation from the original. In the event that the original chunk is unavailable the replica system will retrieve a replica and use that data to reconstruct the original file.

3) Encryption

Each chunk's data is encrypted in the client machine and sent via secure communication using RSA [27] and Advanced

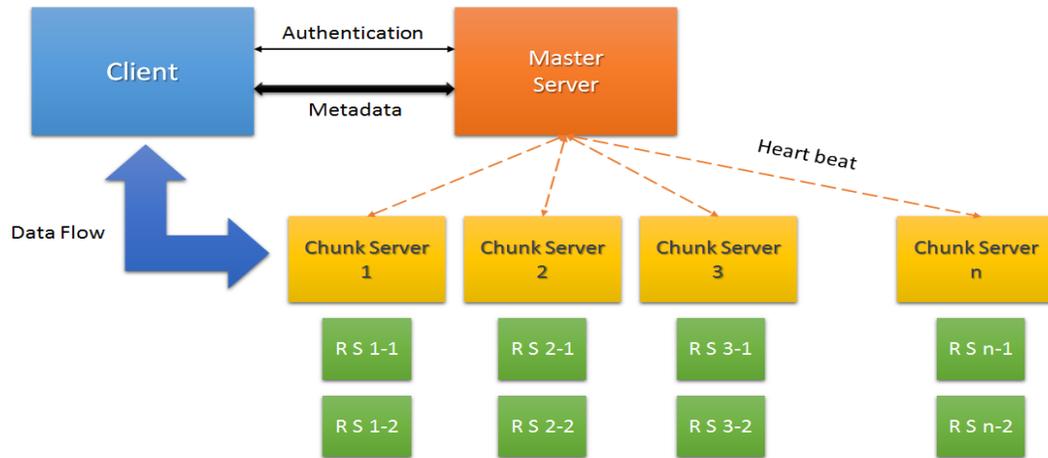


Fig. 1. Treasure Island Security Framework Abstract Model

Encryption Standard (AES) encryption [28]. The key differentiation between the method proposed in this work and traditional approaches is the use of a distributed key approach as opposed to the use of Public Key Infrastructure (PKI).

4) Distributed Key and Sequential Addressing

The use of a distributed key methodology has been driven primarily by two factors, a desire to increase security while introducing a level of flexibility whereby different security levels exist within the system. In terms of the latter, this type of security granularity facilitates varying degrees of file and user level security as opposed to a methodology within which the level of security within the system as whole is the only manipulable value.

Distributed key methodology as proposed within the DKASA system involves the distribution of a cryptographic key into four isolated parts. The first two parts of the key are stored in the master server and the client, the third part is stored in each chunk (n) and the final part of the key is stored in the previous chunk (n-1).

A file to be stored in the private cloud will be divided into a series of sequentially addressed chunks with distinct, appended headers and footers. The header of each encrypted chunk contains the following information:

- 128 bit local deciphering key
- 128 bit remote deciphering key
- The address of the next chunk
- 128 bit status code identifying the originality property of the chunk
- 1024 bits of audit data

This header data is used by the Cloud Management Server (CMS) and the user's client during the file retrieval process to locate file chunks, decipher them and rebuild the original file using the distributed key and sequential addressing approach.

The full key necessary to decrypt each encrypted chunk is produced as a result of the concatenation of the parts of the key stored on the master server, the client, the current chunk server and the previous chunk server as illustrated in Fig. 2.

Upon successful completion of this process an interim copy of the file is available to the user on the client machine. After the user completes file manipulation (read, update, delete, etc) based on the mutation and lease method as employed within the GFS the chunks are stored on new servers. The complete file access algorithm is as follows:

procedure FileRetrieval()

- 1) Client machine sends authentication request to master server
- 2) Master server checks and approves client
- 3) Master server sends SID and service lists to client
- 4) Client asks master server for the address of the first chunk
- 5) Master server sends first chunk server address and first chunk (n-1) code part to client
- 6) Client send chunk server request, 128 bits of deciphering key, code part (n-1) and first chunk address
- 7) Based on internal algorithm client partition is reinterpreted to new deciphering code
- 8) Client reads and deciphers first chunk from the file server based on the reconstructed key
- 9) LOOP: Client refers to the next server based on the read data from the current server
- 10) Client reads and deciphers chunk from the file
- 11) server based on the reconstructed key
- 12) IF: File is complete
- 13) END LOOP
- 14) END IF
- 15) END LOOP:

end FileRetrieval

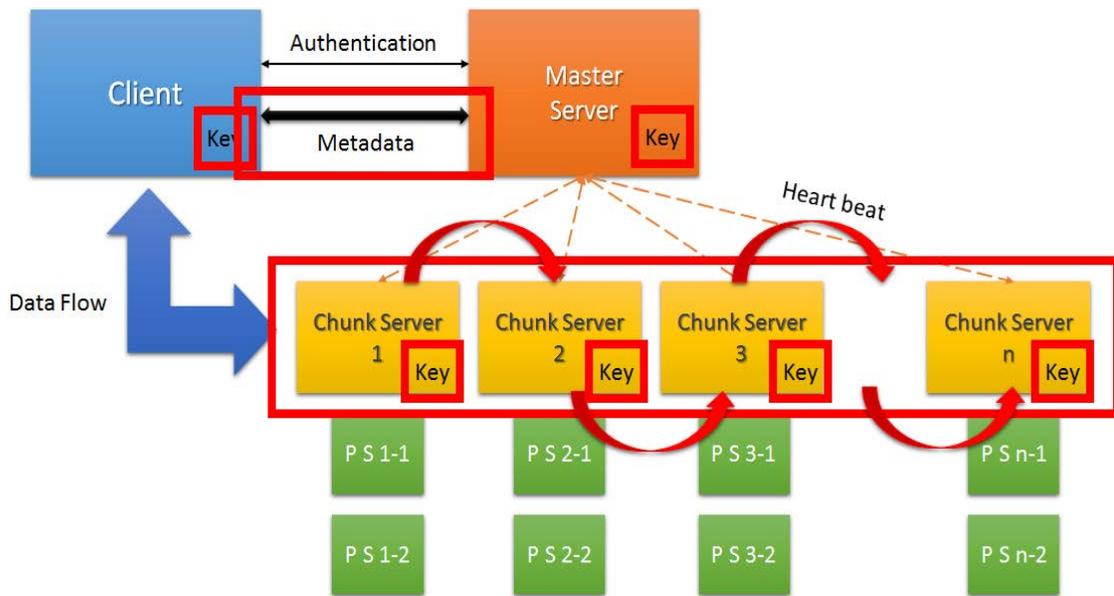


Fig. 2. TISF Sequential Addressing and Distributed Key Methodology

B. Security Risk Model

Our approach adds security beyond that found in security schemes using PKI and single password methodologies in that the likelihood of system compromise from a single attack is largely eliminated. We evaluate this claim using the previously outlined assumptions and the presumption of a single file broken down into many chunks (N), each possessing many replicas (K) within a distributed key architecture. We define the distributed key as

$$D_K = D_1 + D_{1-1} + D_{CM} + D_{MS} \quad (1)$$

All four isolated parts of the key are necessary to construct the full cryptographic key required to decipher each chunk. To evaluate the likelihood of compromise it is thus necessary to calculate the availability of each of the four key components with respect to their individual locations; the master server, the client, the original chunk and the previous chunk. Within the working system the master server is constantly operational therefore the availability of this component to an attacker is equal to 100% or

$$P_A(D_{MS}) = 1 \quad (2)$$

For the client machine the window for an attack is based on the total time of connection. For the purpose of this analysis we assume a client connection duration that is represented as TCM . For the final two key components, it is necessary for an attacker to successfully attack two different chunks, the current chunk and the previous chunk, to assemble all of the components necessary to reconstruct the full cryptographic key. The probability of a successful attack in this scenario is

$$C_n = \frac{\{(N_s - b) 2-2\}}{\{N_s! / (N_s - 2)! 2!\}} \quad (3)$$

Where the denominator is the number of all choices and the numerator is the likelihood that an attacker successfully selects the server containing the first chunk. Thus the chance of

gaining access to all the necessary items for deciphering a chunk will be:

$$P_{FA} = \frac{\{(N_s - b) 2-2\}}{\{N_s! / (N_s - 2)! 2!\}} \times \{TCM/86400\} \quad (4)$$

The availability of a file for a user, in the event of a server failure or malware attack, depends upon the number of replica chunks that exist for each original. When there are $K < n$ replica chunks a high risk situation exists for the integrity of the user data in that file chunks exist without a backup. Where $K = n$ all chunks have a backup therefore a full copy of the entire file exists. The total number of full file backups may be determined by $(K / (n+K))$. The chance of successful file retrieval after any malware attack server failure is thus:

$$AI = K P_p / (n+K) \quad (5)$$

where P_p equals availability of a server in the system, K the quantity of replica control chunks, and n the number of chunks per file.

IV. CONCLUSION AND FUTURE WORK

In this paper we have introduced a security framework for private clouds called the Treasure Island Security Framework (TISF) which is based upon a Distributed Key and Sequentially Addressing Distributed file system. We have introduced DKASA and the methodology behind its proposed implementation while evaluating the security risks inherent in our approach. We believe that our proposed approach enhances both data availability and integrity while providing a higher degree of security and backup control at both the user and file level. Perhaps the most significant advantage of the DKASA cloud as proposed is the avoidance of the most common public cloud security and data availability issues; issues which have been chronicled exhaustively in both the press and the related scientific literature. Our subsequent work will seek to evaluate our claims within a pilot project which will be documented in a future paper.

REFERENCES

- [1] R. Prodan and S. Ostermann, "A survey and taxonomy of infrastructure as a service and web hosting cloud providers", in 10th IEEE/ACM International Conference on Grid Computing, Alberta, Canada, 2009, pp. 17-25.
- [2] K. Ren, C. Wang and Q. Wang, "Security Challenges for the Public Cloud," IEEE Internet Computing, vol. 16, pp. 69–73, Jan. 2012.
- [3] R. Iglesias, R. Nicholls, A. Travis and W. Henderson, "Private Clouds with No Silver Lining: Legal Risk in Private Cloud Services," Digiworld Economic Journal, no. 85, pp. 125–140, 1st Q. 2012.
- [4] L. Kaufman, "Data Security in the World of Cloud Computing," IEEE Security and Privacy Journal IEEE, vol. 7, pp. 61-64, Jul. 2009.
- [5] S. Subashini and V. Kavitha, "A Survey on Security Issues in Service Delivery Models of Cloud Computing," Journal of Network and Computer Applications, vol. 34, no. 1, pp. 1-11, Jan. 2011.
- [6] J. Galante, O. Kharif and P. Alpeyev, "PlayStation security breach shows Amazon's cloud appeal for hackers," The Seattle Times, May, 16, 2011. [Online], Available: http://seattletimes.com/html/business/technology/2015071863_amazoncloudhackers17.html, [Accessed Jun. 29, 2013].
- [7] B. T. Horowitz, "Microsoft Adds HIPAA Compliance Features to Azure for Cloud Health Data," eWeek.com, para. 1, Jul. 27, 2012. [Online]. Available: <http://www.eWeek.com/c/a/Health-Care-IT/Microsoft-Adds-HIPAA-Compliance-in-Windows-Azure-for-Cloud-Health-Data-446671/>. [Accessed Jun. 29, 2013]
- [8] C. Babcock, "Time to Believe in 'Private Clouds'," Information Week, vol. 28, no. 12, pp. 27-30, Apr. 2009.
- [9] K. W. Nafi, T. S. Kar, S. A. Hoque and M. M. A. Hashem, "A Newer User Authentication, File encryption and Distributed Server Based Cloud Computing security architecture," International Journal of Advanced Computer Science and Applications, vol. 3, no. 10, pp. 181-186, Oct. 2012.
- [10] T. Wood, A. Gerber, K. K. Ramakrishnan, P. Shenoy and J. Van der Merwe, "The case for enterprise-ready virtual private clouds," in Proc. of the 2009 Conference on Hot Topics in Cloud Computing, San Diego, CA, 2009, pp. 4-9.
- [11] D. W. Chadwick, M. Casenove and K. Siu, "My private cloud—granting federated access to cloud resources," Journal of Cloud Computing: Advances, Systems and Applications, vol. 2, no. 1, pp. 1-16, Feb. 2013.
- [12] A. A. Karahroudy. "Security Analysis and Framework of Cloud Computing with Parity-Based Partially Distributed File System." M.S. thesis, East Carolina University, Greenville, USA, 2011.
- [13] R. Geambasu, A. A. Levy, T. Kohno, A. Krishnamurthy and H. M. Levy, "Comet: An active distributed key-value store," in 9th USENIX Symposium on Operating Systems Design and Implementation - OSDI, British Columbia, Canada, 2010 pp. 323-336.
- [14] J. Resch. "Authenticating Cloud Storage with Distributed Keys," presented at the Storage Developer Conference (SNIA), Santa Clara, CA, 2011.
- [15] S. P. Maan and M. Sharma, "Social Engineering: A Partial Technical Attack," International Journal of Computer Science Issues (IJCSI), vol. 9, no. 2, pp. 557-559, Mar. 2012. ----MAYBE FIND A SOURCE THAT IS BETTER WRITTEN
- [16] B. Ives, K. R. Walsh and H. Schneider, "The domino effect of password reuse," Communications of the ACM, vol. 47, no. 4, pp. 75-78, Apr.. 2004.
- [17] S. P. Miller, B. C. Neuman, J. I. Schiller and J. H. Saltzer, "Kerberos authentication and authorization system" in In Project Ahtna Technical Plan. 1987 ----FIX FORMAT OF CITATION
- [18] S. M. Bellovin, M. Merritt, "Limitation of the Kerberos authentication system," ACM SIGCOMM Computer Communication Review, vol. 20, no. 5, p. 119-132, Oct. 1990.
- [19] D. Solo, R. Housley and W. Ford, "X. 509 public key infrastructure certificate and CRL profile," Jan. 1999 [Online]
- [20] C. Ellison and B. Schneier, "Five risks of PKI: What you're not being told about public key infrastructure," Computer Security Journal, vol. 16, no. 1, pp. 1-7, Nov. 2000.
- [21] A. Waqar, A. Raza, H. Abbas and M. Khurram Khan, "A Framework for Preservation of Cloud Users' Data Privacy using Dynamic Reconstruction of Metadata," Journal of Network and Computer Applications, vol. 36, no. 1, pp. 235-248, Jan. 2013.
- [22] S.E. Arnold, "MapReduce, Chubby and Hadoop", KM World, vol. 19, no. 10, pp. 1-18, Nov. 2010.
- [23] S. Ghemawat, H. Gobioff and S. Leung, "The Google File System," in Proc. ACM SIGOPS Operating Systems Review, Bolton Landing, NY, USA, 2003, pp. 29-43.
- [24] M. Bhandarkar, "MapReduce programming with Apache Hadoop", IEEE International Symposium on Parallel & Distributed Processing, Taipei, Taiwan, 2010, p. 1.
- [25] K. Shvachko, Hairong Kuang, S. Radia and R. Chansler, "The Hadoop Distributed File System", in IEEE 26th Symposium on Mass Storage Systems and Technologies, Reno, USA, 2010, pp. 1-10.
- [26] T. White, Hadoop: The Definitive Guide, Sebastopol, CA: O'Reilly Media, 2009.
- [27] R. L. Rivest, A. Shamir and L. Adleman, "A Method for Obtaining Digital Signature and Public-Key Cryptosystems," Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, MA, Nov. 1977.
- [28] J. Daemen, V. Rijmen, "Announcing the Advanced Encryption Standard (AES)," Federal Information Processing Standards Publication 197, Nov. 2001.

An Open Cloud Model for Expanding Healthcare Infrastructure

Sherif E. Hussein

Computer and Systems Department, Mansoura University,
Mansoura, Egypt.

Hesham Arafat

Vice Dean of Engineering and Head of Computer and
Systems Department, Mansoura University, Mansoura,
Egypt.

Abstract—with the rapid improvement of computation facilities, healthcare still suffers limited storage space and lacks full utilization of computer infrastructure. That not only adds to the cost burden but also limits the possibility for expansion and integration with other healthcare services. Cloud computing which is based on virtualization, elastic allocation of resources, and pay as you go for used services, opened the way for the possibility to offer fully integrated and distributed healthcare systems that can expand globally. However, cloud computing with its ability to virtualize resources doesn't come cheap or safe from the healthcare perspective. The main objective of this paper is to introduce a new strategy of healthcare infrastructure implementation using private cloud based on OpenStack with the ability to expand over public cloud with hybrid cloud architecture. This research proposes the migration of legacy software and medical data to a secured private cloud with the possibility to integrate with arbitrary public clouds for services that might be needed in the future. The tools used are mainly OpenStack, DeltaCloud, and OpenShift which are open source adopted by major cloud computing companies. Their optimized integration can give an increased performance with a considerable reduction in cost without sacrificing the security aspect. Simulation was then performed using CloudSim to measure the design performance.

Keywords—Cloud Computing; OpenStack; Openshift; Cloudsim; e-health

I. INTRODUCTION

Although much research effort has been put into the design and development of novel e-Health services and applications, their interoperability and integration remain challenging issues. Health services deal with large amount of private data which needs to be both fully protected and readily available to clinicians. However, health care providers often lack the capability to commit the financial resources for either the research or the infrastructure required. In addition, the public is often skeptical about whether IT systems can be trusted with private clinical information. Past failures have fuelled a culture to restrict any innovation by both members of the public and the politicians responsible for health services [1].

Cloud computing has recently appeared as a new computing paradigm, which promises virtually unlimited resources. Customers rent resources based on the pay-as-you-go model and thus are charged only for what they use. Opposite to other service models, in-house Picture Archiving and Communication System (PACS) and application service

provider PACS [2], cloud computing offers relatively lower cost, higher reliability and scalability as shown in table 1.

TABLE I. STRENGTHS AND WEAKNESS OF IN-HOUSE PACS, ASP PACS, AND CLOUD PACS.

Computing Facilities	Service Model	Strengths	Weakness
Dedicated (Local)	In-House PACS	- Smaller network expenses - Control over data - Clear ownership of data - Fast data transmission	- Need for in-site IT expertise - Funded out from the investment budget - Need for new technology and upgrading
Shared (Hosted)	Application Service Provider PACS	- predictable costs - Use of offsite IT expertise - Scalability - Possibility to share data between structures.	- Not economical with small examination numbers - Greater network expenses
Public Cloud	- IaaS resources - PaaS development platform - SaaS PACS utilities	- Billed per study and/or megabyte - Economical with both small and large examination number - High reliability and Scalability	- Slow data transmission - security risk

The cloud system can be divided mainly into three main layers. The infrastructure as a service (IaaS) is the lowest level which delivers computing infrastructure as a service to end users. IaaS is typically provided as a set of APIs that offer access to infrastructure resources. The APIs allow creating virtual machine instances, or to store or retrieve data from a storage or database. The main benefit of virtualized infrastructure, which is offered as a service, is scalability. The commercial model is to pay for the infrastructure that is actually used. This means the designer doesn't have expensive servers idling, and if there is a spike in the visitor numbers, the designer doesn't have to worry if the hardware will cope. He simply scales the infrastructure up and down as needed. Since the servers are virtual, it's much easier to create a new one than it was to add a new box to a server farm. IaaS provides users with a way to monitor, manage, and lease resources by deploying virtual machine (VM) instances on those resources. Amazon EC2, Eucalyptus, Nimbus, OpenStack and Open

Nebula are examples of cloud infrastructure implementations [3].

In Platform as a Service (PaaS), we move a step further. We no longer have to deal with every element of the infrastructure; instead, we can regard the system as one solid platform. For example, in the case of IaaS, if the designer has a website that suddenly requires more capacity because the amount of visitors increased, he would typically fire up more virtual machine instances. With PaaS, this is no longer the case; the platform will scale up and down as necessary and takes care of the implementation details.

The platform is typically represented as a single box. Since the platform usually acts as if it were a single box, it's much easier to work with, and generally there is no need to change much in the application to be able to run on a PaaS environment. PaaS doesn't only offer cpu, memory or file storage; but also offers other parts of the infrastructure, such as databases, either in the form of a scaling traditional RDBMS system, or one of the 'NoSQL' databases that are currently gaining momentum due to its ability to distribute large amount of data over the cloud infrastructure [4]. The third available service model goes another step further in the realm of abstraction. We no longer care about infrastructure as with IaaS, nor do we care about the platform, as with PaaS. Where in the past, software was installed on the desktop, now with software as a service (SaaS); the user just creates an account and is ready to use the applications, in the comfort of the web browser. As with SaaS, we only need to worry about the application we're dealing with. Good examples of SaaS are cloud PACS utilities which can offer services for imaging centers, reading physicians, primary care clinics, and hospital management [5].

cloud, Hybrid architectures provide another choice, a middle ground.

For example, hybrid architecture could move computations to the cloud while keeping sensitive data in a secure database that resides in the private network [6].

In this paper, we provide a closer look into the implementation of healthcare infrastructure using a private cloud based on OpenStack. That was accompanied by setting a PaaS on OpenShift for connecting the private cloud to other health care services, mobile and web applications for clinicians and patients, and resources provided by different cloud providers. That hybrid design was assessed using CloudSim to measure its performance.

The rest of this paper is organized as follow; section 2, introduces a brief note about private cloud implementation, section 3, explains the potentials of expanding the private cloud infrastructure by incorporating other public cloud providers, section 4, discusses a medical application that took advantage of the implemented hybrid cloud, section 5, provides a simulation analysis to assess the performance benefits behind the proposed implementation, and finally we concluded the paper by a discussion section.

II. PRIVATE CLOUD IMPLEMENTATION

A private cloud implementation aims to avoid many of the objections including control over hospital and patients' data, worries about security, and issues connected to regulatory compliance. Because a private cloud setup is implemented safely within the corporate firewall, it remains under the control of the IT department. However, the hospital implementing the private cloud is responsible for running and managing IT resources instead of passing that responsibility on to a third-party cloud provider. Hospitals initiate private cloud projects to enable their IT infrastructure to become more capable of quickly adapting to continually evolving healthcare needs and requirements.

Launching a private cloud project involves analyzing the need for a private cloud, formulating a plan for how to create a private cloud, developing cloud policies for access and security, deploying and testing the private cloud infrastructure, and training employees and partners on the cloud computing project. To create a private cloud project strategy, a hospital identifies which of its healthcare practices can be made more efficient than before, as well as which repetitive manual tasks can be automated via the successful launch of a cloud computing project. By creating a private cloud strategy, the resulting cloud will be able to deliver automatic, scalable server virtualization, providing the benefits of automated provision of resources and the optimal use of hardware within the IT infrastructure [7]. It is important for the private cloud implementation process to analyze and ensure the proper processes and policies are in place to successfully build a secure private cloud. Research and acquire the private cloud infrastructure and cloud-enabling software that will be used, such as OpenStack, CloudStack, and Eucalyptus. Ensure the hypervisor that will manage the virtual machines and virtualized storage are available or can be purchased and installed.

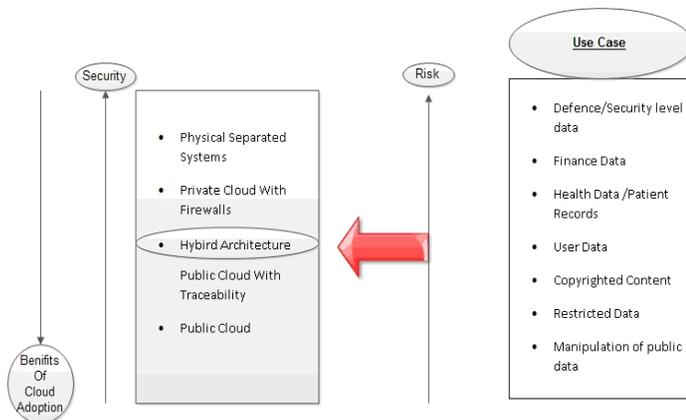


Fig. 1. The relation between security restrictions against the benefits of cloud adoption.

As shown in Figure 1, there is a spectrum of use cases where one end consists of use cases that manipulates public data, hence have very low risk associated with them, while the other end consists of use cases that manipulates credit cards, Social Security Numbers, or ultra sensitive data like nuclear tests.

For the low end of the spectrum, Cloud Computing is an obvious choice, while for the high end; Cloud Computing might never be used. However, instead of public and private

A. The cloud management software

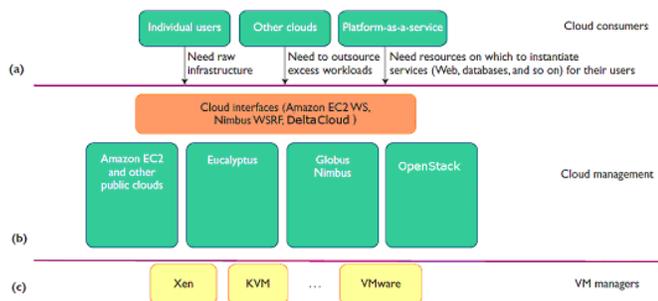


Fig. 2. The cloud ecosystem for building private clouds. (a) Cloud consumers need flexible infrastructure on demand. (b) Cloud management provides remote and secure interfaces for creating, controlling, and monitoring virtualized resources on an infrastructure-as-a-service cloud. (c) VM managers provide simple primitives (start, stop, suspend) to manage VMs on a single host.

Cloud management is the software and technologies designed for operating and monitoring applications, data, and services residing in the cloud as shown in figure 2. Cloud management tools help ensure hospital's cloud computing-based resources are working optimally and properly interacting with users and other services. Cloud management strategies typically involve numerous tasks including performance monitoring (response times, latency, uptime, etc.), security and compliance auditing and management, and initiating and overseeing disaster recovery and contingency plans. With cloud computing growing more complex and a wide variety of private, hybrid, and public cloud-based systems and infrastructure already in use, a hospital's collection of cloud management tools needs to be just as flexible and scalable as its cloud computing strategy. Choosing the appropriate cloud platform, however, can be difficult. They all have pros and cons.

We have decided to compare the capabilities of CloudStack, Eucalyptus, and OpenStack as the most notable open source systems available. Both Eucalyptus and CloudStack's application programming interface (API) provides compatibility with Amazon Web Services' Elastic Compute Cloud (EC2), the world's most popular public cloud. While OpenStack, supports public clouds built by its major vendors. OpenStack is backed by Dell, IBM, RackSpace the second leading IaaS provider after Amazon, NASA, HP the supplier of ARM cloud servers, Canonical the supplier of ubuntu which is tightly integrated with OpenStack in every release and the main operating system for ARM cloud servers [8]. The simple implementation and support of OpenStack for ARM cloud servers favored its choice for our private cloud implementation as it increases the possibility for a lesser cost and an ever growing performance [9, 10].

OpenStack refers to a collection of open-source software packages designed for building public and private clouds. OpenStack is implemented as a set of Python services that communicate with each other via message queue and database.

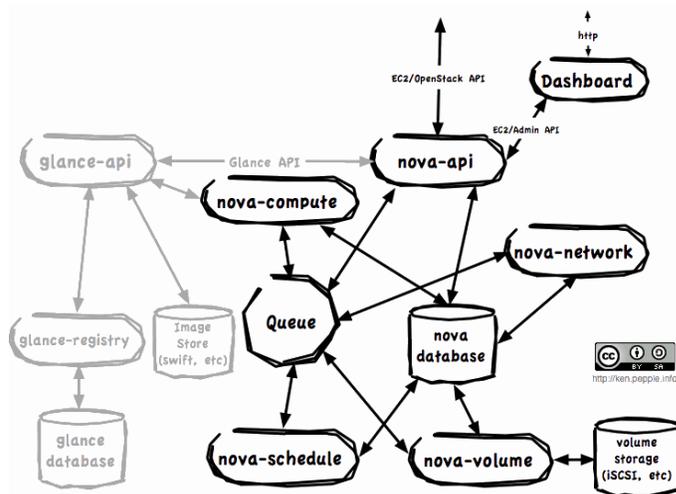


Fig. 3. OpenStack Architecture.

Figure 3 shows a conceptual overview of the OpenStack architecture, with the OpenStack Compute components (which provides an IaaS) bolded and OpenStack Glance components (which stores the virtual machine images) shown in a lighter color. The nova-api service is responsible for fielding resource requests from users. Currently, OpenStack implements two APIs: the Amazon Elastic Compute Cloud (EC2) API, as well as its own OpenStack API [11]. The nova-schedule service is responsible for scheduling compute resource requests on the available compute nodes. The nova-compute service is responsible for starting and stopping virtual machine (VM) instances on a compute node. The nova-network service is responsible for managing IP addresses and virtual LANs for the VM instances. The nova-volume service is responsible for managing network drives that can be mounted to running VM instances. The queue is a message queue implemented on top of RabbitMQ [12] which is used to implement remote procedure calls as a communication mechanism among the services. The database is a traditional relational database such as MySQL used to store persistent data shared across the components. While, the dashboard implements a web-based user interface.

B. The hypervisor

A hypervisor, also called a virtual machine manager, is a program that allows multiple operating systems to share a single hardware host. Each operating system appears to have the host's processor, memory, and other resources all to itself. However, the hypervisor is actually controlling the host processor and resources, allocating what are needed to each operating system in turn and making sure that the guest operating systems (virtual machines) cannot disrupt each other. Most OpenStack development is done with the KVM and XEN hypervisors. KVM however is free and easier to deploy than free XEN, well supported by every major distribution and is adequate for most cloud deployment scenarios (for example, EC2-like cloud with ephemeral storage) while storage support is improving for KVM. It is also simpler and more stable to use with OpenStack and is fully packed by ubuntu 12 Linux that is why KVM was chosen over XEN for our private cloud [13].

C. Legacy migration

Legacy physical server could be migrated into OpenStack cloud using a persistent volume so that the migrated VM is backed by persistent storage. The function provided by the legacy server was the same afterwards and had the manageability benefits from running on an infrastructure platform.

Two things were created: a volume that would contain the data from legacy server and a 'working' VM as a way to work with that volume throughout the migration process. Horizon dashboard has been used to create volumes. The volume size had to be large enough to store all the file systems from legacy system, with appropriate growth, and included any swap-space that might be needed. The 'working' instance was then launched to set up the new volume. This instance had port 22 open under access and security, and ssh key pair was placed onto the system. Then, under 'volumes' the volume was attached to the working instance. The rsync was used to synchronize files between the legacy system and the new VM. The rsync tool was perfect for this for a few reasons: If interrupted, it can resume where it left off. Further, it can be throttled so that data can trickle over to the new volume without impacting the performance of the legacy server. For minimal downtime two separate rsyncs were done. The first one would sync over the vast majority of the data from the legacy server to the volume. The second one was done while both servers were down in order to sync the final changes from the legacy server right before the replacement VM was booted.

The next step was to make the volume bootable by adding a bootloader at the beginning of the volume. OpenStack provides firewalling, so we needed to set up a special security-group for VM that allowed all expected traffic to reach the server. This was done under "Access & Security" in horizon. Before committing to the change-over we had to verify that newly-created bootable volume did indeed work as expected. Once VM was setup correctly, the volume was unmount from the working VM, and legacy server was powered down. New VM system was finally launched.

D. Private cloud security measures

Dome9 secures OpenStack cloud servers and makes them virtually invisible to hackers. That automation closes firewall ports like RDP and SSH, and enables on-demand secure access with just one click. With Dome9, OpenStack Security Groups, centralizing policy management were automated within Dome9 Central. It has as well an SaaS management console for the entire cloud infrastructure. In addition, inter- and intra-group security rules gave ultimate flexibility and granularity. With Dome9 Cloud Connect, setting up Dome9 with OpenStack took less than a minute. Using Dome9 Account, we could add OpenStack supported region, and enter credentials. It was then connected to the private via API to manage all of Security Groups.

III. HYBRID CLOUD IMPLEMENTATION

Public clouds involve the use of third party servers where the user is typically charged on the usage basis. It helps cut the user's capital expenditure significantly, while providing the user with greater flexibility and scalability. However, the

advantages of the public cloud come at the cost of poorer performance and increased risks to data and applications. Private clouds attempt to solve the problem by providing cloud installation on-site with better performance and security, coming at increased capital expenditure and reduced flexibility. Therefore, hybrid clouds attempt to bridge the gap by providing the best of both worlds. Enterprises that use the hybrid cloud typically have a private cloud that handles the performance-sensitive core applications, while using the public cloud for scaling and non-core applications. For instance, the mail server and collaboration related components can be kept on the public cloud, while keeping the patient's database and large files in the private cloud. Another reason is that some applications are highly suited for public cloud, while some other legacy applications might not. Private clouds could be less robust than a public cloud managed by a reputed service provider. If the designer have a natural or manmade disaster attacking his hospital site, his private cloud infrastructure might become crippled. The designer can use the public cloud as a fail-over in that case. Thus, the designer might want to have a hybrid approach [14].

Although OpenStack integrates well with almost all IaaS providers, that doesn't stop our implementation from taking full advantages from those IaaS provider that don't support OpenStack. The key was to use a PaaS that could communicate with many different IaaS through a cloud interface. The cloud interface is the holygrail of cloud computing as it bridges the gap between different IaaS suppliers. On the higher level there is a unified API that could be called from a PaaS while on the lower level there are drivers specific to each IaaS and implemented with the vendor without compromising their codes. That cloud interface and the closely integrated PaaS are now reality with deltacloud and openshift thanks to the ever growing competition between cloud services providers to control the cloud computing market [15].

PaaS was initially conceived as a hosted solution for web applications. However the conceptual design forces the Cloud-compliant applications to accept several restrictions: (1) use the APIs exposed by PaaS owners; (2) use the specific programming paradigm that is adequate for the type of applications allowed by the PaaS; and (3) use the programming languages supported by the PaaS owner. These constraints are still valid for most current PaaS offers (e.g., Google AppEngine, Azure, Heroku, Duostack, XAP, Cast, CloudBees, and Stackato). A first step towards the developer freedom in building new PaaS independent applications was done by DotCloud which lets developers build their own software stack for a certain application. This solution is unfortunately not free and is currently based only on EC2 [16].

On the other hand, Deltacloud is an API developed by Red Hat and the Apache Software Foundation that abstracts differences between clouds. So, Deltacloud provides one unified REST-based API that can be used to manage services on any cloud. While each IaaS cloud is controlled through an adapter called "driver" and provides its own API. Drivers exist for the following cloud platforms: Amazon EC2, Fujitsu Global Cloud Platform, GoGrid, OpenNebula, Rackspace, OpenStack, RHEV-M, RimuHosting, Terremark and VMware vCloud. Next to the 'classic' front-end, it also offers CIMI and EC2

front-ends. Deltacloud is used in applications such as Aeolus to prevent the need to implement cloud-specific logic [17].

Deltacloud has some dependencies that need to be installed before its installation as Deltacloud server relies on a number of external rubygems and other libraries. However, once all the dependencies have been installed the Deltacloud server installation is done in one step [18].

OpenShift takes care of all the infrastructure, middleware, and management and allows the developer to focus on what they do best: designing and coding applications. It takes a No-Lock-In approach to PaaS by providing built-in support for Node.js, Ruby, Python, PHP, Perl, and Java. In addition, OpenShift is extensible with a customizable cartridge functionality that allows enterprising developers to add any other language they wish such as Clojure and Cobol. In addition to this flexible, no-lock-in, language approach, OpenShift supports many of the popular frameworks that make development easier including frameworks ranging from Spring, to Rails, to Play. OpenShift is designed to allow Developers to work the way they want to work by giving them the languages, frameworks and tools they need for fast and easy application development [19].

IV. THE MEDICAL IMAGE WEB APPLICATION

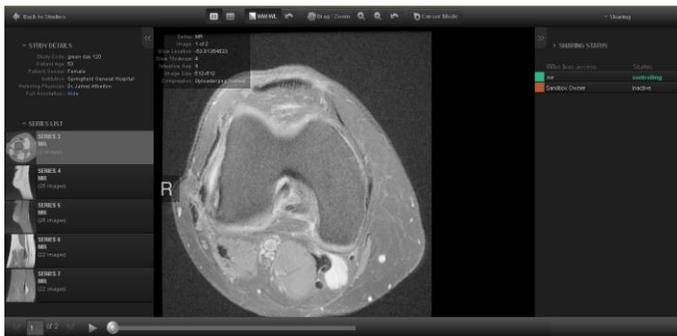


Fig. 4. Web application interface.

Healthcare generates a tremendous amount of data each day (CT, MRI, US, PET, SPECT, Mammography, X-Ray ... etc) and consumes quickly the hospitals storage space. For the reason of combining private and public cloud resources, a web application has been designed to store medical images on the private cloud using MongoDB which automatically migrate images older than one month to the public cloud. The subjects information were removed by stripping out the PHI (Protected Health Information) to conform to HIPAA standard and a sharding key was selected for our database that combines exam date and exam id. Compression, upload, delete, retrieval, and viewing were all integrated into the web application. Hibernate OGM; java and SWING were used to manage the backend on OpenShift with the ability to detect which device is trying to access the database. That allows smart phones as well as desktops to access the data for consultation purposes as shown in figure 4. We now detail the design of the Medical Image Web Application (MIWA) to guarantee the Atomicity, Consistency, Isolation, and Durability properties. Each of the properties is discussed individually [20].

A. Atomicity

The Atomicity property requires that either all operations of a transaction complete successfully, or none of them does. To ensure Atomicity, for each transaction issued, MIWA is using shards for actually storing data, mongos processes for routing requests to the correct data, and config servers, for keeping track of the cluster's state. As soon as an agreement to "COMMIT" is reached, the mongos processes can simultaneously return the result to the web application and complete the second phase.

B. Consistency

The consistency property requires that a transaction, which executes on a database that is internally consistent, will leave the database in an internally consistent state. Consistency is typically expressed as a set of declarative integrity constraints. We assume that the consistency rule is applied within the logic of transactions. Therefore, the consistency property is satisfied as long as all transactions are executed correctly.

C. Isolation

The Isolation property requires that the behavior of a transaction is not disturbed by the presence of other transactions that may be accessing the same data items concurrently. The MIWA decomposes a transaction into a number of sub-transactions, each accessing a single data item. Thus, the isolation property requires that if two transactions conflict on any number of data items, all their conflicting sub-transactions must be executed sequentially, even though the sub-transactions are executed in multiple mongos processes.

D. Durability

The Durability property requires that the effects of committed transactions cannot be undone and would survive server failures. In our case, it means that all the data updates of committed transactions must be successfully written back to the back-end cloud storage service. The main issue here is to support mongos processes failures without losing data. For performance reasons, the commit of a transaction does not directly update data in the cloud storage service but only updates the in-memory copy of data items in the shards. Instead, each mongos process issues periodic updates to the cloud storage service. During the time between a transaction commit and the next checkpoint, durability is ensured by the replication of data items across several shards. After checkpoint, we can rely on the high availability and eventual consistency properties of the cloud storage service for durability.

E. Security

Cryptographic modules supplied by HP Atalla were used in the medical image web application to encrypt healthcare data and reduce the risk of data encryption and reputation damage without sacrificing performance using high-performance hardware security modules. Those Data Security solutions meet the highest government and financial industry standards-including NIST, PCI-DSS and HIPAA/HITECH-protect sensitive data and prevent fraud. HP Enterprise Secure Key Manager (ESKM) and Atalla Network Security Processors (NSP) provided robust security, high performance and

transparency while ensuring comprehensive, end-to-end network security.

V. SIMULATION

Evaluation of alternative designs or solutions for Cloud computing on real test-beds is not easy due to several reasons. Firstly, public Clouds exhibit varying demands, supply patterns, system sizes, and resources (hardware, software, network) [19]. Due to such unstable nature of Cloud resources, it is difficult to repeat the experiments and compare different solutions. Secondly, there are several factors which are involved in determining performance of Cloud systems or applications such as user's Quality of Service (QoS) requirements, varying workload, and complex interaction of several network and computing elements. Thirdly, the real experiments on such large-scale distributed platforms are considerably time consuming and sometimes impossible due to multiple test runs in different conditions. Therefore, a more viable solution is to use simulation frameworks which will enable controlled experimentation, reproducible results and comparison of different solutions in similar environments. Despite the obvious advantages of simulation in prototyping applications and developing new scheduling algorithms for Cloud computing, there are a few simulators for modeling real Cloud environments. For evaluating a scheduling algorithm in a Cloud computing environment, a simulator should allow users to define two key elements: (i) an application model specifying the structure of the target applications in Clouds, typically in terms of computational tasks and data communication between tasks; (ii) a platform model of Cloud computing data centers specifying the nature of the available resources and the network by which they are interconnected. Clouds currently deploy wide variety of applications both from industrial enterprises and scientific community [21]. In terms of the platform, Cloud computing is quite different from traditional distributed computing platforms defined by service-oriented features such as resource elasticity, multiple-level of services and multi-tenancy of resources.

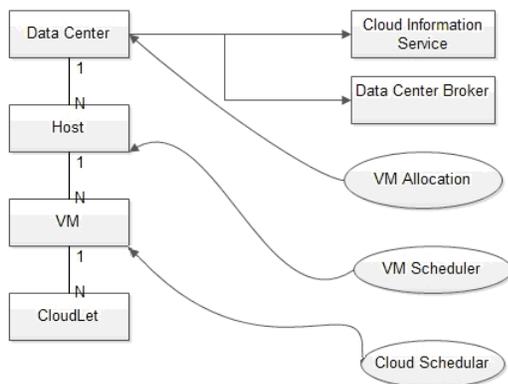


Fig. 5. The main parts and relations of CloudSim.

The experiments in this research were performed on the CloudSim cloud simulator which is a framework for modeling and simulating the cloud computing infrastructures and services [22]. The CloudSim simulator has many advantages: it can simulate many cloud entities, such as datacenter, host and broker. It can also offer a repeatable and controllable

environment. And we do not need to take too much attention about the hardware details and can concentrate on the algorithm design. The simulated datacenter and its components can be built by coding and the simulator is very convenient in algorithm design [23]. The main parts which relate to the experiments in this research and the relationship between them are shown in Figure 5 while the functions of those components are explained in table 2.

TABLE II. CLOUDSIM COMPONENTS AND THEIR FUNCTIONS [24]

CloudSim Component	Function
Cloud Information Service	It is an entity that registers, indexes and discovers the resource.
Datacenter	It models the core hardware infrastructure, which is offered by Cloud providers.
Datacenter Broker	It models a broker, which is responsible for mediating negotiations between SaaS and Cloud providers.
Host	It models a physical server.
Vm	It models a virtual machine which is run on Cloud host to deal with the cloudlet.
Cloudlet	It models the Cloud-based application services.
VmAllocation	A provisioning policy which is run in datacenter level helps to allocate VMs to hosts.
VmScheduler	The policies required for allocating process cores to VMs. It is run on every Host in Datacenter.
CloudletScheduler	It determines how to share the processing power among Cloudlets on a virtual machine. It is run on VMs.

The application simulates an IaaS provider with an arbitrary number of datacenters. Each datacenter is entirely customizable. The user can easily set the amount of computational nodes (hosts) and their resource configuration, which includes processing capacity, amount of RAM, available bandwidth, power consumption and scheduling algorithms. The customers of the IaaS provider are also simulated and entirely customizable. The user can set the number of virtual machines each customer owns, a broker responsible for allocating these virtual machines and resource consumption algorithms. Each virtual machine has its own configuration that consists of its hypervisor, image size, scheduling algorithms for tasks (here known as cloudlets) and required processing capacity, RAM and bandwidth.

The simulation scenario models a network of a private and a public cloud (HP's cloud). The public and the private clouds were modeled to have two distinct data centers. A CloudCoordinator in the private data center received the user's applications and processed (queue, execute) them in a FCFS basis. To evaluate the effectiveness of a hybrid cloud in speeding up tasks execution, two test scenarios were simulated: in the first scenario, all the workload was processed locally within the private cloud. In the second scenario, the workload (tasks) could be migrated to public clouds in case private cloud resources (hosts, VMs) were busy or unavailable. In other words, second scenario simulated a CloudBurst by integrating the local private cloud with public cloud for handing peak in service demands. Before a task could be submitted to a public cloud (HP), the first requirement was to load and instantiate the VM images at the destination. The number of images

instantiated in the public cloud was varied from 10% to 100% of the number of hosts available in the private cloud. Task units were allocated to the VMs in the space-shared mode. Every time a task finished, the freed VM was allocated to the next waiting task. Once the waiting queue ran out of tasks or once all tasks had been processed, all the VMs in the public cloud were destroyed by the CloudCoordinator. The private cloud hosted 20 machines. Each machine had 2 GB of RAM, 10TB of storage and one CPU run 1000 MIPS. The virtual machines created in the public cloud were based on an HP's small instance (2 GB of memory, 2 virtual cores, and 60 GB of instance storage). We considered in this evaluation that two virtual cores of a small instance has the same processing power as the local machine. The workload sent to the private cloud was composed of 2,000 tasks. Each task required between 20 and 22 minutes of processor time. The distributions for processing time were randomly generated based on the normal distribution. Each of the 2,000 tasks was submitted at the same time to the private cloud.

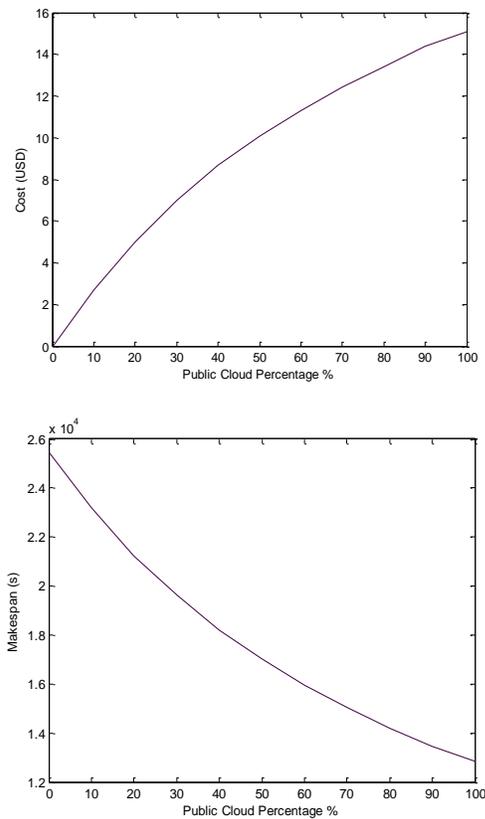


Fig. 6. The relation between cost in USD and the public cloud percentage (up). The relation between makespan in seconds and public cloud percentage (down).

Figure 6 shows the makespan of the tasks that were achieved for different combination of private and public cloud resources. The pricing policy was designed based on the HP's small instances (US\$ 0.042 per instance per hour) business model. It means that the cost per instance is charged hourly. Thus, if an instance runs during 1 hour and 1 second, the amount for 2 hours (US\$ 0.084) will be charged [25].

This experiment showed that the adoption of a hybrid public/private Cloud computing environments could improve productivity of the healthcare organization. With this model, organizations can dynamically expand their system capacity by leasing resources from public clouds at a reasonable cost.

VI. DISCUSSION AND CONCLUSION

Cloud computing is quickly becoming a dominant model for end-users to access centrally managed computational resources. Through this work in extending OpenStack, we have demonstrated the feasibility of providing healthcare users with access to heterogeneous computing resources using a hybrid cloud computing model.

Open cloud computing is not only a low cost choice for implementation but also provides the designer with a vast number of choices. That low cost solution for resources allocation can solve both the limited storage space and the lack of full utilization of computer infrastructure that healthcare always suffered from. We have to be aware that users' requirements may be very different and so the optimal infrastructure will vary. The ability to select suitable resources from different cloud providers can increase performance and lower cost considerably. That was achieved using Deltacloud and OpenShift which offers a communication layer between web applications, mobile applications and the different cloud providers. Simulation was an important step to measure the feasibility of our design and showed better makespan when public cloud took a higher workload share.

REFERENCES

- [1] David Villegas, Norman Bobroff, Ivan Rodero, Javier Delgado, Yanbin Liu, Aditya Devarakonda, Liana Fong, S. Masoud Sadjadi, Manish Parashar, "Cloud federation in a layered service model," *Journal of Computer and System Sciences*, Volume 78, Issue 5, September 2012, Pages 1330-1344.
- [2] Pierre Boiron, Valère Dussaux, "Healthcare Software as a Service: the greater Paris region program experienc The so-called "Région Sans Film" program," *Enterprise Distributed Object Computing Conference Workshops (EDOCW)*, 2011 15th IEEE International, 29 Aug.- 2 Sep. 2011, pp. 247-251, Paris, France.
- [3] P. Sempolinski and D. Thain, "A comparison and critique of Eucalyptus, OpenNebula and Nimbus," in *Proceedings of the 2010 IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom)*, 2010.
- [4] Ivo Jansch and Vito Chin. *Php/architect's Guide to PHP Development in the Cloud*. Blue Parabola, Canada, 2011.
- [5] Trieu C. Mohindra, Ajay Karve, Alexei a. Segal, Alla Chieu, "Dynamic Scaling of Web Applications in a Virtualized Cloud Computing Environment," in *2009 IEEE International Conference on e-Business Engineering*, 2009, pp.: 281-286.
- [6] Perera, S.; Kumarasiri, R.; Kamburugamuva, S.; Fernando, S.; Weerawarana, S.; Fremantle, P., "Cloud Services Gateway: A Tool for Exposing Private Services to the Public Cloud with Fine-grained Control," *Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW)*, 2012 IEEE 26th International, 2012 , pp.: 2237 - 2246.
- [7] F. Quesnel and A. L`ebre, "Cooperative dynamic scheduling of virtual machines in distributed systems," in *Proceedings of the 6th Workshop on Virtualization in High-Performance Cloud Computing (VHPC '11) Euro-Par 2011, Bordeaux, France, Aug. 2011*.
- [8] new extreme low-energy server technology. (n. d.). Retrieved from <http://h17007.www1.hp.com/us/en/iss/110111.aspx>
- [9] Open source software for building private and public clouds. (n. d.). Retrieved from <http://www.openstack.org>

- [10] Corradi, A., Foschini, L., Povedano-Molina, J., Lopez-Soler, J.M., "DDS-Enabled Cloud Management Support for Fast Task Offloading," Computers and Communications (ISCC), 2012 IEEE Symposium on, 1-4 July 2012, pp.: 67 - 74, 2012.
- [11] Crago, S.; Dunn, K.; Eads, P.; Hochstein, L.; Dong-In Kang; Mikyung Kang; Modium, D.; Singh, K.; Jinwoo Suh; Walters, J.P., "Heterogeneous Cloud Computing," Cluster Computing (CLUSTER), 2011 IEEE International Conference on, pp.:378 - 385.
- [12] RabbitMQ - Messaging that just works (n. d.). Retrieved from <http://www.rabbitmq.com>
- [13] Kernel Based Virtual Machine (n. d.). Retrieved from <http://www.linux-kvm.org>
- [14] J. Tordsson, R. S. Montero, R. Moreno-Vozmediano, and I. M. Llorente, "Cloud brokering mechanisms for optimized placement of virtual machines across multiple providers," Future Generation Comp. Syst., vol. 28, no. 2, pp. 358–367, 2012.
- [15] J. Lucas-Simarro, R. Moreno-Vozmediano, R. Montero, and I. Llorente, "Dynamic placement of virtual machines for cost optimization in multicloud environments," in Proceedings of the 2011 International Conference on High Performance Computing & Simulation (HPCS 2011), July 2011, pp. 1–7.
- [16] dotCloud - One home for all your apps (n. d.). Retrieved from <https://www.dotcloud.com>
- [17] The Aeolus Project (n. d.). Retrieved from <http://www.aeolusproject.org>
- [18] Deltacloud (n. d.). Retrieved from <http://deltacloud.apache.org>
- [19] Develop and scale apps in the cloud (n. d.). Retrieved from <https://openshift.redhat.com>
- [20] Genqiang Gu, Qingchun Li, Xiaolong Wen, Yun Gao, Xuejie Zhang, "An overview of newly open-source cloud storage platforms," 2012 IEEE International Conference on Granular Computing (GrC-2012), August 2012, pp. 142-147.
- [21] Inçki, K., Ari, I., Sozer, H., "A Survey of Software Testing in the Cloud", Software Security and Reliability Companion (SERE-C), 2012 IEEE Sixth International Conference on, pp.: 18 - 23.
- [22] Wei Liu, Feiyan Shi, Wei Du and Hongfeng Li "A Cost-Aware Resource Selection for Data intensive Applications in Cloud-oriented Data Centers" in the International Journal of Information Technology and Computer Science 2011.
- [23] The Cloud Computing and Distributed Systems (CLOUDS) Laboratory, University of Melbourne <http://www.cloudbus.org/cloudsim>
- [24] Yuxiang Shi; Xiaohong Jiang and Kejiang Ye, "An Energy-Efficient Scheme for Cloud Resource Provisioning Based on CloudSim," Cluster Computing (CLUSTER), 2011 IEEE International Conference on, 26-30 Sept. 2011, pp.: 595 - 599.
- [25] Wickremasinghe, B.; Calheiros, R.N.; Buyya, R., "CloudAnalyst: A CloudSim-Based Visual Modeller for Analysing Cloud Computing Environments and Applications," Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on Communication, Networking & Broadcasting ; Computing & Processing (Hardware/Software), 2010 , pp.: 446 – 452.
- [26] HP Cloud Pricing (27 Jan. 2013). Retrieved from <https://www.hpcloud.com/pricing#Compute>
- [27] J. Diaz, G. v. Laszewski, F. Wang, and G. Fox, "Abstract Image Management and Universal Image Registration for Cloud and HPC Infrastructures," IEEE Cloud. 2012.
- [28] Nordin, M.I., Amin, A.H.M., Shah, S.N.M., "Agent based Resource Broker for medical informatics application in clouds", Computer & Information Science (ICCIS), 2012 International Conference on, pp.: 802 – 807, Volume: 2, 12-14 June 2012.
- [29] C. Quinton, R. Rouvoy, and L. Duchien, "Leveraging Feature Models to Configure Virtual Appliances," in 2nd International Workshop on Cloud Computing Platforms (CloudCP'12), (Bern, Switzerland), Apr. 2012.
- [30] M. Rahman , R. Thulasiram , P. Graham, "Differential time-shared virtual machine multiplexing for handling QoS variation in clouds," Proceedings of the 1st ACM multimedia international workshop on Cloud-based multimedia applications and services for e-health, Nov. 2012, Nara, Japan.
- [31] B. Nicolae, G. Antoniu, L. Boug'e, D. Moise, and A. Carpen-Amarie, "BlobSeer: Next-generation data management for large scale infrastructures," J. Parallel Distrib. Comput., vol. 71, pp. 169–184, February 2011.
- [32] Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AHA, Conteh L, Kirkwood BR, Meek SR, "Mobile Health (mHealth) Approaches and Lessons for Increased Performance and Retention of Community Health Workers in Low- and Middle-Income Countries: A Review," J Med Internet Research, 2013;15(1).
- [33] Mell, P.; Grance, T. (2010), "The NIST Definition of Cloud Computing," Communications of the ACM, Vol. 53 (2010) No. 6, pp. 50-50.
- [34] von Laszewski, G., Diaz, J., Fugang Wang, Fox, G.C., "Comparison of Multiple Cloud Frameworks," Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on Communication, Networking & Broadcasting ; Components, Circuits, Devices & Systems ; Computing & Processing Hardware/Software, 2012 , pp.: 734 - 741.
- [35] Lori M. Kaufman. Data security in the world of cloud computing. IEEE Security and Privacy, 7:61–64, 2009.
- [36] Youseff, L.; Butrico, M.; Da Silva, D. (2008), "Toward a Unified Ontology of Cloud Computing," the Grid Computing Environments Workshop, 2008. GCE '08, Austin, Texas, USA, pp. 1-10.

A Novel Algorithm for Improving the ESP Game

Mohamed Sakr

Department of Computer Science
Faculty of Computers and
Information, Menoufia University
Shebin El-Kom, Egypt

Hany Mahgoub

Department of Computer Science
Faculty of Computers and
Information, Menoufia University
Shebin El-Kom, Egypt

Arabi Keshk

Department of Computer Science
Faculty of Computers and
Information, Menoufia University
Shebin El-Kom, Egypt

Abstract—one of the human-computation techniques is games with a purpose (GWAP) and microtask crowdsourcing. These techniques can help in making the image retrieval (IR) be more accurate and helpful. It provides the IR system's database with a rich of information by adding more descriptions and annotations to images. One of the systems of human-computation is ESP Game. ESP Game is a type of games with a purpose. In the ESP game there has been a lot of work was proposed to solve many of the problems in it and make the most benefit of the game. One of these problems is that the ESP game neglects player's answers for the same image that don't match. This paper presents a new algorithm to use neglected data to generate new labels for the images. We deploy our algorithm at the University of Menoufia for evaluation. In this trial, we first focused on measuring the total number of labels generated by our Recycle Unused Answers For Images algorithm (RUAI). In our evaluation of the RUAI algorithm we present a new evaluation measure we called it quality of labels measure. This measure identifies the quality of the labels in compared to the pre-qualified labels. The results reveal that the proposed algorithm improved the results in compared to the ESP game in all cases.

Keywords—ESP game; Games with a purpose; Human computation; crowdsourcing

I. INTRODUCTION

There are problems which are difficult to be processed by computers such as those related to artificial intelligence. These problems are easy to be solved by the human brain power. Human computation is the idea of solving difficult problems using human intelligence. Some of these problems are related to artificial intelligence (AI) or image recognition.

Games with a purpose (GWAP) are one of the human computation [1,2]. GWAP are a way to make useful of the human desire to be entertained. Several GWAP systems have been proposed for image annotation and commonsense reasoning. Von Ahn and Dabbish [3] classified GWAP into three game-structure templates that generalize successful instances of human computation games: output-agreement games, inversion-problem games, and input-agreement games. Yuen et al. [4] added output-optimization game to these three templates.

ESP game is one of the GWAP systems. ESP game was the first systems to clarify the advantages of using human computation and GWAP systems. It is example of output-agreement games and is a two player's game for labeling images [5]. Barnard et al. [6] reported that labeling images has proven to be a hard problem for computer vision, but it is something that humans can do easily. It has been shown that

the image labels collected through the ESP game are usually of good quality. Moreover, the game results allow more accurate image retrieval, help users block inappropriate images (e.g., pornographic content), and improve web accessibility (e.g., the labels can help visually impaired people surf web pages [7]). In order to humans to label images there must be some sort of motivation. One type of motivation is entertainment, which is achieved in the ESP game. In the ESP game the players are chosen randomly and are assigned the same image. Each player doesn't know the other player and the two players can't communicate with each other. The only thing they have in common is the image that they play with. Each player is asked to give description to that image and has to guess what the other player is typing for each image to win the game and go to the next image. Once the two players have entered the same word, this word becomes the label for the image. The easiest way for both players to type the same string is by typing something related to the common image. The round lasts for 2.5 minutes.

During the round the players try to describe as many images as they can. The players get number of points for each image they label. If the players agree on 15 images they get a large number of bonus in points. Once there is a difficult *image* that the players can't agree on they both can press the Pass button. The game is attached with a scoreboard, with the names of players with the highest scores. Empirical studies of other peer-production systems have shown that points are a key feature in motivating users [8].

One type of GWAP systems and output-agreement games is ESP game. During the play in the ESP game it appears anecdotally that people coordinate on the same words, but the other words are neglected. In this work we are concentrating on ESP game and on solving one problem of the ESP Game that the player's answers for the same image that don't match in the same game are neglected. This paper presents a new algorithm to use these neglected data to generate new labels for the images.

The rest of this paper is organized as follows. Section II presents a review of related works on problems of the ESP game. Section III presents our algorithm to solve the problems of neglected data in ESP game. The results and simulation analysis of our proposed algorithm are presented in section IV. Section V provides conclusions and future work.

II. RELATED WORK

ESP game is one of the successful applications of the games that harvest human intelligence and time to solve tasks,

which is difficult by computer. In this work we will show that, although the idea underlying the game is an extremely powerful one, more care needs to be taken in the design as the game uses only the answers of the players that match and neglects the other answers.

After analyzing the ESP game we notice the following problems:

- Informative labels: many of the labels from the ESP game are redundant and not very informative (“man” and “guy”). Many labels are generic and not descriptive (“building” and “terraced house”). Many labels can be expected and generated automatically (“water”, “blue”, “sky” and “clouds” are all related)
- How to measure the system’s productivity: Test if the system is productive and give informative labels with good quality and acceptable quantity.
- How to select the next puzzle to play with: How to select the next puzzle (next image to play with) select one with most descriptions or with least descriptions

In the ESP game there has been a lot of work was proposed to solve many of the previous problems in it and make the most benefit of the game. In the next paragraph we try to present some of the previous work that is proposed to solve the problems of the ESP game.

Weber et al. [9] notice that the ESP game failed to collect informative labels so they proposed a language model to generate probabilities to the next labels to be added given the pre-added labels as training data. Chen et al. [10] proposed anew metric called system gain, use analysis to study the properties GWAP systems and implemented a new puzzle selection strategy to improve the GWAP systems. Jain and Parkes [11] presented game theoretic analysis for the ESP game, and they investigated the equilibrium behavior under different incentive mechanisms and provided guidelines to design incentive mechanisms. Von Ahn and Dabbish [3] suggested a set of evaluation metrics, such as throughput, lifetime play, and expected contribution, to determine whether ESP-like GWAP systems are successful. Ho et al. [12] also notice that the set of labels determined from the ESP game for an image, are not very diverse, and develop a three-player version of the ESP game that involves the addition of a “blocker” to type in words that the other two players cannot use to match. In this work we address the informative labels problem and how to generate new labels with no need to extra un-useful game rounds between players.

III. RECYCLE UNUSED ANSWERS FOR IMAGES ALGORITHM (RUAI)

After analyzing the ESP game, the previous problems and there solutions, we found that in some time when the players play the game they enter informative labels and these labels when they are not agreed upon they are neglected and trashed. So the ESP game throws away the unused answers. In this section, we present the (RUAI) algorithm which recycles the player’s answers to make use of these informative answers in the situations where they are neglected as shown in Fig.1 and Fig.2.

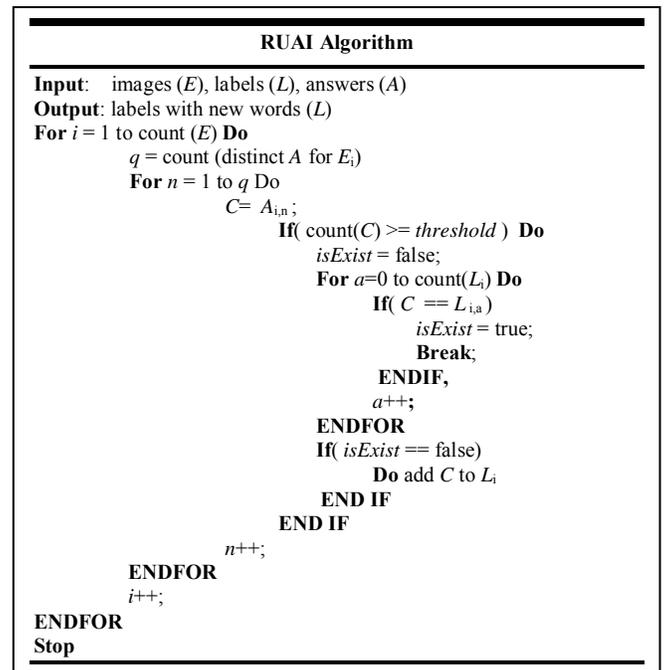


Fig. 1. The RUAI Algorithm.

There are two types of algorithms offline and online. Online algorithm runs during the game while the players are playing. Offline algorithm runs after the players finish playing the game. The RUAI algorithm is categorized as offline algorithm as it runs on the data from the players after they play the game not during the game (the algorithm starts to run after the players finish playing). Also, the algorithm runs on all the players and their data not only on the two players of the game. The scenario of algorithm work is done as follow:

- Step 1: get all the images, its corresponding answers and its labels from the database.
- Step 2: for each image get all the distinct answers.
- Step 3: for each answer calculate the count of its occurrence and test it against a given number which represent the number of players that agree on that answer. This number ranges from 2 to m (threshold). Threshold is decided by the user when searching for a given query in the database. When the threshold is increased it means that the resulted images will be more relevant to the search query (give me the images related to the query Q with accuracy X where Q is the query that the user entered and X is the threshold).
- Step 4: if the count is bigger or equal than the threshold we will check if the answer is in the labels for that images if no we will insert the answer as a new label for that image. If yes go to the next answer.
- Step 5: if the count is smaller than the threshold we will go to the next answers.
- Step 6: after iterating between all the answers for this image we will go to the next image and redo the steps from 2 to 5.

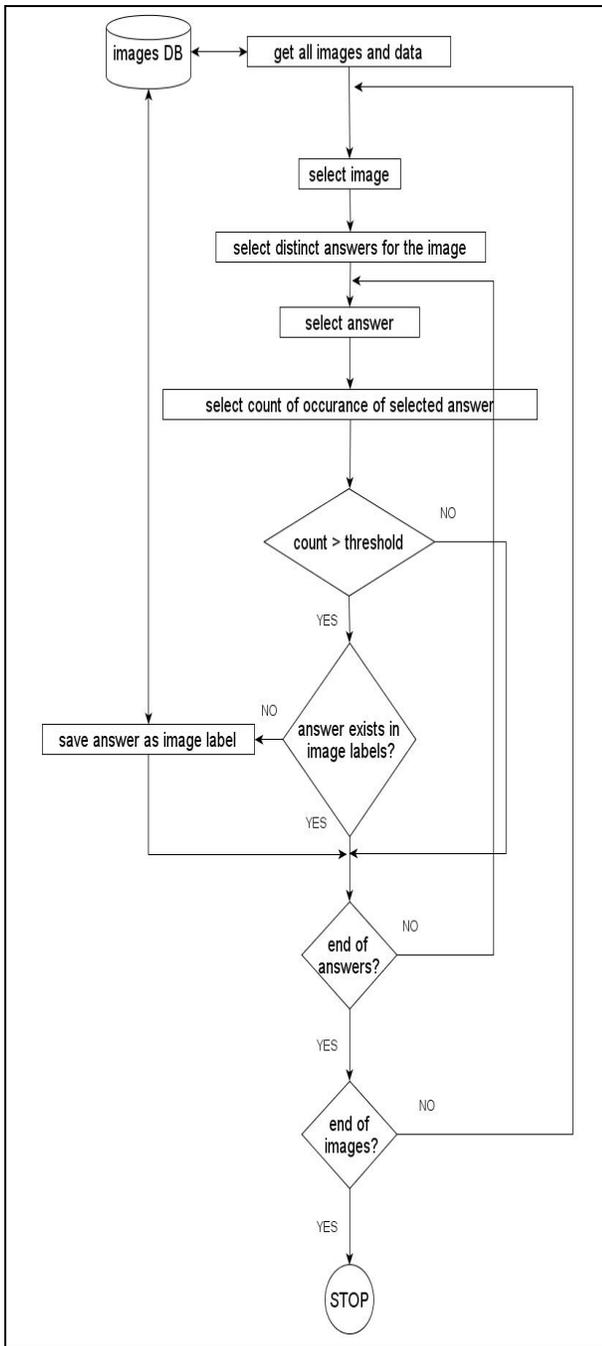


Fig. 2. Flowchart of the RUAI algorithm.

First we will get all the images E_i , all its corresponding answers and all its labels from the database. Second, for each image E_i we will get all distinct answers A_n we will calculate the count of its occurrence. After that, we will test the count if it is bigger or equal the threshold or not. If yes we will look through the labels L for the image E_i to check if that answer is exist or not. If it doesn't exist it will be inserted to the labels L of image E_i then go to the next answer, repeat the previous steps and so on until the end of all image's E_i answers. BUT if it exists it will go to the next answer for image E_i , repeat the previous steps and so on until the end of all

image's E_i answers After the end of all image's E_i answers it will go to the next image, repeat the steps and so on until the end of images E .

A. Case study of the RUAI algorithm

This case study illustrates the work of the (RUAI) algorithm. Suppose there are four players P1, P2, P3 and P4 that are playing the game as shown in Fig.3. P1 plays with P2 and P3 play with P4. The four players are playing on the same image I1. P1 entered the words B and D. P2 entered the words A, B and C. As the ESP game the label for the image will be B and the answers A, C and D will be neglected. The other game is between P3 and P4. P3 enters the words Z and D. P4 enters the words A and Z. As the ESP game the label for the image will be Z and the answers A and D will be neglected. From the ESP Game the labels for the image I1 will be B and Z as shown in Fig.4 (a). Now we will perform our algorithm which will iterate on every image as shown in Fig.4.

In this case study the RUAI algorithm would be performed on I1. For each distinct answers on that image which will be (A, B, C, D and Z) test count of occurrence of each distinct answer against the threshold for now it will be 2. Count of A, B, C, D and Z will be 2, 2, 1, 2 and 2 as shown in Fig.4 (b). The labels A, B, D and Z all their counts are bigger than or equal threshold 2 so all will be labels for the image. But B and Z are already inserted as labels by the ESP game so the new labels A and D will be inserting into the database as shown in Fig.4(c) (d).

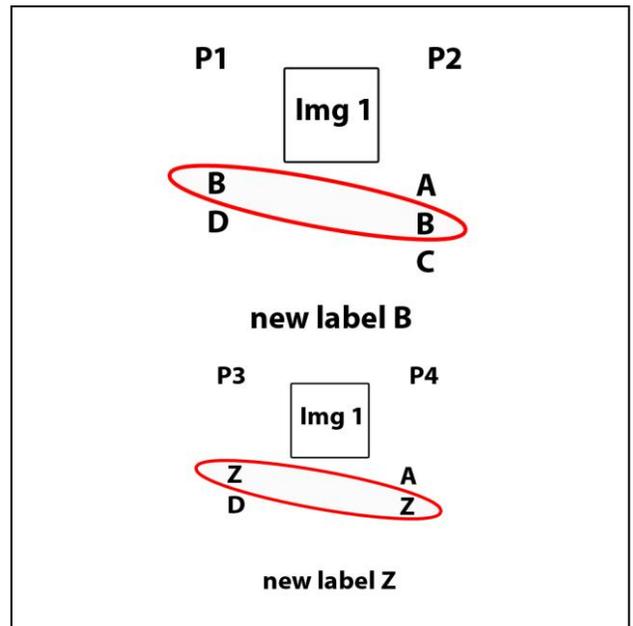


Fig. 3. Case study Game scenario

IV. RESULTS AND SIMULATION ANALYSIS

At the beginning we start to get a data-set of images and labels to work on. A number of crowdsourcing data-sets are available for research. For example, von Ahn et al. contributed a list of 100,000 images with English labels from their ESP Game [13]. We used von Ahn et al. data-set. First, we integrate only the images to our system and not the labels.

Image	labels
<i>I1</i>	<i>B,Z</i>

(a) The image and its labels from the ESP game

Distinct answer	count
A	2
B	2
C	1
D	2
Z	2

(b) Distinct answers for the image I1 and the count of each answer.

Distinct answer	count
A	2
D	2

(c) Distinct answers for the image I1 with count >=2(threshold) and not in fig.4(a)

Image	labels
I1	B,Z,A,D

(d) image and new labels after run RUAI algorithm

Fig. 4. RUAI data table

We deploy our system at the University of Menoufia for evaluation using Java 1.6 as programming language, Mysql [14] server as database management server and run the server over Intel core i7 with 4 GB Ram PC on windows 7-64 operating system.

In this trial, we first focused on measuring the total number of labels generated by our RUAI algorithm. We sent emails to the staff of the University. We advertised the system as a free game and provided a link to the system. Sixteen users signed up initially (many of them were research students and demonstrators). The results of the proposed algorithm and the ESP game are presented in Table 1.

The results show that the total number of images that were described is 56 images. Total number of answers users entered was 736 answers. Total number of labels that our prototype of the ESP Game generated was 155 labels. Total numbers of labels that our proposed algorithm generated were 198 labels. So our algorithm generated new 43 labels of the images.

TABLE I. THE RESULTS OF THE RUAI ALGORITHM AND THE ESP GAME FOR SIXTEEN USERS.

Method	Total Images	Total Answers	Total Labels
ESP Game	56	736	155
RUAI algorithm	56	736	198

In our evaluation of the RUAI algorithm we present a new evaluation measure we called it *quality of labels measure*. This measure identifies the quality of the labels compared to pre-qualified labels.

To compute the quality of labels measure we first compare the labels results from our RUAI algorithm with the labels in von Ahn et al. data-set and compute the total number of labels resulted from the RUAI algorithm which is exists in von Ahn et al. data-set then calculate the percentage of them. This computation is done for each image then at the end we compute the average. The mathematical formula of the quality of labels measure is illustrated as in (1)

$$Quality\ of\ labels\ measure = \frac{\sum_{i=1}^N \frac{length(L_{iRUAI} \cap L_{idata-set})}{length(L_{idata-set})}}{N} \quad (1)$$

$L_{idata-set}$: Set of labels for image *i* from von Ahn et al. data-set

L_{iRUAI} : Set of labels for image *I* from RUAI algorithm

N : Total number of images

Length: is a function to calculate the length of a given set

The algorithm result's shows that about 78% of labels in von Ahn et al. data-set were generated by our RAUI algorithm. We noticed that in the previous work of evaluating the quality of the labels is done manually by asking group of people to describe images and compare the results or by giving them questions for a given image to know if the labels for that image are correct. In this paper the evaluation of the label's quality is done automatically by using the quality of labels measure formula. The advantages of using this formula is to reduce the time and cost.

Due to the time limitation, we did not observe the users for longer period to give more answers. However, we believe if we deploy the system to a larger demographic, our algorithm would produces even more promising results fuelled by the network effect.

V. CONCLUSION

This paper presents a new algorithm to generate new labels for the images with no need to extra game rounds. The algorithm overcame some of the problems in the ESP game by using neglected player's answers for the same image that don't match. Also we present a new evaluation measure which called quality of labels measure. This measure identifies the quality of the labels compared to pre-qualified labels. The using of this measure improved the time and saved the cost. The results of comparing RUAI algorithm and the ESP game reveal that the RUAI algorithm is much better than the ESP game in all cases.

In the future work we intend to generate new labels for the images using the data mining techniques. Furthermore we intend to evaluate the performance of our algorithm in terms of its efficiency and scalability.

REFERENCES

- [1] C. Harris and P. Srinivasan, "Comparing Crowd-Based, Game-Based, and Machine-Based Approaches in Initial Query and Query Refinement Tasks," *Lecture Notes in Computer Science*, vol. 7814, pp. 495-506, 2013.
- [2] S. Thaler, E. Simperl and S. Wolger, "An Experiment in Comparing Human-Computation Techniques," *Internet Computing*, IEEE, vol. 16(5), pp. 52-58, October 2012.
- [3] L. Ahn and L. Dabbish, "Designing games with a purpose," *Communications of the ACM*, vol. 51(8), pp. 58-67, August 2008.
- [4] M. Yuen, L. Chen and I. King, "A Survey of Human Computation Systems," *Proceeding CSE '09 Proceedings of the 2009 International Conference on Computational Science and Engineering*, vol. 4, pp. 723-728, August 2009.
- [5] L. Ahn and L. Dabbish, "Labeling images with a computer game," *Proceeding CHI '04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 319-326, 2004.
- [6] K. Barnard, P. Duygulu, D. Forsyth, N. Freitas, D. Blei and M. Jordan, "Matching Words and Pictures," *Journal Of Machine Learning Research*, vol.3, pp. 1107-1135, 2003.
- [7] J. Bigham, R. Kaminsky, R. Ladner, O. Danielsson and G. Hempton, "WebInSight : Making Web Images Accessible," *Proceeding Assets '06 Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, pp. 181 - 188, 2006.
- [8] K. Nam, M. Ackerman and L. Adamic, "Questions in , Knowledge iN ? A Study of Naver ' s Question Answering Community," *Conference on Human Factors in Computing Systems, CHI 09, ACM*, April 2009.
- [9] S. Robertson, M. Vojnovic and I. Weber, "Rethinking the ESP Game," *Proceeding CHI EA '09 CHI '09 Extended Abstracts on Human Factors in Computing Systems*, pp. 3937-3942, September 2009.
- [10] L. Chen, B. Wang and K. Chen, "The Design of Puzzle Selection Strategies for GWAP Systems," *Journal Concurrency and Computation: Practice & Experience*, vol. 22(7), pp. 890-908, May 2010.
- [11] S. Jain and D. Parkes, "A Game-Theoretic Analysis of Games with a Purpose," *Proceeding WINE '08 Proceedings of the 4th International Workshop on Internet and Network Economics*, pp. 342-350, 2008.
- [12] C. Ho, T. Chang, J. Lee, J. Hsu and K. Chen, "KissKissBan: a competitive human computation game for image annotation," *Proceeding HCOMP '09 Proceedings of the ACM SIGKDD Workshop on Human Computation*, pp. 11-14, 2009.
- [13] ESP Game dataset:
<http://server251.theory.cs.cmu.edu/ESPGame100k.tar.gz>
- [14] MySQL Server <http://dev.mysql.com/downloads/mysql/>

A quadratic convergence method for the management equilibrium model

Jiayi Zhang
Feixian School
Linyi University
Feixian, Shandong, P.R.China

Abstract—in this paper, we study a class of methods for solving the management equilibrium model. We first give an estimate of the error bound for the model, and then, based on the estimate of the error bound, propose a method for solving the model. We prove that our algorithm is quadratically convergent without the requirement of existence of a non-degenerate solution.

Keywords—Management equilibrium model; estimation of error bound; algorithm; quadratic convergence

I. INTRODUCTION

The management equilibrium model seeks a vector $(x^*, y^*) \in R^{2n}$ such that

$$x^* \geq 0, y^* \geq 0, (x^*)^T y^* = 0, Mx^* - Ny^* = Qz^* + q, \quad (1)$$

Where $M, N \in R^{m \times n}$, $Q \in R^{m \times l}$, $q \in R^m$, and there exists $z^* \in R^l$. The model originated from equilibrium problems in economic management, engineering, etc. Applications of complementarity problems from the field of economics include general Walrasian equilibrium, spatial price equilibria, invariant capital stock, market equilibrium, optimal stopping, and game-theoretic models. In engineering, the complementarity problems also plays a significant role in contact mechanics problems, structural mechanics problems, obstacle problems mathematical physics, Elastohydrodynamic lubrication problems, traffic equilibrium problems (such as a pathbased formulation problem, a multicommodity formulation problem, network design problems), etc. [1,2] For example, the equilibrium of supply and demand in an economic system is often depicted as a complementary model between two decision variables. As another example, the typical Walras' Law of competition equilibrium in economic transactions can also be converted to complementary model between price and excess demand [3].

Recently, many effective methods have been proposed to solve (1) [4-6]. The basic idea of these methods is to convert (1) into an unconstrained or a simply constrained optimization problem. As we known, if the Jacobian matrix at a solution to (1) is non-singular, then it is guaranteed that the Levenberg-Marquardt (L-M) algorithm is quadratically convergent [5,6]. Lately, Yamashita and Fukushima have proved that the condition for the local error bound to hold is weaker than the

non-singularity of the Jacobian matrix [7]. This motivates the establishment of an error bound for (1). The establishment of LCP error bound has been extensively studied (see literature review [8]). For example, Mangasarian and Ren have given an error bound under the R_0 -matrix condition [9]. Clearly, (1) is a generalization of LCP, which prompts whether or not the LCP error bound can be generalized to (1). For this reason, we focus on the establishment of an error bound for (1), design a smooth algorithm for solving (1) using the error bound, and analyze the convergence of the algorithm as well as the rate of convergence.

In section 2, we give primarily an equivalent conversion of (1). In section 3, using a new residual function, we establish an error bound for (1) under more general conditions. In section 4, based on the established error bound, we propose a smooth algorithm for solving (1), and prove that the given algorithm is quadratically convergent without the requirement of existence of a non-degenerate solution. Compared with the convergence of algorithms in [5,6], the condition is weaker.

Now we give some notations. The inner product of vectors $x, y \in R^n$ is written as $x^T y$. Let $\|\cdot\|$ be the Euclidean norm. For ease of presentation, we write (x, y, z) for column vector $(x^T, y^T, z^T)^T$, and use $dist(\xi, \omega^*)$ for the shortest distance from vector ξ to a closed convex set ω^* .

II. EQUIVALENT CONVERSION OF THE MANAGEMENT EQUILIBRIUM MODEL

We give in this section an equivalent conversion of (1). For convenience, let $\xi^* \square (x^*, y^*, z^*) \in R^{n+n+l}$. Then, (1) can be converted equivalently to the following problem : Find ξ^* such that

$$\begin{cases} A\xi^* \geq 0, B\xi^* \geq 0, \\ (A\xi^*)^T (B\xi^*) = 0, \\ (M, -N, -Q)\xi^* - q = 0, \end{cases} \quad (2)$$

Where $A = (I, 0, 0)$, $B = (0, I, 0)$. Let ω^* be the set of solutions of (2) and assume that it is nonempty.

We have the following conclusion.

Theorem 1.1 Vector $(x^*, y^*) \in R^{2n}$ is a solution to (1) if and only if there exists $z^* \in R^l$ such that $\xi^* = (x^*, y^*, z^*)$ is a solution to (2).

III. ESTIMATION OF THE ERROR BOUND OF THE MANAGEMENT EQUILIBRIUM MODEL

This section mainly establishes the error bound for the management equilibrium model. First, we give some related results, the definition of projection operator and its related properties.

Theorem 2.1 For a given positive constant ρ , there exists a constant $\eta_1 > 0$ such that

$$\text{dist}(\xi, \omega^*) \leq \eta_1 r(\xi), \forall \xi \in \omega, \|\xi\| \leq \rho,$$

where $r(\xi) = \|\min\{A\xi, B\xi\}\|$,

$$\omega = \{\xi \in R^{2n+l} \mid (M, -N, -Q)\xi = q\}.$$

Proof. Assume that the theorem does not hold. Then there exists a sequence $\{\xi^k\}$, such that for any positive integer k , we have

$$\text{dist}(\xi^k, \omega^*) > kr(\xi^k) \geq 0.$$

That is,

$$\frac{r(\xi^k)}{\text{dist}(\xi^k, \omega^*)} \rightarrow 0, k \rightarrow \infty, \quad (3)$$

where $\xi^k \in \omega$, and $\|\xi^k\| \leq \rho$. Since sequence $\{\xi^k\}$ is bounded, and $r(\xi)$ is continuous, together with (3), we have $r(\xi^k) \rightarrow 0, k \rightarrow \infty$. In addition, sequence $\{\xi^k\}$ has a convergent subsequence $\{\xi^{k_i}\}$. Let $\xi^{k_i} \rightarrow \bar{\xi} (k_i \rightarrow \infty)$, where $\bar{\xi} \in \omega^*$. We have the following conclusion.

$$\frac{r(\xi^{k_i})}{\|\xi^{k_i} - \bar{\xi}\|} \rightarrow \beta (k_i \rightarrow \infty). \quad (4)$$

Where, β is a positive constant.

On the other hand, from (3) we have

$$\frac{r(\xi^{k_i})}{\|\xi^{k_i} - \bar{\xi}\|} \leq \frac{r(\xi^{k_i})}{\text{dist}(\xi^{k_i}, \omega^*)} \rightarrow 0 (k_i \rightarrow \infty).$$

This contradicts with (4), hence the theorem is proved. \square

We give in the following the error bound established by Hoffman^[10].

Lemma 2.1 For a polyhedral cone

$P = \{x \in R^n \mid D_1x = d_1, D_2x \leq d_2\}$, where $D_1 \in R^{l \times n}, D_2 \in R^{m \times n}, d_1 \in R^l, d_2 \in R^m$, there exists a constant $c > 0$, such that

$$\text{dist}(x, P) \leq c[\|D_1x - d_1\| + \|(D_2x - d_2)_+\|], \forall x \in R^n.$$

Now, we also give the definition of projection operator and its related properties^[11]. For a nonempty closed convex set $S \subseteq R^n$, the orthogonal projection from vector $x \in R^n$ onto S is

$$P_S(x) := \arg \min \{\|y - x\| \mid y \in S\},$$

and it has the following property.

Lemma 2.2 For any vectors $u, v \in R^n$, we have

$$\|P_S(u) - P_S(v)\| \leq \|u - v\|.$$

Using Theorem 2.1, Lemma 2.1 and Lemma 2.2, we have the main result.

Theorem 2.2 For any positive constant ρ , there exists a constant $\eta_2 > 0$ such that

$$\text{dist}(\xi, \omega^*) \leq \eta_2 (\|(M, -N, -Q)\xi - q\| + r(\xi)), \quad \|\xi\| \leq \rho,$$

Where $r(\xi) = \|\min\{A\xi, B\xi\}\|$.

Proof. For any vector $\xi \in R^{2n+l}$, there exists $\bar{\xi} \in \omega$, such that $\|\xi - \bar{\xi}\| = \text{dist}(\xi, \omega)$. From Lemma 2.1, there exists a constant $c_1 > 0$, such that

$$\text{dist}(\xi, \omega) \leq c_1 \|(M, -N, -Q)\xi - q\|.$$

Furthermore,

$$\begin{aligned} & \|r(\xi) - r(\bar{\xi})\| \\ &= \|\min\{A\xi, B\xi\} - \min\{A\bar{\xi}, B\bar{\xi}\}\| \\ &= \|(A\xi - P_{R_+}(A\xi - B\xi)) - (A\bar{\xi} - P_{R_+}(A\bar{\xi} - B\bar{\xi}))\| \\ &\leq \|A(\xi - \bar{\xi})\| + \|(P_{R_+}(A\xi - B\xi) - P_{R_+}(A\bar{\xi} - B\bar{\xi}))\| \\ &\leq \|A(\xi - \bar{\xi})\| + \|(A\xi - B\xi) - (A\bar{\xi} - B\bar{\xi})\| \\ &\leq 2\|A(\xi - \bar{\xi})\| + \|B\xi - B\bar{\xi}\| \\ &\leq (2\|A\| + \|B\|)\text{dist}(\xi, \omega), \end{aligned}$$

Where the second inequality is based on Lemma 2.2. Combined with the above formula, we have

$$\|r(\bar{\xi})\| \leq \|r(\xi)\| + (2\|A\| + \|B\|)\text{dist}(\xi, \omega) \quad (5)$$

From (5) and Theorem 2.1, we have

$$\begin{aligned} & \text{dist}(\xi, \omega) \\ & \leq \text{dist}(\xi, \omega) + \text{dist}(\bar{\xi}, \omega^*) \leq \text{dist}(\xi, \omega) + \eta_1 r(\bar{\xi}) \\ & \leq \text{dist}(\xi, \omega) + \eta_1 [r(\xi) + (2\|A\| + \|B\|)\text{dist}(\xi, \omega)] \\ & \leq [\eta_1(2\|A\| + \|B\|) + 1]\text{dist}(\xi, \omega) + \eta_1 r(\xi) \\ & \leq [\eta_1(2\|A\| + \|B\|) + 1]c_1 \|(M, -N, -Q)\xi - q\| + \eta_1 r(\xi) \\ & \leq \eta_2 [\|(M, -N, -Q)\xi - q\| + r(\xi)], \end{aligned}$$

Where $\eta_2 = \max\{[\eta_1(2\|A\| + \|B\|) + 1]c_1, \eta_1\}$. \square

In the following we using Fischer function ([12]) to establish another error bound. Define $\phi: R^2 \rightarrow R^1$ and

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b, \forall a, b \in R.$$

It has the following property:

$$\phi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0,$$

In addition, Tseng^[13] gives the following conclusion.

Lemma 2.3

$$\begin{aligned} & [2 - \sqrt{2}] |\min(a, b)| \leq \phi(a, b) \\ & \leq (\sqrt{2} + 2) |\min(a, b)|. \end{aligned}$$

For any vectors $a, b \in R^n$, define a vector-valued function $\Psi(a, b) = (\phi(a_1, b_1), \phi(a_2, b_2), \dots, \phi(a_n, b_n))$. Based on this mapping, (2) can be converted into the following equation

$$\Phi(\xi) := \begin{pmatrix} \Psi(A\xi, B\xi) \\ (M, -N, -Q)\xi - q \end{pmatrix} = 0,$$

Clearly, using Lemma 2.3 and Theorem 2.2, it is easy to have the following result.

Theorem 2.3 For any given positive constant ρ , there exists a constant $\eta_3 > 0$ such that

$$\text{dist}(\xi, \omega^*) \leq \eta_3 \|\Phi(\xi)\|, |\xi| \leq \rho.$$

As function $\Phi(x)$ is not smooth, let $\Psi_t: R^2 \rightarrow R$ denote smooth Fisher-Burmeister function

$$\Psi_t(a, b) = \sqrt{a^2 + b^2 + 2t^2} - a - b,$$

Where $t > 0$ is a smooth parameter. For ease of presentation, let

$$\Gamma(x, y, t) = (\Psi_t(x_1, y_1), \dots, \Psi_t(x_n, y_n))^T,$$

Where $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T$,

And $p(a, b, t) = \Psi_t(a, b)$. We define mapping

$$F: R^{2n+1} \times (0, +\infty) \rightarrow R^{n+m} \times (0, +\infty),$$

That is,

$$F(\xi, t) = \begin{pmatrix} \Gamma(A\xi, B\xi, t) \\ (M, -N, -Q)\xi - q \\ t \end{pmatrix}.$$

Let $f(\xi, t) = F(\xi, t)^T F(\xi, t) = \|F(\xi, t)\|^2$.

Obviously, $\xi^* \in \omega^* \Leftrightarrow (\xi^*, 0)$ is a solution to $F(\xi, t) = 0$.

Therefore we construct a smooth method to solve $F(\xi, t) = 0$, and assume that the set of solutions to $F(\xi, t) = 0$ is ω_t^* .

First we give the following properties of $p(a, b, t)$ [14,15].

Lemma 2.4 Function $p(a, b, t)$ has the following properties:

- a) On $R^2 \times (0, +\infty)$, function $p(a, b, t)$ is continuously differentiable, and strongly semi-smooth, that is, $p(a + \Delta a, b + \Delta b, t + \Delta t) - p(a, b, t) - V^T(\Delta a, \Delta b, \Delta t) = O\|(\Delta a, \Delta b, \Delta t)\|^2$, $\forall (a, b, t) \in R^2 \times [0, +\infty)$,

Where $V \in \partial p(a + \Delta a, b + \Delta b, t + \Delta t)$, and ∂p is the Clarke generalized gradient of p .

- b) $\forall (a, b, t) \in R^2 \times (0, +\infty)$, we have $|\Psi(a, b) - p(a, b, t)| = |\Psi(a, b) - \Psi_t(a, b)| \leq \sqrt{2} t$.

Based on Lemma 2.4, we have the following result.

Theorem 2.4 Function $F(\xi, t)$ has the following properties:

- a) On $R^{2n+1} \times (0, +\infty)$, function $F(\xi, t)$ is continuously differentiable, locally Lipschitz continuous, and strongly semi-smooth, that is, there exist constants $L_1 > 0, L_2 > 0, b_1 > 0$ such that

$$\|F(\xi + \Delta\xi, t + \Delta t) - F(\xi, t)\| \leq L_1 \|(\Delta\xi, \Delta t)\|, \quad (6)$$

$$\begin{aligned} & \|F(\xi + \Delta\xi, t + \Delta t) - F(\xi, t) - H^T(\Delta\xi, \Delta t)\| \\ & \leq L_2 \|(\Delta\xi, \Delta t)\|^2, \forall (\xi, t) \in R^{2n+1} \times (0, +\infty), \quad (7) \\ & H \in \partial F(\xi + \Delta\xi, t + \Delta t), \end{aligned}$$

$$\begin{aligned} & \forall (\Delta\xi, \Delta t) \in N(0, b_1) \\ & = \{(\Delta\xi, \Delta t) \mid \|(\Delta\xi, \Delta t)\| \leq b_1, t + \Delta t \geq 0\}, \end{aligned}$$

Where $\partial F(\xi, t)$ is the Clarke generalized gradient of $F(\xi, t)$.

b) For $(\xi^*, 0) \in \omega_i^*$, there exists a neighbourhood

$$N((\xi^*, 0), b_2) = \{(\xi, t) \mid \|(\xi, t) - (\xi^*, 0)\| \leq b_2, t \geq 0\},$$

And a constant $c_1 > 0$, for any

$$(\xi, t) \in N((\xi^*, 0), b_2),$$

We have

$$\text{dist}((\xi, t), \omega_i^*) \leq c_1 \|F(\xi, t)\|. \quad (8)$$

Proof. The result of (i) follows from Lemma 2.1 directly.

(ii) For any $\|\xi\| \leq \rho$, there exists a constant $b_3 > 0$, such that

$$\begin{aligned} & \text{dist}(\xi, \omega^*) \leq \eta_2 \|\Phi(\xi)\|, \\ & \forall \xi \in N(\xi, b_2) = \{\xi \mid \|\xi - \xi^*\| \leq b_2\}, \end{aligned}$$

Let $\text{dist}(\xi, \omega^*) = \|\xi - \bar{\xi}\|$, where $\bar{\xi} \in \omega^*$.

From Lemma 2.4(ii), we have

$$\begin{aligned} & \left| \|\Phi(\xi)\| - \|\Phi_t(\xi)\| \right| \leq \|\Phi(\xi) - \Phi_t(\xi)\| \\ & \leq \sqrt{2n} t, \end{aligned}$$

Where $\Phi_t(\xi) := \begin{pmatrix} \Psi_t(A\xi, B\xi) \\ (M, -N, -Q)\xi - q \end{pmatrix}$, for any

$$(\xi, t) \in N((\xi^*, 0), b_2) = \{(\xi, t) \mid \|(\xi, t) - (\xi^*, 0)\| \leq b_2\}$$

We have

$$\begin{aligned} & \text{dist}((\xi, t), \omega^*) \leq \|(\xi, t) - (\bar{\xi}, 0)\| \leq \|\xi - \bar{\xi}\| + t \\ & \leq \eta_2 \|\Phi(\xi)\| + t \leq \eta_2 \|\Phi_t(\xi)\| + (\sqrt{2n}\eta_2 + 1)t \\ & \leq (\sqrt{2n}\eta_2 + 1)(\|\Phi_t(\xi)\| + t) \\ & \leq \sqrt{2}(\sqrt{2n}\eta_2 + 1)(\sqrt{\|\Phi_t(\xi)\|^2 + t^2}) \\ & \leq \sqrt{2}(\sqrt{2n}\eta_2 + 1)\|F(\xi, t)\|, \end{aligned}$$

where $c_1 = \sqrt{2}(\sqrt{2n}\eta_2 + 1)$. \square

IV. ALGORITHM AND CONVERGENCE

In this section, we give a smooth and convergent algorithm for solving (1), and using the error bound established in section 2, prove the quadratic convergence of the given smooth algorithm without the condition of existence of a non-degenerate solution.

Algorithm 3.1

Step 1: Choose parameters $\sigma \in (0, 1)$, $\rho > 0$ and $\varepsilon \geq 0$, initial value $(\xi, t)^0 \in R^{2n+l+1}$, and $\|(\xi, t)^0\| \leq \rho$. Let $k = 0$.

Step 2: Stop if $\|\nabla f(\xi^k, t^k)\| \leq \varepsilon$; otherwise, turn to Step 3.

Step 3: Choose the Jacobian matrix H^k of $F(\xi^k, t^k)$, and let $d^k = (\Delta\xi^k, \Delta t^k)$ be the solution to the following strict quadratic programming

$$\begin{aligned} & \min \quad \theta^k(d) \\ & \text{s.t.} \quad \|(\xi, t)^k + d\| \leq \rho, |\Delta t| \leq \frac{1}{1+\mu^k} t^k \end{aligned} \quad (9)$$

Where

$$\begin{aligned} \theta^k(d) & = \|F(\xi^k, t^k) + H^k d\|^2 + \mu^k \|d\|^2, \\ \mu^k & = \sigma \|F(\xi^k, t^k)\|^2. \end{aligned}$$

Step 4: Let $\xi^{k+1} := \xi^k + \Delta\xi^k$,

$$t^{k+1} := t^k + \Delta t^k, \quad k := k + 1, \quad \text{turn to Step 2.}$$

In the following convergence analysis, assume that Algorithm 3.1 generates an infinite sequence. We have the following result.

Theorem 3.1 Assume that Algorithm 3.1 generates a sequence $\{(\xi^k, t^k)\}$. If the initial value is close sufficiently to $\{(\xi^*, 0)\}$, which is a solution to $F(\xi, t) = 0$, then $\{\text{dist}((\xi^k, t^k), \omega^*)\}$ converges quadratically to 0, i.e., sequence $\{\xi^k\}$ converges quadratically to $\bar{\xi} \in \omega^*$.

Proof. Let $\tau := (\xi, t)$, $\tau^* := (\xi^*, 0)$. For any tiny $\delta > 0$, define

$$B_\delta(\tau^*) := \left\{ (\xi, t) \in R^{2n+l} \times (0, +\infty), \left\| (\xi, t) - (\xi^*, 0) \right\| \leq \delta \right\},$$

In the following we prove the theorem in three steps.

First we prove the following result.

If $\tau^k \in B_{\delta/2}(\tau^*)$, then

$$\|d^k\| \leq c_2 \text{dist}(\tau^k, \omega_i^*), \quad (10)$$

$$\|F(\tau^k) + H^k d^k\| \leq c_3 \text{dist}(\tau^k, \omega_i^*)^2, \quad (11)$$

Where $c_2 > 0, c_3 > 0$ are constants.

Let the closest point in ω_i^* to τ^k be $\bar{\tau}^k$, that is,

$$\|\tau^k - \bar{\tau}^k\| = \text{dist}(\tau^k, \omega_i^*) \quad (12)$$

Let $\bar{d}^k = \tau^k - \bar{\tau}^k$. As d^k is the globally optimal solution to (9), we have

$$\theta^k(d^k) \leq \theta^k(\bar{d}^k) = \theta^k(\tau^k - \bar{\tau}^k). \quad (13)$$

Since $\tau^k \in B_{\delta/2}(\tau^*)$, we have

$$\begin{aligned} \|\bar{\tau}^k - \tau^*\| &\leq \|\bar{\tau}^k - \tau^k\| + \|\tau^k - \tau^*\| \\ &\leq \|\tau^k - \tau^*\| + \|\tau^k - \tau^*\| \leq \delta. \end{aligned}$$

Hence, $\bar{\tau}^k \in B_\delta(\tau^*)$. From the definition of μ^k , (8) and (12), we have

$$\begin{aligned} \mu^k &= \sigma \|F(\tau^k)\|^2 \geq \sigma c_1^{-2} \text{dist}(\tau^k, \omega_i^*)^2 \\ &= \sigma c_1^{-2} \|\tau^k - \bar{\tau}^k\|^2 \end{aligned} \quad (14)$$

Using (12) - (14) and (7), together with the definition of $\theta^k(d)$, we know that

$$\begin{aligned} \|d^k\|^2 &\leq [1/u^k] \theta(d^k) \leq [1/u^k] \theta(\tau^k - \bar{\tau}^k) \\ &= [1/u^k] [\|F(\tau^k) + H^k(\tau^k - \bar{\tau}^k)\|^2 + \mu^k \|\tau^k - \bar{\tau}^k\|^2] \\ &= [1/u^k] \|F(\tau^k) + F(\bar{\tau}^k) + H^k(\tau^k - \bar{\tau}^k)\|^2 \\ &\quad + \|\tau^k - \bar{\tau}^k\|^2 \leq [1/u^k] L_2^2 \|\tau^k - \bar{\tau}^k\|^4 + \|\tau^k - \bar{\tau}^k\|^2 \\ &\leq \left\{ 1 / \left[\sigma c_1^{-2} \|\tau^k - \bar{\tau}^k\|^2 \right] \right\} L_2^2 \|\tau^k - \bar{\tau}^k\|^4 + \|\tau^k - \bar{\tau}^k\|^2 \\ &\leq \left\{ L_2^2 / \left[\sigma c_1^{-2} \right] \right\} \|\tau^k - \bar{\tau}^k\|^2 + \|\tau^k - \bar{\tau}^k\|^2 \\ &\leq \left\{ L_2^2 / \left[\sigma c_1^{-2} \right] + 1 \right\} \|\tau^k - \bar{\tau}^k\|^2 \leq c_2 \text{dist}(\tau^k, \omega_i^*)^2 \end{aligned}$$

Where $c_2 = \left\{ L_2^2 / \left[\sigma c_1^{-2} \right] + 1 \right\}$. Then (10) holds.

From the definition of $\theta^k(d)$, we know

$$\|F(\tau^k) + H^k d^k\|^2 \leq \theta^k(d^k). \quad (15)$$

In addition, from (13), (7) and the definition of $\theta^k(d)$, we have

$$\begin{aligned} \theta^k(d^k) &\leq \theta^k(\tau^k - \bar{\tau}^k) \leq \|F(\tau^k) + F(\bar{\tau}^k) + H^k(\tau^k - \bar{\tau}^k)\|^2 \\ &\quad + \mu^k \|\tau^k - \bar{\tau}^k\|^2 \leq L_2^2 \|\tau^k - \bar{\tau}^k\|^4 + \mu^k \|\tau^k - \bar{\tau}^k\|^2 \end{aligned} \quad (16)$$

From (6), we also have

$$\begin{aligned} \mu^k &= \sigma \|F(\tau^k)\|^2 = \sigma \|F(\tau^k) - F(\bar{\tau}^k)\|^2 \\ &\leq \sigma L_1^2 \|\tau^k - \bar{\tau}^k\|^2 \end{aligned}$$

Together with (15) and (16), we have

$$\begin{aligned} \|F(\tau^k) + H^k d^k\|^2 &\leq \theta^k(d^k) \\ &\leq L_2^2 \|\tau^k - \bar{\tau}^k\|^4 + \mu^k \|\tau^k - \bar{\tau}^k\|^2 \\ &\leq L_2^2 \|\tau^k - \bar{\tau}^k\|^4 + \sigma L_1^2 \|\tau^k - \bar{\tau}^k\|^4 \\ &\leq (L_2^2 + \sigma L_1^2) \|\tau^k - \bar{\tau}^k\|^4 \\ &= c_3^2 \|\tau^k - \bar{\tau}^k\|^4 = c_3^2 \text{dist}(\tau^k, \omega_i^*)^4 \end{aligned}$$

Where $c_3 = \sqrt{L_2^2 + \sigma L_1^2}$.

Next, for any natural number k , if $\tau^k, \tau^{k-1} \in B_{\delta/2}(\tau^*)$, there exists $c_4 > 0$, such that

$$\text{dist}(\tau^k, \omega_i^*) \leq c_4 \text{dist}(\tau^{k-1}, \omega_i^*)^2 \quad (17)$$

In fact, since $\tau^k, \tau^{k-1} \in B_{\delta/2}(\tau^*)$, and $\tau^k = \tau^{k-1} + d^{k-1}$, together with (8), we have

$$\begin{aligned} \|F(\tau^{k-1} + d^{k-1})\| &= \|F(\tau^{k-1}) + H^{k-1} d^{k-1}\| \\ &\leq \|-F(\tau^{k-1} + d^{k-1}) + F(\tau^{k-1}) + H^{k-1} d^{k-1}\| \\ &\leq L_2 \|d^{k-1}\|^2 \end{aligned}$$

That is,

$$\begin{aligned} \|F(\tau^{k-1} + d^{k-1})\| &\leq L_2 \|d^{k-1}\|^2 \\ &\quad + \|F(\tau^{k-1}) + H^{k-1} d^{k-1}\| \end{aligned} \quad (18)$$

Using (18), (8), (10) and (11), we know

$$\begin{aligned} \text{dist}(\tau^k, X^*) &\leq c_1 \|F(\tau^k)\| \\ &= c_1 \|F(\tau^{k-1} + d^{k-1})\| \\ &\leq c_1 L_2 \|d^{k-1}\|^2 + c_1 \|F(\tau^{k-1}) + H^{k-1} d^{k-1}\| \\ &\leq c_1 L_2 c_2^2 \text{dist}(\tau^k, \omega_i^*)^2 + c_1 c_3 \text{dist}(\tau^k, \omega_i^*)^2 \\ &\leq (c_1 L_2 c_2^2 + c_1 c_3) \text{dist}(\tau^k, \omega_i^*)^2 \\ &= c_4 \text{dist}(\tau^k, \omega_i^*)^2 \end{aligned}$$

Where $c_4 = c_1 L_2 c_2^2 + c_1 c_3$.

Last, we prove that the condition of (17) holds. That is, for a positive constant

$\lambda = \min \{ \delta / [2(1 + 2c_2)], 1 / (2c_4) \}$, when initial value $\tau^0 \in B_\lambda(\tau^*)$, for any natural number k , we have $\tau^k \in B_{\delta/2}(\tau^*)$.

We prove the above result by mathematical induction.

When $k = 0$, from the way λ is chosen, we know $\lambda \leq \delta / 2$, and then $\tau^0 \in B_{\delta/2}(\tau^*)$. Now assume that $k \geq 0$, $\tau^m \in B_{\delta/2}(\tau^*)$, for $m = 0, 1, 2, \dots, k$, We prove in the following $\tau^{k+1} \in B_{\delta/2}(\tau^*)$.

$$\begin{aligned} & \| \tau^{k+1} - \tau^* \| = \| \tau^k + d^k - \tau^* \| \\ & \leq \| \tau^k - \tau^* \| + \| d^k \| \leq \| \tau^{k-1} + d^{k-1} - \tau^* \| + \| d^k \| \\ & \leq \| \tau^{k-1} - \tau^* \| + \| d^k \| + \| d^{k-1} \| \dots \\ & \leq \| \tau^0 - \tau^* \| + \sum_{m=0}^k \| d^m \| \leq \lambda + c_2 \sum_{m=0}^k \text{dist}(\tau^m, \omega_i^*) \end{aligned}$$

where, from (10), the last inequality holds. In addition, since

$\tau^m \in B_{\delta/2}(\tau^*)$, $m = 0, 1, 2, \dots, k$, together with (17),

we have

$$\text{dist}(\tau^m, \omega_i^*) \leq c_4 \text{dist}(\tau^{m-1}, \omega_i^*)^2, m = 0, 1, 2, \dots, k$$

Hence

$$\begin{aligned} & \text{dist}(\tau^m, \omega_i^*) \leq c_4 \text{dist}(\tau^{m-1}, \omega_i^*)^2 \\ & \leq c_4 c_4^2 \text{dist}(\tau^{m-2}, \omega_i^*)^2 \dots \\ & \leq c_4 c_4^2 \dots c_4^{2^{m-1}} \text{dist}(\tau^0, \omega_i^*)^{2^m} \\ & \leq c_4^{2^m - 1} \| \tau^0 - \tau^* \|^{2^m} \leq c_4^{2^m - 1} \lambda^{2^m} \end{aligned}$$

From the above formula, and the way λ is chosen, we know that $\lambda \leq 1 / (2c_4)$, $\lambda \leq \delta / [2(1 + 2c_2)]$,

And

$$\begin{aligned} & \| \tau^{k+1} - \tau^* \| \leq \lambda + c_2 \sum_{m=0}^k \text{dist}(\tau^m, \omega_i^*) \\ & \leq \lambda + c_2 \sum_{m=0}^k c_4^{2^m - 1} \lambda^{2^m} \leq \lambda + c_2 \lambda \sum_{m=0}^k c_4^{2^m - 1} \lambda^{2^m - 1} \\ & \leq \lambda + c_2 \lambda \sum_{m=0}^k \left(\frac{1}{2} \right)^{2^m - 1} \leq \lambda + c_2 \lambda \sum_{m=0}^{\infty} \left(\frac{1}{2} \right)^m \\ & \leq (1 + 2c_2) \lambda \leq \delta / 2 \end{aligned}$$

Now Theorem 3.1 is proved. \square

NOTE: Theorem 3.1 shows that the given smooth algorithm has the property of quadratic convergence without the condition of existence of a non-degenerate solution. This is a new result.

V. CONCLUSIONS

In this paper, we propose an algorithm for solving the management equilibrium model. Under without the requirement of nondegenerate solution, we also show that the algorithm is quadratic convergence based on error bound estimation instead of the nonsingular assumption just as was done in [5,6]. This conclusion can be viewed as extension of previously known result in [5, 6]. How to use the algorithm to solve the practical management based on the computer, this is a topic for future research.

ACKNOWLEDGMENT

The authors wish to give their sincere thanks to the editor and the anonymous referees for their valuable suggestions and helpful comments which improved the presentation of the paper.

REFERENCES

- [1] FERRIS M C, PANG J S. Engineering and economic applications of complementarity problems[J]. Society for industrial and applied mathematics, 1997, 39(4): 669-713.
- [2] COTTLE R W, PANG J S, STONE R E. The linear complementarity problem[M]. New York: Academic press, 1992.
- [3] WALRAS L. Elements of Pure Economics [M]. London : Allen and Unwin, 1954.
- [4] FACCHINEI F, PANG J S. Finite-Dimensional Variational Inequalities and Complementarity Problems[M]. New York: Springer, 2003.
- [5] REN Q J, SUN H C. A Quadratically Convergent Algorithm for the Generalized Linear Complementarity Problem[J]. International Mathematical Forum, 2007, 60: 2971-2981.
- [6] SUN H. A non-smooth quadratically convergent algorithm for solving horizontal linear complementarity problems [J]. Journal of Sichuan Normal University, 2007, 5: 560-564.
- [7] YAMASHITA N , FUKUSHIMA M. On the rate of convergence of the Levenberg-Marquardt method[J]. Computing [Suppl], 2001, 15: 239-249.
- [8] PANG J S. Error bounds in mathematical programming [J]. Math . Programming, 1997, 79: 299-332.
- [9] MANGASARIAN O L , REN J. New improved error bound for the linear complementarity problem[J]. Math. Programming, 1994, 66: 241-255.
- [10] HOFFMAN A J. on the approximate solutions of linear inequalities[J]. J. Res. National Bureau of Standards, 1952, 49: 263-265.
- [11] ZARANTONELLO E. H., Projections on Convex Sets in Hilbert Space and Spectral Theory, Contributions to Nonlinear Functional Analysis, New York: Academic Press, 1971.
- [12] FISCHER A. A special Newton-type optimization method[J]. Optim., 1992, 24: 269-284.
- [13] TSENG P. Growth behavior of a class of merit function for the nonlinear complementarity problem[J]. J. Optim. Theory Appl., 1996, 89: 17-37.
- [14] ENGELKE S, KONZOW C. Improved smoothing type method for the solution of linear programs[J]. Preprint, Institute of Applied Mathematics, University of Hamburg, Hamburg, March, 2000.
- [15] QI L, SUN D. Smoothing function and smoothing Newton method for complementarity problems and variational inequality problems[J]. Reprint, University of New South Wales, Sydney, Australia, 1998.

The Bitwise Operations Related to a Fast Sorting Algorithm

Krasimir Yordzhev

Faculty of Mathematics and Natural Sciences
South-West University
Blagoevgrad, Bulgaria

Abstract—in the work we discuss the benefit of using bitwise operations in programming. Some interesting examples in this respect have been shown. What is described in detail is an algorithm for sorting an integer array with the substantial use of the bitwise operations. Besides its correctness we strictly prove that the described algorithm works in time $O(n)$. In the work during the realization of each of the examined algorithms we use the apparatus of the object-oriented programming with the syntax and the semantics of the programming language C++

Keywords—bitwise operations; programming languages C/C++ and Java; sorting algorithm

I. INTRODUCTION

The use of bitwise operations is a powerful means during programming with the languages C/C++ and Java. Some of the strong sides of these programming languages are the possibilities of low level programming. Some of the means for this possibility are the introduced standard bitwise operations, with the help of which it is possible to directly operate with every bit of an arbitrary variable situated in the random access memory of the computer. In the current article we are going to describe some methodical aspects for work with the bitwise operations.

As an interesting example of application of the bitwise operations comes the realised by us algorithm for sorting an integer array, for which we strictly prove its correctness and the fact that this algorithm will use operations included in the standard of the programming language C++. A main role in the realisation of the algorithm plays the bitwise operations.

II. BITWISE OPERATIONS

The bitwise operations can be applied for integer data type only. For the definition of the bitwise operations and some of their elementary applications could be seen, for example, in [2], [5] for C/C++ programming languages and in [4], [7] for Java programming language.

We assume, as usual that bits numbering in variables starts from right to left, and that the number of the very right one is 0.

Let x and y be integer variables or constants and let z be integer variables of one type, for which w bits are needed. Let x and y be initialized (if they are variables) and let the assignment $z = x \& y$; (bitwise AND), or $z = x | y$; (bitwise inclusive OR), or $z = x \wedge y$; (bitwise exclusive OR), or $z = \sim x$; (bitwise NOT) be made. For each $i = 0, 1, 2, \dots, w-1$, the

new contents of the i -th bit in z will be as it is presented in the Table I.

TABLE I. BITWISE OPERATIONS IN PROGRAMMING LANGUAGES C/C++ AND JAVA

i -th bit of					
x	y	$z = x \& y$	$z = x y$	$z = x \wedge y$	$z = \sim x$
0	0	0	0	0	1
0	1	0	1	1	1
1	0	0	1	1	0
1	1	1	1	0	0

In case that k is a nonnegative integer, then the statement $z = x \ll k$; (bitwise shift left) will write in the $(i+k)$ bit of z the value of the k bit of x , where $i = 0, 1, \dots, w-k-1$, and the very right k bits of x will be filled by zeroes. This operation is equivalent to a multiplication of x by 2^k .

The statement $z = x \gg k$ (bitwise shift right) works the similar way. But we must be careful if we use the programming language C or C++, as in various programming environments this operation has different interpretations: somewhere k bits of z from the very left place are compulsory filled by 0 (logical displacement), and elsewhere the very left k bits of z are filled with the value from the very left (sign) bit; i.e. if the number is negative, then the filling will be with 1 (arithmetic displacement). Therefore it is recommended to use unsigned type of variables (if the opposite is not necessary) while working with bitwise operations (see also Example 3). In the Java programming language, this problem is solved by introducing the two different operators: $z = x \gg k$ and $z = x \ggg k$ [4], [7].

Bitwise operations are left associative.

The priority of operations in descending order is as follows: \sim (bitwise NOT); the arithmetic operations $*$ (multiply), $/$ (divide), $\%$ (remainder or modulus); the arithmetic operations $+$ (addition) - (subtraction); the bitwise operations \ll and \gg ; the relational operations $<$, $>$, $<=$, $>=$, $==$, $!=$; the bitwise operations $\&$, \wedge and $|$; the logical operations $\&\&$ and $\|\|$.

III. SOME ELEMENTARY EXAMPLES OF USING THE BITWISE OPERATIONS

Example 1: To compute the value of the i -th bit (0 or 1) of an integer variable x we can use the function:

```
int BitValue(int x, unsigned int i) {  
    int b = ((x & 1<<i) == 0) ? 0 : 1;  
    return b;  
}
```

Example 2: 2 Directly from the definition of the operation bitwise shift left (<<) follows the efficiency of the following function computing 2^n , where n is a nonnegative integer:

```
unsigned int Power2(unsigned int n) {  
    return 1<<n;  
}
```

Example 3:3 The integer function $f(x) = x \% 2^n$ implemented using operation bitwise shift right (>>).

```
int Div2(int x, unsigned int n) {  
    int s = x<0 ? -1 : 1;  
    /* s = the sign of x */  
    x = x*s;  
    /* We reset the sign bit of x */  
    return (x>>n)*s;  
}
```

When we work with negative numbers we must consider that in the computer the presentation of the negative numbers is through the so called true complement code. The following function gives us how to code the integers in the memory of the computer we work with. For simplicity we are going to work with type short, but it is not a problem for the function to be overloaded for other integer types, too.

Example 4:4 A function showing the presentation of the numbers of type short in the memory of the computer.

```
void BinRepl(short n) {  
    int b;  
    int d = sizeof(short)*8 - 1;  
    while (d>=0) {  
        b = 1<<d & n ? 1 : 0;  
        cout<<b;  
        d--;  
    }  
}
```

Some experiments with the function BinRepl are given in Table II.

TABLE II. PRESENTATION OF SOME NUMBERS OF TYPE SHORT IN THE MEMORY OF THE COMPUTER

An integer of type short	Presentation in memory
0	0000000000000000
1	0000000000000001
-1	1111111111111111
2	0000000000000010
-2	1111111111111110
$16 = 2^4$	0000000000010000
$-16 = -2^4$	1111111111110000
$26 = 2^4 + 2^3 + 2$	000000000011010
$-26 = -(2^4 + 2^3 + 2)$	111111111100110
$41 = 2^5 + 2^3 + 1$	000000000101001
$-41 = -(2^5 + 2^3 + 1)$	111111111010111
$32767 = 2^{15} - 1$	0111111111111111
$-32767 = -(2^{15} - 1)$	1000000000000001
$32768 = 2^{15}$	1000000000000000
$-32768 = -2^{15}$	1000000000000000

Compare the function presented in Example 4 to the next function presented in Example 5.

Example 5:5 A function that prints a given integer in binary notation.

```
void DecToBin(int n) {  
    if (n<0) cout<<'-' ;  
    /* Prints the sign - , if n<0: */  
    n = abs(n);  
    int b;  
    int d = sizeof(int)*8 - 1;  
    while ( d>0 && (n & 1<<d) == 0 ) d--;  
    /* Skips the insignificant zeroes at the beginning: */  
    while (d>=0) {  
        b = 1<<d & n ? 1 : 0;  
        cout<<b;  
        d--;  
    }  
}
```

Example 6:6 The following function calculates the number of 1 in a given integer n written in a binary notation. Here again we ignore the sign of the number (if it is negative) and we work with its absolute value.

```
int NumbOf_1(int n) {  
    n = abs(n);
```

```
int temp=0;
int d = sizeof(int)*8 - 1;
for (int i=0; i<d; i++)
    if (n & 1<<i) temp++;
return temp;
}
```

IV. BITWISE SORTING

In this section we are going to suggest a fast algorithm for sorting an arbitrary integer array. And since during its realisation we are substantially going to use bitwise operations, we will call it "**Bitwise sorting**". We will prove that the bitwise sorting works in time $O(n)$, where n is the size of the array. This is an excellent evaluation regarding the criterion time. For comparison below we give some of the most famous sorting algorithms and their evaluations by criterion time [1], [6], [8].

- **Selection sort** – works in time $O(n^2)$;
- **Bubble sort** – works in time $O(n^2)$;
- **Bidirectional bubble sort (Cocktail sort)** – works in time $O(n^2)$;
- **Insertion sort** – works in time $O(n^2)$;
- **Merge sort** – works in time $O(n \log n)$;
- **Tree sort** – works in time $O(n \log n)$;
- **Timsort** – works in time $O(n \log n)$;
- **Counting sort** – works in time $O(n + m)$, where m is a second parameter, giving the number of the unique keys;
- **Bucket sort** – works in time $O(n)$.

Notes:

1) Timsort has been developed for use with the programming language Python[??].

2) For Counting sort $O(n + m)$ additional memory is necessary.

3) For Bucket sort $O(m)$ additional memory is necessary, where m is another parameter, giving the number of the unique keys, and it is also necessary to have knowledge of the nature of the sorted data which goes beyond the functions "swap" and "compare".

V. PROGRAMME CODE OF THE ALGORITHM

The algorithm created by us, described with the help of programming language C++, is shown below (algorithm 1). Due to some obvious reasons, first we create a function which sorts an array whose elements are either only nonnegative, or only negative. The second function divides the given array into

two disjoint subarrays respectively only with negative and only with nonnegative elements. After sorting each one of these subarrays, we merge them so that we obtain one finally sorted array.

Algorithm 1:7 Bitwise sorting.

```
/* The first function sorts integer elements with the same signs: */
```

```
template <class T>
void BitwiseSort1(T A[],int n){
int t;
/* t - size of the type T in bits */
t = sizeof(T)*8;
T A0[n], A1[n];
/* A0 remembers the elements for which the k-th bit is 0 */
/* A1 remembers the elements for which the k-th bit is 1 */
int n0; // size of A0
int n1; // size of A1
for (int k=0; k<t-1; k++) {
/* k - number of the bit. The numeration starts from 0.
Does not check the sign bit */
n0=0;
n1=0;
for (int i=0; i<n; i++){
/* checks the k-th bit of the i-th element of the array: */
if (A[i] & 1<<k) {
A1[n1] = A[i];
n1++;
}
else {
A0[n0] = A[i];
n0++;
}
}
/* We merge the two arrays. As a result we obtain an array
whose elements if the k-th bit is equal to 0 are at the beginning,
and if it is equal to 1 at the end. */
for (int i=0; i<n0; i++)
A[i] = A0[i];
for (int i=0; i<n1; i++)
A[n0+i] = A1[i];
}
```

```

}
/* The second function sorts the whole array */
template <class T>
void BitwiseSort(T A[],int n){
    T Aminus[n], Aplus[n];
    /* Aminus[n] - An array with the negative values of A */
    /* Aplus[n] - An array with the nonnegative values of A */
    int Nm = 0, Np = 0;
    /* Nm -- number of elements written in Aminus */
    /* Np -- number of elements written in Aplus */
    for (int i=0; i<n; i++)
        if (A[i] < 0) {
            Aminus[Nm] = A[i];
            Nm++;
        }
        else {
            Aplus[Np] = A[i];
            Np++;
        }
    /* Sorts the negative elements: */
    BitwiseSort1(Aminus,Nm);
    /* Sorts the nonnegative elements */
    BitwiseSort1(Aplus,Np);
    /* We merge the two arrays: */
    for (int i=0; i<Nm; i++)
        A[i] = Aminus[i];
    for (int i=Nm; i<n; i++)
        A[i] = Aplus[i-Nm];
}

```

VI. EVALUATION OF THE ALGORITHM

As a main disadvantage of the algorithm described by us comes the fact that it is applicable only to arrays of integers or symbols (type char). This is because for it we substantially use bitwise operations, which are applicable only over integer types of data. But this disadvantage is compensated by its high speed. As we will see below, algorithm 1 works in time $O(n)$, where n is the number of the elements which are subjected to sorting.

Except through the multiple experiments which we have made, with the help of the following theorem we will prove the correctness of the algorithm created by us.

Theorem 1:8 During every execution of algorithm 1:7 with an arbitrary input array of integers, as a result a sorted array is obtained.

Proof. It is enough to prove that function BitwiseSort1 works so as to fulfill the conditions of the theorem.

Let $A = \{a_0, a_1, \dots, a_{n-1}\}$ be an arbitrary integer array with length n and let $A^{(k)} = \{a_0^{(k)}, a_1^{(k)}, a_{n-1}^{(k)}\}$ be the array which is obtained after iteration with number k , where $k = 0, 1, \dots, t-2$, $t = \text{sizeof}(T)*8$, i.e. t is equal to the number of the bits which every element of A occupies in the memory of the computer.

Let x be an integer. For every natural number $k = 0, 1, 2, \dots$ we define the functions:

$$\mu_k(x) = x \% 2^k,$$

where just like in programming languages C/C++ and Java the operator % denotes the remainder during integer division. Apparently $\mu_{s-1}(x) = x$ if the absolute value of the integer x can be written with no more than s digits 0 or 1 in a binary notation. Therefore in order to prove that as a result of the work of the algorithm the array $A^{(t-2)}$ is sorted, it is enough to prove that the array

$$\overline{A^{(t-2)}} = \{\mu_{t-2}(a_0^{(t-2)}), \mu_{t-2}(a_1^{(t-2)}), \dots, \mu_{t-2}(a_{n-1}^{(t-2)})\}$$

is sorted. Applying inductive reasoning, we will prove that for every s , such that $0 \leq s < t-1$, the array

$$\overline{A^{(s)}} = \{\mu_s(a_0^{(s)}), \mu_s(a_1^{(s)}), \dots, \mu_s(a_{n-1}^{(s)})\}$$

is sorted.

When $s = 0$ the assertion follows from the fact that during iteration with number 0 ($k = 0$), $A^{(0)}$ is ordered so that first come all elements of the array which in their binary notation end in 0, followed by all elements of the array which in their binary notation end in 1.

We assume that for a certain natural number s , $0 \leq s < t-2$ the array

$$\overline{A^{(s)}} = \{\mu_s(a_0^{(s)}), \mu_s(a_1^{(s)}), \dots, \mu_s(a_{n-1}^{(s)})\}$$

is sorted. But then analysing the work of the algorithm in $(s+1)$ -th iteration, it is easy to see that the array $A0$, which is obtained from $\overline{A^{(s)}}$ taking in the same row only these elements of $A^{(s)}$ having 0 in bit with number $s+1$, is a sorted array. Analogously we prove that the array $A1$ is sorted and in

bit with number $s + 1$ on each of its elements stands 1. Then the array

$$\overline{A^{(s+1)}} = \{\mu_{s+1}(a_0^{(s+1)}), \mu_{s+1}(a_1^{(s+1)}), \dots, \mu_{s+1}(a_{n-1}^{(s+1)})\},$$

which is obtained from the merger of the arrays A0 and A1 where the elements of A0 precede the elements of A1, is sorted. And with this we have proven the theorem.

Theorem 2:9 Algorithm 1 described with the help of programming language C++ works in time $O(n)$.

Proof. The assertion of the theorem follows from the fact that in function BitwiseSort1 we have only two nested loops. In the inner loop exactly n iterations are performed, and in every iteration once the operation $\&$ (bitwise conjunction), once the operation \ll (bitwise shift left), once the if statement, once the assignment statement and once the increment statement are performed. Each of the aforesaid operations is performed in constant time. The outer loop does $t - 2$ iterations, where t is a constant, and in every iteration besides the inner loops there are also two assignment operations.

In function BitwiseSort the division of the array into two disjuncting subarrays is performed apparently in time $O(n)$. The newly obtained two arrays are sorted in total time $O(n)$. The following merger of the two sorted arrays with total length n is apparently also performed in time $O(n)$.

REFERENCES

- [1] J. Darlington, A Synthesis Of Several Sorting Algorithms. Acta Informatica, Volume 11, Number 1, December 18, 1978, 1-30.
- [2] S. R. Davis, C++ For Dummies, IDG Books Worldwide, 2000.
- [3] M. L. Hetlan, Python Algorithms: Mastering Basic Algorithms In The Python Language, Apress, 2010.
- [4] D. Flanagan, JAVA In A Nutshell, O'Reilly, 2002.
- [5] B. W. Kernigan, D. M. Ritchie, The C Programming Language, AT T Bell Laboratories, 1998.
- [6] D. E. Knuth, The Art Of Computer Programming. Volume 3: Sorting And Searching, Second Edition, Addison-Wesley, 1998.
- [7] H. Schildt, Java 2 A Beginner's Guide, Mcgraw-Hill, 2001.
- [8] R. Sedgewick, Algorithms In C++, Parts 1-4 Fundamentals, Data Structures, Sorting, Searching, Third Edition, Addison Wesley Longman, 1999

Efficient Role Assignment Scheme for Multichannel Wireless Mesh Networks

Sunmyeng Kim

Department of Computer Software Engineering
Kumoh National Institute of Technology
Gumi, South Korea

Hyun Ah Lee

Department of Computer Software Engineering
Kumoh National Institute of Technology
Gumi, South Korea

Abstract—a wireless mesh network (WMN) is cost-effective access network architecture. The performance of multi-hop communication quickly reduces as the number of hops becomes larger. Nassiri et al. proposed a Molecular MAC protocol for autonomic assignment and use of multiple channels to improve network performance. In the Molecular MAC protocol, each node forms a shortest path-spanning tree to a gateway node linked to a wired Internet. After a tree is formed, the nodes with an even-numbered depth and an odd-numbered depth are assigned with the roles of a nucleus and an electron, respectively. After such roles are assigned, each nucleus selects an idle channel. However, this protocol has the following drawback; since the nodes with an even-numbered depth are assigned with the role of a nucleus, there are many nuclei in the topology. The number of assigned channels tends to increase, since each nucleus selects an idle channel that is not currently being occupied by its neighboring nuclei. In wireless communications networks, channels are very important resources. Thus, it is necessary to assign the minimum number of channels as little as possible. To do so, this paper proposes an efficient role assignment scheme, which can reduce the number of assigned channels by reducing the number of nodes assigned as nuclei and preventing nodes within the transmission range of each other from becoming nuclei. Based on various simulation results, the proposed scheme was verified.

Keywords—role assignment; multichannel; mesh network

I. INTRODUCTION

A wireless mesh network (WMN) is a cost-effective access network architecture. It is a promising wireless technology for numerous applications. It is gaining significant attention as a possible way for Internet service providers (ISPs) and carriers in order to provide wireless broadband service. In WMNs, nodes are consisted of mesh routers and mesh clients [1]. Mesh routers have minimal mobility and form the backbone of WMNs. When a node is within the transmission range of another node, they are considered as neighbors, and there is a wireless link between them. Some nodes, called gateways, are connected to the wired network, which connects the WMNs to rest of the Internet.

The packets sent by end users travel through wireless mesh networks over multiple hops. And gateway nodes relay traffic to and from the wired Internet. The performance of multi-hop communication quickly reduces as the number of hops becomes larger due to intra-flow and inter-flow interferences [2-4]. This is because a wireless link is shared among

neighboring nodes.

With a single channel, nodes operate on the same channel. Therefore, network performance decreases. Nodes can be equipped with multiple radios and channels. This means there is a unique frequency used for each wireless hop, and thus enables separation of wireless collision domain and reduces the interference and contention. This can significantly improve network performance without bandwidth degradation.

The design of the MAC protocol is the most likely challenge in WMNs. An interesting problem in WMNs is how to efficiently utilize multiple channels. Several MAC protocols for handling multiple channels have been proposed in the literature [5-17].

The algorithms proposed in [5] select channels for the mesh radios to minimize interference within the mesh network and between the mesh network and co-located wireless networks. A new method for the interference estimation is suggested. All available channels are periodically monitored by each mesh node and measured information of internal channel usage and external interference is shared with mesh nodes within interference range. Duarte et al. use game theory to design a systematic approach to utilize partially overlapped channels in WMNs while minimizing the adverse effect of adjacent channel interference [6]. In [7], both centralized and distributed algorithms are presented, which aim to minimize the number of pairs of links that are interfering. The Molecular MAC protocol was proposed to organize the mesh network according to the molecular analogy [8-12]. It divides the network into atoms with nucleus nodes operating on fixed channels and electrons that dynamically switch channels between neighbor nuclei. Electrons can be shared by numerous atoms.

In the Molecular MAC protocol, each node forms a shortest path-spanning tree to a gateway node linked to a wired Internet. After a tree is formed, the nodes with an even-numbered depth and an odd-numbered depth are assigned with the roles of a nucleus and an electron, respectively. After such roles are assigned, each nucleus selects an idle channel that is not currently being occupied by its neighboring atoms with the assistance from electrons in the same atom.

Since a node with an even-numbered depth is a nucleus in a spanning tree, the Molecular MAC protocol can support many nuclei, in which nodes within the transmission range of each other can be nuclei. The number of assigned channels in the Molecular MAC protocol tends to increase, since each nucleus

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0025495).

selects an idle channel that is not currently being occupied by its neighboring nuclei. In wireless communications networks, channels are very important resources. Thus, it is necessary to assign the minimum number of channels as little as possible. To do so, this paper proposes an efficient role assignment scheme, which can reduce the number of assigned channels by reducing the number of nodes assigned as nuclei and preventing nodes within the transmission range of each other from becoming nuclei. The proposed scheme adopts the basic operating principles of the Molecular MAC protocol, though the role assignment scheme is revised. For role assignment, the Molecular MAC protocol uses the shortest path-spanning tree, but the proposed scheme introduces a new metric by considering three components affecting network performance, and assigns the role of each node based on it. After the role assignment in the proposed scheme, nodes assigned as nuclei select channels with the help of electrons as in the Molecular MAC protocol.

The paper is organized as follows. We give a brief introduction of the Molecular MAC protocol and its role assignment problem in Section II. In Section III, the proposed scheme is presented in detail. In Section IV, performance studies are carried out through simulation results. Finally, we draw a conclusion in Section V.

II. RELATED WORK

This Section is designed to briefly specify the Molecular MAC protocol and then touch on certain problem relating to the role assignment.

A. Molecular MAC Protocol

The IEEE 802.11 wireless network functions well in the infrastructure mode. In addition, it can provide a fair bandwidth to all users by slightly modifying a channel access method. Nevertheless, the IEEE 802.11 network may incur numerous problems on a multi-hop network. The Molecular MAC protocol expands the IEEE 802.11 network in order to transmit data packets on the multi-hop network effectively.

Since the IEEE 802.11 access method works well on a single-hop network, the Molecular MAC protocol divides a wireless mesh network into different spatially distributed atoms. Each atom uses a channel not used by other neighboring atoms. An atom is composed of one nucleus and several electrons, and a nucleus selects a channel to be used by its own atom. Any node within an atom's boundary plays a role as an electron and belongs to a neighboring atom. An electron directly communicates with its nuclei; however, since there is no direct link between electrons, a direct communication among them cannot be implemented. In addition, due to no direct link among nuclei, direct communications among them are also not possible. Therefore, the communication among neighboring electrons can be handled by nuclei while the communication among nuclei can be executed by neighboring electrons.

In the Molecular MAC protocol, each node is assigned with the role as a nucleus or an electron while each nucleus selects a channel to be used by its own atom. Accordingly, each node forms a shortest path-spanning tree to a gateway node linked to a wired Internet. After a tree is formed, the nodes with an even-

numbered depth and an odd-numbered depth are assigned with the roles of a nucleus and an electron, respectively. After such roles are assigned, each nucleus requests its own electrons for channel information. Channel information includes a list of active channels and their activities. A list of active channels includes the numbers of active channels in the corresponding electrons' parent atoms. A channel activity is a parameter that is designed to indicate how many data packets are transmitted on each channel in the active channel list, which is expressed in the number of packets transmitted. Requested by a nucleus, each electron makes up a list of active channels, measures each channel's activities, and accordingly responds to the nuclei. Then, a nucleus receives certain responses from all neighboring electrons and accordingly selects a channel according to a subsequent rule. 1) After a list of active channels is received from all neighboring electrons, a channel is randomly chosen out of those non-active channels. 2) If every channel is currently occupied, a channel with the least activity is selected. Once a nucleus allocates a channel, its neighboring electrons use the channel. The corresponding electrons use all channels allocated by their atoms' nuclei.

Fig. 1 illustrates a fundamental structure of the Molecular MAC protocol. As shown in the figure, there are 2 atoms, 2 nuclei and 6 electrons. The atom 1 includes the nucleus N1 and the electrons (E1, E2, E3 and E4), and uses the channel 1. The atom 2 includes the nucleus N2 and the electrons (E3, E4, E5 and E6), and uses the channel 2. The electrons E3 and E4, shared by the two atoms, use both channel 1 and channel 2. If a neighboring atom's nucleus requests information for the allocation of a channel, the corresponding electrons transmit their channel activity information. Furthermore, the electrons E1/E2, E3/E4 and E5/E6 also transmit the active channel lists including the channel 1, the channels 1 & 2 and the channel 2, respectively, to the nucleus. Once all channels are allocated, communications are processed as follows: As shown in the figure, when certain data are to be transmitted from E1 to E6, E1 transmits the data to N1 using the channel 1 while N1 transmits the data back to E3 or E4. Then, E3 or E4 transmits the data to N2 using the channel 2 while N2 transmits the data back to E6. The neighboring electrons, E1 and E2, do not directly communicate with each other but communicate via the nucleus N1.

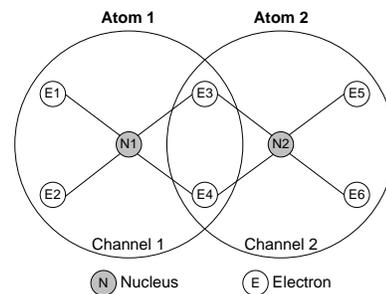


Fig. 1. Basic architecture of the Molecular MAC protocol

B. Role Assignment Problem

There is a problem in role assignment scheme of the Molecular MAC protocol. First of all, each node forms a shortest path-spanning tree to a gateway node connected to the

wired Internet. After forming the tree, a node of which depth is an even number is assigned as a nucleus, and a node of which depth is an odd number as an electron. Since this scheme prefers the shortest path to the gateway node, the number of nuclei tends to increase, and that of assigned channels also increases. In addition, it only takes the even-numbered depth into account in role assignment. Nodes within the transmission range of each other can become nuclei, so the channel reuse ratio is low. Consequently, the number of assigned channels is increased.

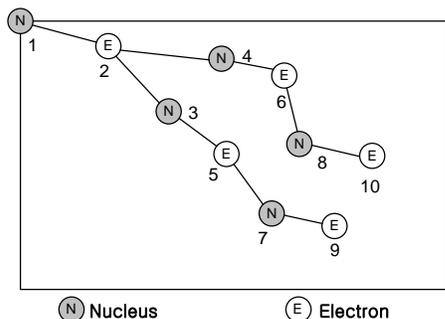


Fig. 2. Role assignment problem

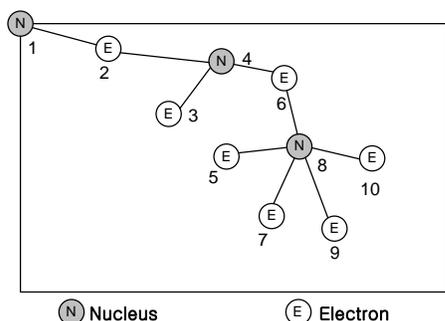


Fig. 3. Example of role assignment with the proposed scheme

Fig. 2 shows a problem of role assignment in the Molecular MAC protocol. There are 10 nodes in the figure, and node 1 is a gateway node connected to the wired Internet. Remaining nodes form a shortest path-spanning tree from node 1 and are assigned roles. Since the depth of nodes 3, 4, 7 and 8 is an even number, the role of a nucleus is assigned to each of these nodes. As the depth of remaining nodes is an odd number, the role of an electron is assigned to each of them. As shown in the figure, nodes 3 and 4, and nodes 7 and 8 are assigned as nuclei, even though they are located within the transmission range of each other. Thus, this scheme causes the waste of channels.

To solve this problem, this paper introduces a new metric, and proposes a new scheme of assigning roles to nodes. With this scheme, it is possible to prevent nodes located within the transmission range of each other from becoming nuclei, and allow the number of electrons supported by a nucleus approach the number which optimizes the network performance. Fig. 3 shows an example of role assignment with the proposed scheme. It shows that the number of nuclei is reduced from 5 to 3 compare to Fig. 2. Accordingly, the number of assigned channels is also reduced. More details are described in Section III.

III. PROPOSED ROLE ASSIGNMENT SCHEME

In this paper, we propose a new scheme to solve the problem of assigning roles to nodes in the Molecular MAC protocol. The proposed scheme adopts the basic operating principle of the Molecular MAC protocol without any modifications, though some parts related to the role assignment of nodes are modified. In the proposed scheme, a node which fully meets the following conditions is assigned as a nucleus: that is, a node with the optimal number of electrons and a node which can send data of electrons linked to it to a gateway node in a timely way. And then, to avoid selecting nodes within the transmission range of each other as nuclei, all nodes located within the transmission range of the node selected as a nucleus are assigned as electrons. To do that, we introduce a new metric, which is called as Expected Transmission Performance (ETP).

A. Overall Flow of Tree-Forming Process

In the proposed scheme, a node forms a tree to a gateway node by using the ETP metric. Each node is assigned one of the following roles: nucleus, electron, unassigned, or candidate nucleus. When the role assignment is completed, every node is assigned the role of a nucleus or an electron. An unassigned node is a node to which the role of nucleus or electron is not assigned in the tree-forming process. A candidate nucleus means a node, which has potential of being a nucleus among unassigned nodes in the role assignment process. Thus, it is also an unassigned node. In the first stage of making a tree, the role of every node is unassigned. A gateway node is assigned as a nucleus in default. The tree-forming process is as follows:

- 1) Every unassigned node within the transmission range of the gateway node is assigned as an electron, and forms a tree to connect to the gateway node.
- 2) Each electron node assigns every unassigned node within its transmission range as a candidate nucleus node.
- 3) For candidate nucleus nodes, ETP metric table as shown in Table I is made.

TABLE I. ETP METRIC TABLE

Electron ID	Metric Value

In the table, an electron ID is the identity of an electron which assigns a node as a candidate nucleus node in Step 2. How to calculate metric values is described in subsection III.B in detail.

- 4) After making an ETP metric table for all candidate nucleus nodes, a candidate nucleus node with the largest metric value is assigned as a nucleus.
- 5) A nucleus node newly assigned in Step 4 assigns every unassigned node within its transmission range as an electron and forms a tree by connecting them to itself.
- 6) The newly assigned nucleus node forms a tree by connecting to a node with electron ID of the largest metric value in Table I.
- 7) It resets every candidate nucleus node to unassigned status, and initializes its ETP metric table.

8) Until the role of nucleus or electron is assigned to all unassigned nodes, Step 2 to 7 is repeated.

B. Calculation of ETP Metric Value

As a nucleus, the proposed scheme selects a node that has optimal number of electrons and sends data of electrons to the gateway node at the maximum data rate. For this end, the following three conditions shall be met to apply the proposed scheme. First, a nucleus shall have optimal number of electrons, because the network performance is deteriorated if the number of electrons connected to the nucleus is too many or too few. Second, a nucleus shall be able to send data at high data rate to connected electrons. Finally, data transmission time from the nucleus to the gateway node shall be short. To satisfy the above conditions, the proposed ETP metric is made of the following three components: NO_{Node} , $DATA_{Rate}$, and TX_{Time} . Each metric component has a value of [0, 1], where „1“ is the best and „0“ is the worst. Thus, a candidate nucleus node with the largest ETP metric value is selected as a nucleus.

ETP metric component NO_{Node} is obtained as the number of unassigned nodes within the transmission range of a candidate nucleus node. Wireless mesh networks can be implemented with various wireless technology including 802.11, 802.15, 802.16, cellular technologies or combinations of more than one type. With the WLAN popularizing, IEEE 802.11 MAC protocol has been adopted as the de-facto medium access control of wireless mesh networks. In IEEE 802.11 MAC protocol, network performance is enhanced following the growth of the number of nodes with access points, but the performance is deteriorated if the number exceeds a certain value. Thus, the number of nodes, which shows the maximum network performance, is defined as the optimal number of nodes. When the number of unassigned nodes within the transmission range of a candidate nucleus node approaches the optimal number of nodes, NO_{Node} approaches 1. On the contrary, if there is a large discrepancy with the optimal number of nodes, then it approaches „0“. NO_{Node} is calculated as follows:

$$NO_{Node} = 1 - \left| \frac{Opt_N - n}{Opt_N} \right| \tag{1}$$

where, n is the number of unassigned nodes within the transmission range of a candidate nucleus node, and Opt_N is the optimal number of nodes. $|x|$ is the absolute value of x . When n is two times or more than Opt_N , this metric component may have a negative value. To avoid this, n is set to $2 \cdot Opt_N$.

When a candidate nucleus node is assigned as a nucleus, ETP metric component $DATA_{Rate}$ is calculated by considering data rates with connecting electrons. The candidate nucleus node measures the distances to unassigned nodes within its transmission range and checks the data rate with each node with the data rate table based on the measured distances. Table II shows the relationship between the rate and the distance. This table is from a reference [18]. Although the reference

showed the rate based on distance between two nodes, this paper is expressed the distance between two nodes after normalizing it. That is, a normalized distance is a value dividing the distance between two nodes by the transmission range of a node. Thus, it is expressed as [0, 1].

TABLE II. DATA RATE ACCORDING TO THE DISTANCE BETWEEN TWO NODES IN OPEN ENVIRONMENT IN 802.11B/G

Normalized Distance	Data Rate(Mbps)
0.125	54
0.213	36
0.300	18
0.498	11
0.649	6
1.000	1

With Table II, data rates between the candidate nucleus node and all unassigned node within its transmission range are checked and average data rate (Avg_{rate}) is calculated as follows:

$$Avg_{rate} = \sum_{i=1}^n Rate_i / n \tag{2}$$

where, n is the number of unassigned nodes within the transmission range of a candidate nucleus node, and $Rate_i$ is the data rate between the candidate nucleus node and an unassigned node i .

ETP metric component $DATA_{Rate}$ is calculated by dividing average data rate by the maximum rate (MAX_{rate}) supported by the network as follows. For example, the maximum rate in Table II is 54 Mbps.

$$DATA_{Rate} = \frac{Avg_{rate}}{MAX_{rate}} \tag{3}$$

Let us define a gateway transmission time and neighbor transmission time as the length of time required to send a data packet from a node to the gateway node and between neighbor nodes, respectively.

ETP metric component TX_{Time} is obtained by using the gateway transmission time and neighbor transmission time.

In Step 5 in subsection III.A, after assigning a candidate nucleus node as a nucleus, the nucleus node assigns unassigned nodes within its transmission range as its electrons, and then connects as a tree. At this time, the nucleus node sends the gateway transmission time to each electron node. In addition, each electron node adds the gateway transmission time received from its nucleus and the neighbor transmission time between itself and its nucleus. In Step 2 in subsection III.A, it sends the sum to the candidate nucleus node.

When forming a tree with unassigned nodes, each node sends the unassigned nodes the gateway transmission time from itself to the gateway node. When receiving the gateway

transmission time, an unassigned node sends another unassigned node the sum of the received gateway transmission time and the neighbor transmission time from itself to the previous node.

Let us denote data rate and neighbor transmission time of two neighboring nodes i and j as $Rate_{i,j}$ and $TxTime_{i,j}$. The relationship between the two is expressed as follows:

$$TxTime_{i,j} = \frac{1}{Rate_{i,j}} \quad (4)$$

Now, let us denote $TxTime_{n \rightarrow 1}$ the gateway transmission time from an arbitrary node n to the gateway node 1. The gateway transmission time is expressed as the sum of neighbor transmission times between neighboring nodes, through which data is passed when sending a data packet from a node to the gateway node through the formed tree, as shown in the following expression:

$$\begin{aligned} TxTime_{n \rightarrow 1} &= TxTime_{n-1 \rightarrow 1} + TxTime_{n,n-1} \\ &= \sum_{i=n}^2 TxTime_{i,i-1} \end{aligned} \quad (5)$$

where, n is an arbitrary node and 1 is the gateway node. $n-1$, $n-2$, ..., and 2 are nodes placed on a path from the arbitrary node to the gateway node in the tree.

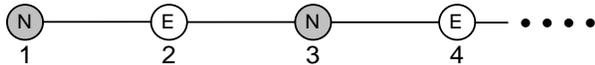


Fig. 4. Example of calculating $TxTime_{n \rightarrow 1}$

Fig. 4 shows an example of the process of calculating $TxTime_{n \rightarrow 1}$. There are 4 nodes in the figure, and node 1 is the gateway node. The gateway node is assigned as a nucleus, which assigns node 2 located within its transmission range as an electron and connects to the tree. And then it sends the gateway transmission time of $TxTime_{1 \rightarrow 1}$ ($= 0$) to node 2. Node 2 assigns node 3 within its transmission range as a candidate nucleus node, and calculates the gateway transmission time ($TxTime_{2 \rightarrow 1}$) from itself to the gateway node 1 and sends the result to node 3. $TxTime_{2 \rightarrow 1}$ is $0 + TxTime_{2,1}$. Node 3 is assigned as a nucleus, and then assigns node 4 within its transmission range as an electron and connects it as a tree. It then calculates the gateway transmission time ($TxTime_{3 \rightarrow 1} = TxTime_{2 \rightarrow 1} + TxTime_{3,2}$) from itself to the gateway node 1, and sends the result to node 4. In addition, node 4 also calculates the gateway transmission time ($TxTime_{4 \rightarrow 1} = TxTime_{3 \rightarrow 1} + TxTime_{4,3}$) and sends the result to the next node.

After every candidate nucleus node calculates the gateway transmission time, the largest value ($TxTime_{max}$) is selected. N is a set of candidate nucleus nodes.

$$TxTime_{max} = \max_{i \in N} TxTime_{i \rightarrow 1} \quad (6)$$

ETP metric component TX_{Time} of an arbitrary node n is calculated by using $TxTime_{max}$ as follows:

$$Tx_{Time} = 1 - \frac{TxTime_{n \rightarrow 1}}{TxTime_{max}} \quad (7)$$

After calculating three metric components NO_{Node} , $DATA_{Rate}$, and TX_{Time} , the ETP metric value (ETP) is obtained as follows:

$$ETP = \alpha \cdot NO_{Node} + \beta \cdot DATA_{Rate} + \gamma \cdot TX_{Time} \quad (8)$$

where, α , β , and γ are weighting factors for each metric component ($\alpha + \beta + \gamma = 1$). ETP has a value of $[0, 1]$, where „1“ is the best.

IV. SIMULATION RESULTS

In this Section, we discuss the simulation results of the proposed scheme. To study the performance of the proposed scheme, we have implemented it. We compare them to the results of the Molecular MAC protocol. We simulated an IEEE 802.11 network with transmission rates of 54 Mbps for data packets and of 6 Mbps for control packets such as RTS, CTS and ACK, respectively. Opt_N is 5. α , β , and γ are set to 0.3, 0.3, and 0.4 respectively.

The spanning tree construction proceeds as follows. First, the network elects a gateway node, which is connected to the wired Internet. Then, the other nodes construct a spanning tree rooted at the gateway node. We place the gateway node on the top-left corner and randomly the other nodes in a simulation topology.

Main performance metrics of interest are the number of nuclei, the number of assigned channels, the number of electrons per nucleus and the number of nuclei per electron. The number of nuclei is the number of nodes assigned as nuclei in the simulation topology. The number of assigned channels is the number of channels assigned to nuclei in the simulation topology. T

hat is, it is the number of channels required to serve all the nodes in the simulation topology. The number of electrons per nucleus is the average number of electrons connected to a nucleus. The number of nuclei per electron is the average number of nuclei connected to an electron. All simulation results are averaged over ten simulations. In the simulation result, the proposed scheme is expressed as ETP.

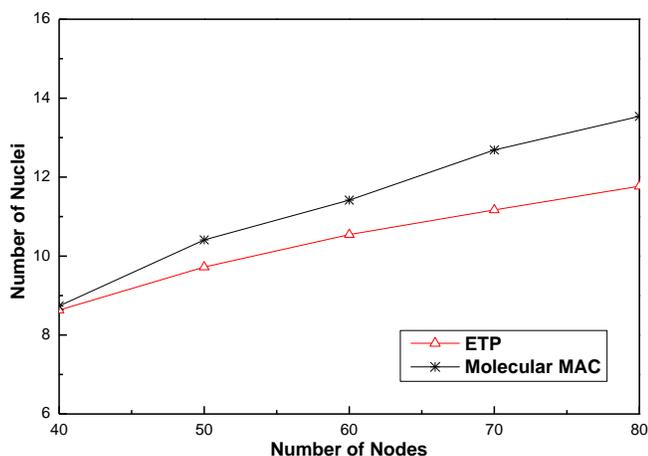


Fig. 5. Number of nuclei according to the number of nodes

Figs. 5, 6, 7, and 8 show change of network performance following the increase of the number of nodes used in the simulation topology. Fig. 5 shows that the number of nodes assigned as nuclei is increasing linearly in proportion to the number of nodes. When the number of nodes is 40, about 22% of the nodes are nuclei in both schemes. As the number of nodes is increasing, the proportion is decreasing progressively. Thus, when the number of nodes is 80, about 14.7% of the nodes are assigned as nuclei in the proposed scheme, while about 16.9% are in the Molecular MAC protocol. Fig. 5 shows that the number of nodes in the proposed ETP scheme is increasing slowly compared to that in the Molecular MAC protocol, and its proportion is also low.

Fig. 6 shows the number of channels assigned to serve all the nodes in the simulation topology. In the Molecular MAC protocol, the number of assigned channels is slowly increasing following the growth of the number of nodes. When the number of nodes is 40, about 3.7 channels are assigned; when it is 80 nodes, 4.8 channels are assigned. However, in the proposed ETP scheme, it is maintained constantly regardless of the number of nodes. As explained in subsection II.B, in the Molecular MAC protocol, each node forms the shortest path-spanning tree to the gateway node, and a node of which depth is even numbered is assigned as a nucleus. Thus, as shown in Fig. 6, the number of nodes assigned as nuclei increases. In addition, only even-numbered depth is considered in role assignment, neighboring nodes within the transmission range of each other become nuclei and use different channels from each other, the channel reuse ratio is lowered. As a result, the number of assigned channels increases. However, the proposed ETP scheme takes the transmission range into account, only a node separated at an appropriate distance is assigned as a nucleus. Thus, there is no possibility that a neighboring node within the transmission range is a nucleus. If two nuclei are separated at a certain distance, then both of them operate without a hitch when they are using the same channel. Thus, the channel reuse ratio increases, and the number of assigned channels is also grown slowly.

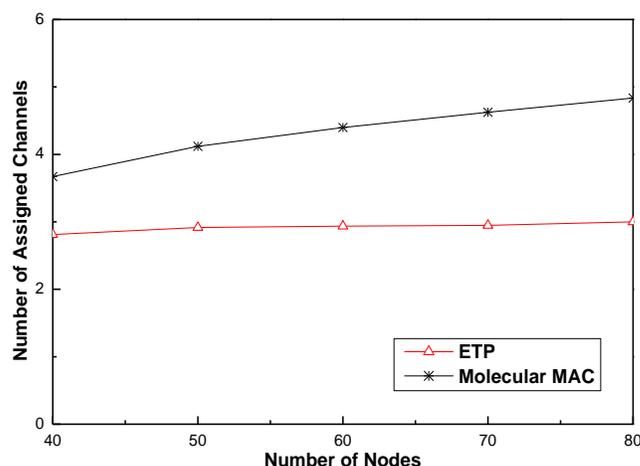


Fig. 6. Number of assigned channels according to the number of nodes

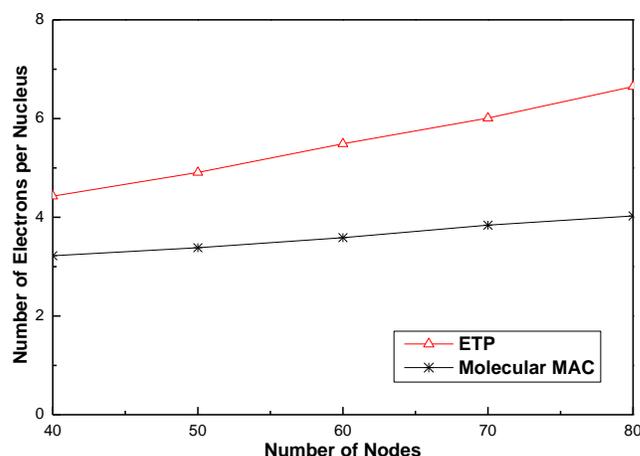


Fig. 7. Number of electrons per nucleus according to the number of nodes

Fig. 7 shows the number of electrons connected to each nucleus following the increase of nodes. As shown in Fig. 5, when the number of nodes is 40, that of nuclei in the proposed ETP scheme is 8.6, while that of the Molecular MAC protocol is 8.7. And the number of electrons is 31.4 and 31.3, respectively. In average, the number of electrons connected to a nucleus can be obtained by dividing the number of electrons by that of nuclei. In theory, the average number of electrons is 3.7 ($=31.4/8.6$) and 3.6 ($=31.3/8.7$), respectively. However, Fig. 7 shows somewhat different results: 4.4 and 3.2, respectively. The cause of such discrepancy is as follows: As shown in Figs. 2 and 3, a leaf node in the tree graph is connected to a nucleus, though an electron in the middle is connected more than two nuclei. In this way, an electron in the middle is connected to several nuclei, so the number of electrons per nucleus counts some redundant nodes, which makes such discrepancy. As shown in Fig. 5, the number of nuclei is increasing more slowly in the proposed ETP scheme than in the Molecular MAC protocol; thus, the number of linked electrons is increasing relatively faster.

REFERENCES

Fig. 8 shows the number of nuclei connected to an electron following the increasing number of nodes. As shown in Figs. 2 and 3, all the electrons are connected to more than one nucleus to communicate. Therefore, the result value shall be more than 1 in both schemes, though the Molecular MAC protocol shows a value of less than 1. The cause is as follows: In the Molecular MAC protocol, a node with even-numbered depth is assigned as a nucleus. However, exceptionally, a leaf node is assigned as an electron not as a nucleus, even if it has even-numbered depth, and then connected to electrons.

That is, the last node with even-numbered depth does not communicate with the nucleus. When such nodes are increasing, there is intra-flow interference, which may deteriorate the network performance. The proposed ETP scheme follows the operating principle of the Molecular MAC protocol basically. However, since the proposed scheme minimizes such nodes, it is able to ameliorate the deterioration of network performance.

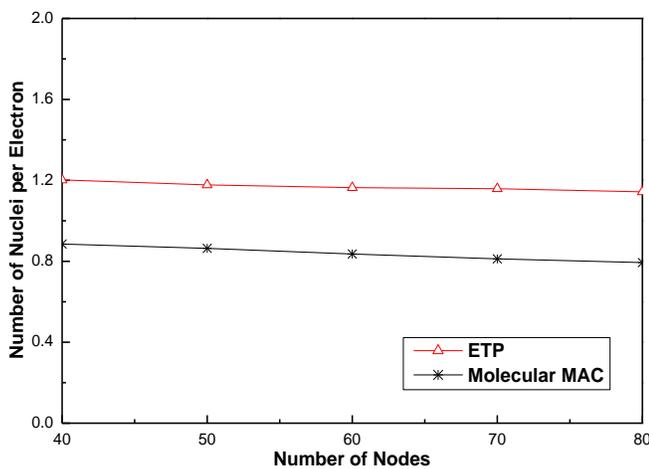


Fig. 8. Number of nuclei per electron according to the number of nodes

V. CONCLUSION

A wireless mesh network (WMN) is a promising wireless technology for numerous applications. The Molecular MAC protocol was proposed by adopting a molecular analogy. In the Molecular MAC protocol, the nodes with an even-numbered depth are assigned with the role of a nucleus, it supports many nuclei. Also, nodes within the transmission range of each other can become nuclei, so the channel reuse ratio is low. Consequently, the number of assigned channels increases. This paper proposes an efficient role assignment scheme, which can reduce the number of assigned channels by reducing the number of nodes assigned as nuclei and preventing nodes within the transmission range of each other from becoming nuclei.

The proposed scheme introduces a new metric by considering three components affecting network performance, and assigns the role of each node based on it. Simulation results show that the proposed scheme has better performance than the Molecular MAC protocol.

[1] I.F. Akyildiza, X. Wang, and W. Wang, "Wireless mesh networks: a survey," Elsevier Computer Networks, vol. 47, no. 4, pp. 445-487, 2005.

[2] A. Iyer and C. Rosenberg, "Understanding the key performance issues with MAC protocols for multi-hop wireless networks," Wireless Communications and Mobile Computing, vol. 6, no. 6, pp. 745-760, 2006.

[3] H. Zhai, X. Chen, and Y. Fang, "Alleviating intra-flow and inter-flow contentions for reliable service in mobile ad hoc networks," in Proceedings of IEEE MILCOM, vol. 3, pp. 1640-1646, Nov. 2004.

[4] M. Karaliopoulos, T. Spyropoulos, and B. Plattner, "Interference-aware routing in wireless multihop networks," IEEE Transactions on Mobile Computing, vol. 10, no. 5, pp. 716-733, May 2011.

[5] Y. Kim, D. Jung, Y. Kim, S. Choi, and J. Ma, "Efficient interference-aware channel allocation in multi-radio wireless mesh networks," in Proceedings of International Conference on Advanced Communication Technology (ICACT 2012), pp. 920-925, 2012.

[6] P.B.F. Duarte, Z.M. Fadlullah, A.V. Vasilakos, and N. Kato, "On the partially overlapped channel assignment on wireless mesh network backbone: a game theoretic approach," IEEE Journal on Selected Areas in Communications, vol. 30, no. 1, pp. 119-127, Jan. 2012.

[7] A. Subramanian, H. Gupta, S. R. Das, and J. Cao, "Minimum interference channel assignment in multi-radio wireless mesh networks," IEEE Transactions on Mobile Computing, vol. 7, no. 12, pp. 1459-1473, Dec. 2008.

[8] F. Theoleyre, B. Darties, and A. Duda, "Assignment of roles and channels for a multichannel MAC in wireless mesh networks," in Proceedings of 18th International Conference on Computer Communications and Networks (ICCCN 2009), pp. 1-6, Aug. 2009.

[9] B. Darties, F. Theoleyre, and A. Duda, "A Divide-and-conquer scheme for assigning roles in multi-channel wireless mesh networks," in Proceedings of IEEE 34th Conference on Local Computer Networks (LCN 2009), pp. 277-280, Zurich, Switzerland, Oct. 2009.

[10] D. Abdelali, F. Theoleyre, A. Bachir, and A. Duda, "Neighbor discovery with activity monitoring in multichannel wireless mesh networks," in Proceedings of IEEE Wireless Communications and Networking Conference (WCNC 2010), pp. 1-6, Apr. 2010.

[11] M. Nassiri, F. Theoleyre, M. Heusse, and A. Duda, "Molecular architecture for autonomic wireless mesh networks," in Proceedings of the ACM student workshop - CoNEXT 2007, Dec. 2007.

[12] F. Theoleyre, M. Heusse, and A. Duda, "Molecular MAC for multichannel wireless mesh networks," in Proceedings of IEEE 6th International Conference on Mobile Adhoc and Sensor Systems (MASS 2009), pp. 110-119, Oct. 2009.

[13] S. Pollak, V. Wieser, and A. Tkac, "A channel assignment algorithm for wireless mesh networks with interference minimization," in Proceedings of Wireless and Mobile Networking Conference (WMNC 2012), pp. 17-21, 2012.

[14] K.N. Ramachandran, E.M. Belding, K.C. Almeroth, and M.M. Buddhikot, "Interference-aware channel assignment in multi-radio wireless mesh networks," in Proceedings of IEEE INFOCOM, pp. 1-12, Apr. 2006.

[15] K. Xing, X. Cheng, L. Ma, and Q. Liang, "Superimposed code based channel assignment in multi-radio multi-channel wireless mesh networks," in Proceedings of ACM MobiCom, pp. 15-26, 2007.

[16] H. Skalli, S. Ghosh, S. Das, L. Lenzini, and M. Conti, "Channel assignment strategies for multiradio wireless mesh networks: issues and solutions," IEEE Communications Magazine, vol. 45, no. 11, pp. 86-95, Nov. 2007.

[17] S. Avallone, G. Stasi, and A. Kassler, "A Traffic-aware channel and rate re-assignment algorithm for wireless mesh networks," IEEE Transactions on Mobile Computing, vol. 12, no. 7, pp. 1335-1348, Jul. 2013.

[18] Cisco Aironet 802.11a/b/g Wireless LAN Client Adapters (CB21AG and P121AG) Installation and Configuration Guide, Appendix A, Technical Specification, <http://www.cisco.com>.

A DIY Approach to Uni-Temporal Database Implementation *

Haitao Yang, Fei Xu, Lating Xia
GuangDong Construction Information Center (GDCIC)
Guangzhou, China

Abstract—when historical versions of data are concerned for a MIS (Management Information System) we naturally might resort to temporal database products. These bi-temporal products, however, are often extravagant and not easily mastered to most of MIS practitioners. Hence we present a plain DIY (do it yourself) solution, the Audit & Change Logs Mechanism-based approach—ACLM, to meet the uni-temporal requirement from restoring historical versions of data. With ACLM programmers can code SQL scripts on demand to trace and replay any snapshot of historical data version via RDBMS built-in functions, they need not to shift away from their usual way of coding stored procedures for data maintenance. Besides, the ACLM approach is compatible with meta-data change, and its additive overhead was instantiated imperceptible for throughputs of routine access with a typical scenario.

Keywords—DIY solution; recurrence; historical snapshot; uni-temporal database; MIS

I. INTRODUCTION

Recent years' practices in developing web creditable MIS for numerous users made us all realized that maintaining enormous records of users-oriented data are an unbearable task to anybody (individuals or organs). Regarding that it is users or clients' own right and responsibility to keep their delivered data valid and complete, we were looking forward to an interactive and sharing pattern that all users involved should honestly maintain their information themselves, which is perhaps the only workable approach. Then, there comes the risk of abuse of self-maintenance right. To prevent such a risk we have no feasible solutions of instant response but can build a final line of defense by an ex post facto measure that is, logging all behaviors of maintaining data and offering a facility to restore or reveal any historical version of concerned data and the responsible manipulators. Herewith we get to the field of uni-temporal database application, and might assume products of temporal DBMS (in short, tDBMS) as a matter of course. tDBMS products, however, are bi-temporally-oriented, and not familiar to most of practitioners in ordinary MISs. Even worse, such products often offer extra functions well beyond need and bring with much greater complexity and higher cost than expected. Upon these considerations, we turned to explore an exercisable and methodological approach (that why we refer to it as DIY -- do it yourself).

II. RELATED WORK AND OUR DIRECTION

In realistic applications, data recorded in databases all have certain time properties either explicit or implicit, at least those indicate when the data are valid and when they are recorded [6]

— the former is a time property with data semantics, classified as the valid-time property, and the latter is a time property with data operation, categorized into the transaction-time (time of manipulating data) property. Contemporary RDBMS products have granted us the ability to straightly store and manage all relational data including temporal data in the same database, notice that data's time property is also a datum. To certain extent, temporal data is an issue of data versions that concerns with data recordation at different times. It is not feasible for all time versions of data are treated in equivalence, since eventually the ever-increasing amount of historical versions will become overwhelming on all aspects of data storage and usage. Approaching such a problem and its related, a study of temporal DBMS has been developed for decades [1]. Despite lots of research on tDBMS, practical tDBMS products are rare, and even more, most of them are in fact an extension of traditional RDBMS [11], in general developing a tDBMS application is still a tough and often individualized task for a MIS (Management Information System) developer team.

In usual practices of developing MIS, we normally design and develop a database application around usage of the newest data version, because in default, people much concern with the current status rather than those historical. If no requirement of recalling a historical "snapshot" (we use this term to refer to *a picture of data at a historical time*), historical statuses of data will be updated or overlaid by the newest one, and everything is just simple as usual. But if the responsibilities of conducting data change is of the concern, e.g., they must be audited or traced afterwards (a lot of crucial MIS applications have this requirement), in which historical statuses of data need to be carefully and explicitly addressed, we actually step in the scope of tDBMS. Up to now, tDBMS approaches in mainstream, such as the famous ATSQL2 proposal are conducted in a direction of treating time properties of data as an abstract or super attribute with special disposal [10], and they are based on relational data processing, of which the most concerned are often associated with their special temporal data type, temporal manipulation on table or column-level, and dedicated temporal relation constraints etc. In these practices, accordingly, special time attribute-oriented extensions to SQL must be introduced to and well supported by tDBMS. Despite temporal SQL-compliant research has been very comprehensive these days, however, the related SQL-level support mechanism, temporal database model theories, etc., are sophisticated and too much for most of applications just involved with some plain temporal requirements as in ordinary MIS. Most of MIS practitioners would rather treat temporal parts of MIS applications in a similar way as in ordinary DBMS programming practice, e.g.,

* This work is sponsored by Guang-Dong Construction Information Center.

assigning each intended time property of data into a concrete data attribute, and so on. Thereby, we prefer a system mode of “traditional RDBMS” + “software”. Here, the “software” could be programs as a part of the hosting application itself, or in an embedded type as a third-product software product (often as a middleware), e.g., the well-known TimeDB [12] is a RDBMS-based embedded middleware for temporal data applications, TimeDB runs as a frontend to the hosting RDBMS (e.g., Oracle) and supports the temporal query language ATSQL2, where finally ATSQL2 statements are compiled by TimeDB into (sequences of) SQL-92 statements which are executed by the underlying RDBMS backend. The pure temporal disposal part (software) of TimeDB is to interface between the temporal usages (delivered in ATSQL2 statements by users) and SQL-92 executions.

As we know, tDBMS products such as TimeDB often store the transaction-time of data in the same data tuple. Such a device is inefficient for OLTP (On-Line Transaction Processing), considering that if data items are frequently updated, the data table where these items reside will soon be overwhelmed by historical versions of data, which we figured as a so-called 99 to 1 % phenomenon, i.e., 99% (symbolizing most) data records are for past statuses while 1% (symbolizing a little proportion) for the current or latest status, while most of accesses to the data table are just for the 1% records. On this aspect, we believe that many MIS practitioners like us would rather try an on-hand and less costly scheme than take an abstruse academic approach or purchase an often too costly or heavy tDBMS product.

For generality and practicability, basing on the above consideration and rules of engineering we believe that a good approach for the issue discussed should be a methodological one with ease of use or duplication, in the other word, a DIY (Do It Yourself) type, and which should adopt an outline pattern of separating historical versions of data from the current one, and provide guidelines for designing fundamental maintenance, management and utility services of data. Along this direction, we start our approach by introducing several key concepts in a simple but typical example about temporal data recordation:

$P1 = (\text{“John”}, 2000, \textit{interval_1})$ | a data record with its valid time indicating *John* has a salary of two thousands dollar.

$P2 = (P1, 11/5/2012\ 4:44\ \text{PM})$ | the above data record $P1$ was created or updated at 11/5/2012 4:44 PM.

The statement from $P1$ is true only for its valid time period *interval_1*, but what from $P2$ that recorded a fact is always true. The valid period *interval_1* can be definite as $[time1, time2]$, or indefinite as $[time1, unknown]$ that spans from *time1* until something happens (e.g., John’s salary is changed or he is dismissed, etc.) when the *unknown* becomes a certain value. Generally we are not likely to maintain valid time properties via an automatic mechanism since they are associated with concrete semantics of the data they modify, as in the above simple example, when and how to make the unknown time known is up to the intelligence of realizing the corresponding event and its relevancy; but it is different with transaction time properties since they simply denote a data manipulation event that can be monitored via certain DBMS built-in mechanisms.

We doubt in nature there is any universal automatic scheme to cope well with storage, management and usage of valid time property of data, despite lots of techniques on related issues. In the other side, for a real relation object its valid time attribute and its other attributes are all in an equal position with respect to relational data theory and application semantics, thus they could be and should be treated equally if convenient.

Further, we clarify three key facts which are often ignored: 1) the **valid-time property** of data virtually can only be actively determined by who understanding the data meaning, perhaps an intelligent software can do this, but developing an intelligent software is far beyond the scope of applied tDBMS research; 2) for web data applications, the responsibility audit about data manipulation could not be done within DBMS since conventionally different web users share a common DBMS account; 3) data structure changes cannot be excluded in real applications, for instance, adding or retiring a field in a data table (in practices, a *relation* is often instantiated as a table of records, an *attribute* as a *field* of record in the table) for one or other reason is allowable.

Accordingly, we have three keynotes for a feasible tDBMS implementation: (1) the valid-time property of data should be considered in the context of data application; (2) the responsibility audit of data manipulation needs participation from higher layers of application outside DBMS; (3) a good implementation mechanism for tDBMS should be compatible with ordinary structure change of data tables. Following these guidelines, towards a generic design scheme for tDBMS-related MIS applications we focus on the **transaction-time** property of data (so we call this approach as of the *uni-temporal* database implementation) and treat it as a common attribute [9], this is contrasted with the nowadays so-called “bitemporal database” [11], i.e., a general temporal database.

III. V+A FRAME FOR TDBMS APPLICATIONS

Along the above decided direction of investigation, the underlying thing is to set up a software frame for tDBMS applications.

Firstly, concerning temporal evolvement of data content we noticed two main disposals of transaction-time of data: (1) using self-contained temporal recordation of data manipulation, often in a vitae form, i.e., each maintenance manipulation on a data item should append its execution time to all data records changed, there is no need for additional logging mechanism; (2) devising dedicated temporal logging mechanism for recording data manipulation activities, which notes down the transaction-time in a log separate from the data table.

Secondly, for the sake of practicability we shall take into account a common phenomenon of meta-data change — change to the composition of data table’s PK (primary key), such as using different component fields or altering the data type of some component fields — which was often ignored by most of tDBMS approaches.

Thirdly, Application requirements on storing, managing, and hereby using time properties of data are versatile, but for usual cases of MIS they can be classified into two types: the time property that is used frequently or routinely should be accessed easily, whereas the others without routine usage can

do with less convenience of access for a much lower cost of implementation.

Accordingly, we proposed the Vita + Audit (in short, V+A) frame. The kernel of V+A frame consists of DRB (direct-retrieval base) + ACLM (audit & change logging mechanism). DRB is used to store usual data content, i.e., current-status data and temporal recordation in a vita form for direct accesses of routine transactions. Contrasting with DRB ordinary services for direct content access, ACLM is for dedicated audit accesses regarding data manipulations with transaction time — it keeps trace of each activity of data manipulation, memorizes into the change log the data snapshot of the data version just before the data manipulation exerts on DRB each time, and at the same time inserts into the audit log a record about what kind of data manipulation and who makes that manipulation. Under ACLM we should not miss any historical version of data being audited though we could not view directly its content in a single SQL manipulation.

In general, data query operations need not being logged in the audit log except for applications with extremely high safety demand since they do not create any new version of data, neither content of an *insert* operation needs recordation since it has no previous version. But the *insert* operation itself shall be recorded in the audit log in order to restore historical versions of data table before the operation timestamp. One main usage of the audit log is to record the responsible subjects of data maintenance – the actual operators from client end (terminal users) instead of those common DBMS accounts on the web data layer. In convention of software industries, user identity certification for web applications is fulfilled before calling functions of web data layer, and normal accesses to a web DBMS are requested via some shared DBMS account. To log the identity of a user (who instructed DBMS to execute a data manipulation) into a record of the audit log, the web user certification information should be passed into a corresponding inner procedure of the web DBMS. Manipulation of changing data content (Update, Delete, or iNsert) and its recordation in the audit and change logs should be treated within a single DBMS transaction as an atomic action (either both succeeded or anything they did will be withdrawn completely afterwards). Such transactions shall be coherently fulfilled through a stored procedure of DBMS script on the intermediate layer.

A. Audit Log

For each application the audit log is unitary, it is used in a way similar to keeping accounts of any change or comment on any data record of DRB: (1) recording any SQL-update, delete, and insert manipulation; and (2) logging any responsible comment on a data snapshot. The former is oriented to generic syntactic audit while the latter is about important semantic audit. The audit log in ACLM is application-oriented, i.e., all data tables from the same application share a unitary audit log. The structure of the audit log is defined as relation *Adt_log* described in Table I (data type in this paper are all given as in Oracle DBMS). Complementarily, a PK specification defined as relation *PK_spec* in Table II is introduced for all involved versions of PK structure of each data table from the same application, where one row of specification is for a member attribute of a PK.

TABLE I. AUDIT LOG (RELATION ADT_LOG)

Seq.	Attribute name	Data type	Remark
1	<i>Audit ID</i>	Varchar(32)	The PK attribute
2	<i>D table name</i>	Varchar(32)	
3	<i>D key value</i>	Varchar2(128)	Convert to String
4	<i>timestamp</i>	date	Time and day
5	<i>comments</i>	Varchar2(512)	
6	<i>Operation type</i>	Char(1)	C/D/U/N
7	<i>Operator id</i>	Varchar(20)	
8	<i>signature</i>	Varchar(172)	Sha1RSA

TABLE II. PK SPECIFICATION (RELATION PK_SPEC)

Seq.	Attribute name	Description	Data type
1	<i>D table name</i>	Be referred in Table I	Varchar(30)
2	<i>PK attribute name</i>	PK member attribute	Varchar(512)
3	<i>PK attribute seq</i>	Sequence number of this member attribute	Integer>=0
4	<i>Struct_Valid_S_time</i>	When this PK structure became valid	Date

Further explanation of the audit log’s definition and related usage are detailed as follows:

- 1) *The basic design of the audit log is applied directly to data tables with a single attribute PK (unitary PK). In relation Adt_log, attributes D_table_name is used to store the name of the data table being audited, D_key_value is used to store the PK value of the data record being audited. As to data tables without a unitary PK, more additive disposal is needed, see later in section IV. In relation PK_spec, PK_attribute_seq=0 is corresponding for unitary PK cases, while PK_attribute_seq>0 for non-unitary PK cases.*
- 2) *In fact, all MISs should know the PK composition of their data tables via something like PK_spec in advance of executing data maintenance. With PK_spec we need not include PK_attribute_name in Adt_log, which can prevent a transition dependency <D_table_name, PK_attribute_name> occurs in Adt_log.*
- 3) *To cope with meta-data changes, Struct_Valid_S_time attribute of relation PK_spec is set to indicate the start time that a version of PK composition became valid. The expired time of a valid PK composition is given subsequently by a next value of Struct_Valid_S_time in sequences for the same data table.*
- 4) *Attribute Operation_type has a set of basic values {D (Delete), U (Update), N (iNsert)} and an extended value C (Comment). Attribute Comment is used to record any responsible literal comment (including endorsement) on the data record being audited, it is left empty (assigned a null value) when Operation_type<>C. Value usages of attribute Comment can be extended and further categorized if needed in applications, e.g., classified into Censor and Verification, etc.*
- 5) *Attribute Timestamp is used to note down the time when the current audit record was created. To avoid ambiguity it is stipulated that the time of a web server’s clock be adopted, and relevantly-logged data changes take effect just after the instant of Timestamp.*

6) Signature is for storing the result of RSA calculation of Hash value of objects being signed by terminal users with their private key [8, 3]. A signed object consists of all content attributes (except maintenance and auxiliary attributes) of the data record under audit, and attributes from seq. 1 to 7 in *Adt_log*. The Hash value of a signed object is computed on the concatenated contents (all converted into the string type) of each involved attribute.

7) Any Update manipulation to alter a PK's value shall be equivalently decomposed into a Delete manipulation on the data record with the present PK value, and a subsequent *iNsert* manipulation of the updated data record with a new PK value.

Locking a data table during submitting a comment on its data record could lower the risk of mismatching the comment with a newer data version that was being created in the same time. Of course, freezing the data table for the whole process of comment action can exclude such a risk completely, but which will bring along with a more serious problem that normal data maintenance might be blocked for an uncertain (at worse often rather long) time by some comment activity, and the situation probably become even worse if the comment right is abused. Thus we in practice shall set a threshold of time limit for locking (e.g., 10 minutes) to avoid involving sophisticated lock/unlock mechanism. Conclusively, we have several more principles of using audit log:

- Applicable to record verification results in a generalized form of literal comment.
- To log each behavior of deleting, inserting, updating or verifying a data record, and the manipulator's digital signature about the essential content of the audit record in the same audit record.
- Separating historical data's storage, i.e., they are kept elsewhere (in the change log).

B. Change Log

The direct usage of change log is to record any data version just before it become outdated, which enables the occurrence of any historical data snapshots later. We shall record in the change log the current value of each data attribute bound for a content change just before the change operation is carried out, and the change operation being taken shall be noted down in the audit log at the same time. The data structure of change log is as defined in Table III, where attribute *valbfchg* stores the value-before-change for attribute *chgfldname* that stores the name of an attribute undergoing a value change, while the expiration time of content of *valbfchg* is indicated by attribute *timestamp* from a correlated record (being correlated through the value of *Audit_ID*) in the audit log *Adt_log*. For example, *Adt_log.timestamp="time1"*, *Chg_log.chgfldname="name"* and *Chg_log.valbfchg="John smith"* specified that data attribute *name* had a value of "John smith" just before time "time1".

The value-before-change of each data attribute (indicated with the content of *chgfldname*) that underwent a value change, except of lob type (Clob/Blob) shall be consistently converted into the string type and then put in attribute *valbfchg*. If a changed attribute is of lob type, its value-before-change shall be deposited in attribute *lob_value* while *valbfchg* is left empty.

For data attributes of binary lob type, we shall use an additive attribute *ContentType* [4] to further specify their content type to facilitate web applications for presenting such content.

TABLE III. CHANGE LOG (RELATION CHG_LOG)

Attribute name	Description
Audit_ID	a PK attribute
Chgfldname	a PK attribute
valbfchg	Direct value before change
datatype	String/clob/blob
Lob_value	If datatype=C/Blob
ContentType	For lob type data
Hash	For lob type data
Chg_act	U/D

C. ACLM Operation

The procedure of ACLM operation is outlined as follows:

1) A web service of data maintenance calls a DBMS stored procedure [7] to execute a dedicated data manipulation (*iNsert* | *Update* | *Delete* | *Comment*), noticing that *C* type operation is writing to the comment attribute of the audit record;

2) The DBMS stored procedure fulfills the data change on each target data table and correspondingly inserts a new audit record into the audit log with the matched type assigned to the attribute *Operation_type* within a single transaction;

3) Triggers of each target data table are ignited [7] to insert corresponding log records into the change log.

This software mechanism is illustrated as Fig. 1.

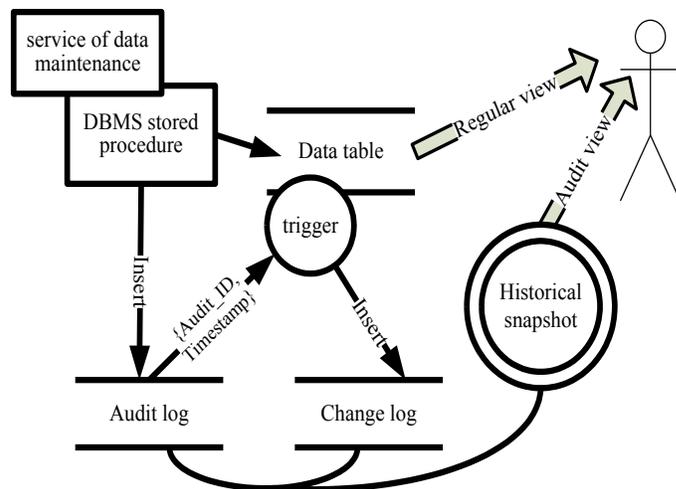


Fig. 1. Diagram of tDBMS software operation

Under ACLM, whenever calling a DBMS stored procedure to fulfill a process of data change (Update, Delete, *iNsert*) or data inspection (Comment) it is requested to specify whether to simultaneously write the audit and change logs, if yes (ACLM function enabled normally) then the calling program shall also ascertain the timestamp of logging a record into the audit log through an interface of the invoked DBMS stored procedure, regarding that a web server's clock is adopted, see paragraph 5) of subsection III.A. Since digital signature can only be made in

client ends where user private keys are available, and attribute *timestamp* in relation *Adt_log* is one of digital signature objects, thus the client end must get the time of web server just before user making digital signature. An individual ACLM operation involves two basic actions: A1 — exerting data changes, and A2 — logging such actions in both the audit and change logs. We propose not to execute A1 and A2 in separate web services, since combining two web services as a transaction (all done or nothing) will involve very sophisticated disposal, e.g., if A1 is successful but A2 is not, A1 has to be rolled back, then we have to pay off the cost of rolling back A1 due to A2; and what is more, sometimes (due to poor communication qualities) we cannot judge if A2 is successful or not (it might succeed but its reply was lost or simply delayed), unless we decide according to a time limit, but specifying a time limit is a trade-off issue, often very subtle.

Thus, we shall request a DBMS to conduct both A1 and A2 via a single web service call, which means whenever submitting a data change call we should have the accompanying *signature* value prepared for the audit record at the same time (must in advance obtain the values of *timestamp* and *Audit_ID* as parts of content to sign).

The above description implies a sequence of steps for carrying out a data change manipulation under ACLM: (1) pre-read content of a target data record to prepare the change action, (2) get the web server's time for digital signature, (3) submit the manipulation request to the hosting DBMS via a web service, (4) log the manipulation into the audit and change logs. Here, we have an order of timestamps: *timestamp* (pre-read) < *timestamp*(sign) < *timestamp*(submit) < *timestamp*(log). For better uniform simplicity, *timestamp* (pre-read) is adopted to substitute the rest of timestamps. This is because:

(1) it wouldn't influence consistency of retrospectively historical snapshots; (2) be competent for re-showing historical snapshots for cases without demand of extremely precise accuracy; (3) it is impossible within a one-off web calling to include into the digital signature a precise time of writing data table. Regarding that the web server functions as the centrum of ACLM, it is proper to grant attribute *timestamp* with the reading of the web server's clock.

Note:

1) Triggers of data table need to read contents of attribute *Audit_ID* and *timestamp* from the corresponding record of the audit log. Each record of the change log correlates to a unique record of the audit log, whereas each record of the audit log correlates to a group of records of the change log except those audit records of non-change type (no changing any existent data, e.g., insert or comment type), records from both logs are correlated via values of a common attribute *Audit_ID*.

2) If a round of data change process begins at a halfway phase (one or several rounds of data change were executed before, but none of them are regarded complete, i.e., all of their execution results are halfway, and saved into their DRB data table temporarily), then we strongly suggest that only enabling ACLM logging function for the first round of operation process

since all midway versions of data change in the same process transaction need not logging.

3) We shall not enforce ACLM function indiscriminately for all data tables without considering the additive overhead. For example, when all historical data versions of a data table are in fact presented as direct content, there is no need to log data changes anyway. In developing ACLM for MIS, we suggest to set up a configuration table of data change audit individually per application to specify together all involved data tables and their involved fields whose value's change need to be logged.

D. Typical Applications of Audit & Change Logs

1) Restore a record's snapshot at an audit timestamp

Let's take a scenario of reverting to a historical snapshot of data record at an audit timestamp. For data table *Tx*, let *Ax* be the audit log record with the timestamp *ts*, *Kx* be the PK value of *Tx*'s data record *Rx* that was audited by *Ax* at *ts*. If *Ax* is of C type, we have to track down along the time axis to the point (if any) when *Rx* underwent a change (update or deletion) after *ts*, see Fig. 2.

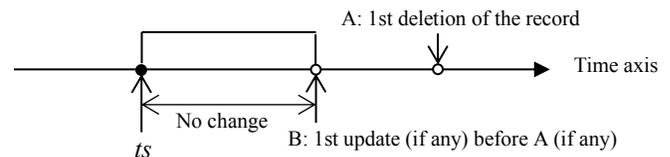


Fig. 2. Snapshot Recurrence Scene concerning a Specific Attribute

```
Create table snapshot(fieldname string not null, value string null, auditID
string null, chgtime timestamp null);
snapshot_D obj=new snapshot_D(ts, Tx, Kx);
Class snapshot_D {
Var timestamp D_time=null;
Public snapshot_D(timestamp tstamp, string tname, string keyval) {
Insert into snapshot (fieldname) select column_name from cols where
table_name=tname;
Var string auditID=null;
Select min(a.timestamp) into D_time From adt_log a Where
a.timestamp>=tstamp and a.D_table_name=tname and
a.D_key_value=keyval and Operation_type="D";
If D_time=null then exit; /* after tstamp Rx was not ever deleted */
Select a.Audit_ID into auditID From Adt_log a Where a.timestamp=D_time
and a.D_table_name=tname and a.D_key_value=keyval and
a.Operation_type="D";
Update snapshot s Set s.value = (Select b.valbfchg from chg_log b Where
b.Audit_ID=auditID and b.Chgflname=s.fieldname);}}
```

Fig. 3. Recurrence of *Rx* snapshot just before *Rx*'s first deletion after *ts*

Regarding that *Ax*'s existence implies *Rx* existed at *ts*, so we can restore the snapshot of *Rx* at *ts* through the following processes (be succinct, no datum of lob type is involved here):

Step 1: We shall check if *Rx* underwent a *D* type change or a PK value change after *ts* — the later is equal to a *D* type change being followed by a *N* type one, since in ACLM it is implemented by deleting the current record and subsequently inserting a record with the new PK value — if yes, then we shall firstly restore the snapshot of *Rx* just before *Rx*'s first deletion after *ts*, e.g., just before point A as in Fig.2.

```
snapshot_rollback roll_obj=new snapshot_rollback(ts, obj.D_time, Tx, Kx);
Class snapshot_rollback{
Public snapshot_rollback(timestamp tstamp, timestamp Dtime, string
tname, string keyval) {
SET ANSI_NULLS OFF;
Create view aud_chg_v AS /* (var >= null) == false */
Select a.Audit_ID, a.D_table_name, a.D_key_value, a.timestamp,
b.Chgfldname, b.valbfchg From adt_log a, chg_log b Where
a.D_table_name=tname and a.D_key_value=keyval and
a.timestamp>=tstamp and not(a.timestamp>=Dtime) and
a.Operation_type="U" and a.audit_ID=b.audit_ID
Update snapshot s Set s.chgtime = (Select min(c.timestamp) From
aud_chg_v c Where c.chgfldname=s.fieldname);
Update snapshot s Set s.auditID = (Select c.Audit_ID From aud_chg_v c
Where c.timestamp=s.chgtime and c.chgfldname=s.fieldname);
Update snapshot s Set s.value = Coalesce((Select b.valbfchg From chg_log
b Where b.Audit_ID=s.auditID and b.chgfldname=s.fieldname),
s.value);}}
```

Fig. 4. Reverting to Rx's snapshot at time ts

The process of this step is illustrated by a procedure as coded in a pseudo-java+SQL language in Fig.3, where it is fulfilled by creating a class instance $obj= snapshot_D(ts, Tx, Kx)$ regarding that $obj.D_time!=null$ indicates *yes*.

Step 2: We shall search the earliest U type change during $[ts, obj.D_time)$ for each attribute of R_x , and roll back together such a change (if any, e.g., at point B as showed in Fig.2) for any attribute whose value was changed during $[ts, obj.D_time)$ to get the snapshot for R_x at time ts . The task of this step is illustrated by a procedure as coded in a pseudo-java+SQL language in Fig.4, where it is carried out by setting $roll_obj=snapshot_rollback(ts, obj.D_time, Tx, Kx)$, regarding that $obj.D_time=null$ means no change of D type was on R_x after ts , and then $not(a.timestamp>=Dtime)$ becomes true accordingly due to $Dtime=null$.

Notice: different fields of a data record might undergo a content change at different times. The correctness of executing SQL scripts in Fig. 3 and 4 (where the procedure of opening database is omitted) rests with that all audits on the same object are sequential, i.e., no more than one audit action exerting on the same object is allowed at the same time.

2) Recuring to a table's snapshot at arbitrary time

To recur to the snapshot of table T_x at a given time ts is to restore exactly all data records that appeared at ts . Supposed ACLM has been functioning since ts , and then the recurring procedure can be outlined as follows:

a) To retrieve those T_x 's records that have not undergone any change since ts , we have

$S_1=\{select * from T_x where (T_x's PK) not in (select D_key_value from adt_log where D_table_name="T_x" and operation_type<>"C" and timestamp>=ts)\}$.

b) To restore $S_2=\{those T_x's records that existed at ts and underwent a change after ts\}$, we shall collect the set S_c of any audit record that logged the first change data operation on the same data record after ts :

$S_c=\{select Audit_ID, operation_type from adt_log where timestamp in (select min(timestamp) from adt_log where D_table_name="T_x" and operation_type<>"C" and timestamp>=ts group by D_key_value)\}$;

and then we exclude those audit IDs whose audit records logged a N type operation (As to a data record, after ts the first

change data operation is of N type implies that the data record didn't exist at ts otherwise it can not be inserted):

$S_a=\{select Audit_ID from S_c where operation_type<>"N"\}$.

c) For any audit record Ax whose ID is in S_a we restore the snapshot of the corresponding data record at Ax 's timestamp through a process described in subsection III.D.1). It is easy to prove that $\forall x \in S_2 (\exists y \in S_a)$ that y logged the first U or D change of x after ts , whereas $\forall y \in S_a (\exists x \in S_2)$ that the first U or D type change of x after ts was logged by the audit record with $ID = y$. As a result of the above process, we can regain each member of S_2 . The $S_1 \cup S_2$ is the wanted.

3) Practical simplicity for efficiency

The above applications are generally-oriented that each time a data record underwent a U type change the ACLM logged only those attributes that had undergone actual content change and their values. In practice, however, it was very clumsy to tell which field's value of data input interface has actually been made different from its existent value in a web submission of data maintenance input, and if such actual value change is judged within a trigger procedure then the execution efficiency of the trigger and thereby the hosted DML operation will be greatly abated. For the sake of simplicity and efficiency, the data input submitted (if accepted) is directly delivered to DBMS for an update operation to replace the existent values of the object attributes respectively without further distinguishing the existent and new values. In addition, to make easier the snapshot recurrence we should set ACLM to log all attributes and their existent values into the change log at the same time whenever a U type operation encountered though at more cost of storage space. Such a disposal can save a lot of computation, regarding that in this way each time an audit record closest behind to the given time is enough for snapshot recurrence without bothering to dig out all involved audit records for all attributes' snapshots that were logged at different timestamp.

IV. TESTING ACLM'S INFLUENCE ON DBMS OPERATIONS

Applying ACLM means appending certain additional audit and log tasks to normal DML operations in exchange for the competence of tracing versions and their responsible persons. There comes an issue of evaluating additive overheads from ACLM. Intuitively, we have two plain measurements for this evaluation: (1) *the perceptible performance decline*, (2) *the increment in comparative execution time*. In fact, as for MIS applications featuring human-computer interaction it is the measurement (1) much more suitable than (2) though the latter is more precise than the former. The *perceptible performance decline* can be well evaluated in terms of success ratios of maintenance operation per unit time for a normal range of request throughput into the hosting DBMS. On this aspect we made a succinct test to checkout if the ACLM influence is acceptable or not: per one minute how many percentage of update or delete operations succeeded for a large scope of operations throughput. The data table for test is defined in Table IV, its content attributes (*Citizen_ID_num*, *Reg_name*, *Reg_text*) were under audit of ACLM. For comparison, during the test twin instances of relation *Cer_Reg* were created such that one instance is ACLM-enabled while the other is non-ACLM, and both were exerted hundreds of thousands *Update*

and *Delete* operations. The test results were depicted in Fig. 5 and 6.

TABLE IV. CER_REG

Seq.	Attribute name	Data type	ACLM
1	<i>Citizen ID num</i>	Number(18)	X
2	<i>Reg name</i>	Varchar(32)	X
3	<i>Reg text</i>	Varchar(1024)	X
4	<i>timestamp</i>	timestamp	
5	<i>Process Status</i>	Char(1)	
6	<i>Lock Person</i>	Varchar(32)	
7	<i>Operator id</i>	Varchar(20)	
8	<i>signature</i>	Varchar(172)	

For a widely comparison of significance, we simulated a broad range of data maintenance frequencies covering and well beyond the statistic scope of *page view* (PV) of our web site (a well known industry website in our province) whose home page was accessed about 350 times per minute (on average the frequency of home page access corresponds to that of the web underlying DBMS inner DML request in the case of data maintenance) at its all time peak -- in Fig.5 and 6 the peak PV value was indicated by the vertical green line. Both Fig.5 and Fig 6. actually recorded how many operations were successfully fulfilled within one minute.

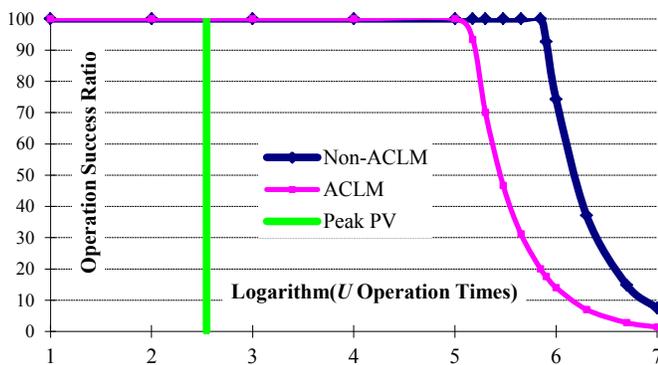


Fig. 5. Test results on Update (U) operations

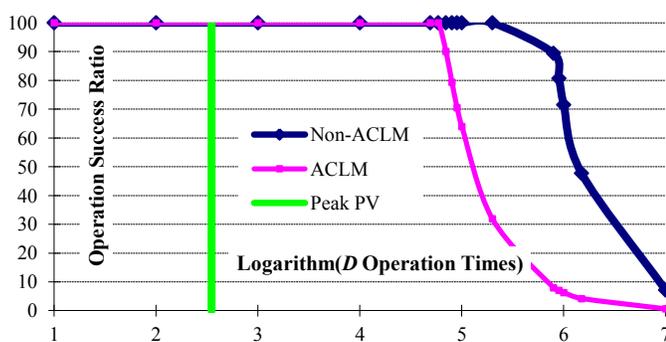


Fig. 6. Test results on Delete (D) operations

As in Fig. 5, both the ACLM-enabled (labeled ACLM) and ACLM-disabled (labeled non-ACLM) cases had similar performance lines (all in values of logarithm to base 10) for *Update* operations. Their success ratios hold 100% from the start at lower throughput of data operations up to some thresholds well above the site peak PV frequency, and then dropped rapidly as the impact of access throughput became very heavy. Although the success ratio of the ACLM-enabled

system dropped relatively earlier, its abrupt decrease point only came about as the access frequency reached an extremely high level, say 150,000 times of inner update operations requested per minute in the test, which is pretty rare for a normal website and can be referred to as an ultra situation. Fig.6 told the same thing about *Delete* operations, where, the ACLM-enabled one encountered a threshold of success-ratio drop at 70,000 times of inner delete operations per minute, and the test case should be referred to as covering ultra situations too. Moreover, the higher the computing power of server and client ends became, the less the MIS performance loss from ACLM would be, regarding that the hardware configuration of the test is quite low (the hosting Oracle 9i DBMS was mounted on a Dell PowerEdge2950 PC server -- Xeon E5430 CPU with a 2.66 GHz frequency and 1G memory), all these showed that the performance decline of routine DBMS operation due to ACLM is normally acceptable or even neglectable.

V. FURTHER CONCERNS ABOUT ATTRIBUTES OF LOGS

A. Log Attributes' Minimization

First, relation *Adt_log* cannot be more simple, the reasons are: (1) attributes other than *Audit_ID* are semantic and all indispensable to specify a data manipulation; 2) the auxiliary attribute *Audit_ID* serves as a foreign key in *Chg_log* to correlate together all data attributes being logged there for the same data manipulation. Next, it is easy to verify that *Adt_log* is in the third normal form (3NF), while *Chg_log* is nearly of 3NF except a functional dependency $\langle Lob_value \rightarrow Hash \rangle$. Although $\langle Lob_value \rightarrow Hash \rangle$ brings some redundancy (a list of *Hash* result), it in return offers a well-balanced performance or efficiency in the value comparison of judging if the content of lob type attribute underwent an actual change, regarding that a lob type attribute could have an unlimited variety of content data sequence length, but it can be stood for by its fixed length *Hash* value in comparison computing.

B. Logs Construction's Completeness

The proposed audit & change logs enable restoring any data record's value (if existed) at an arbitrary given time *ts* for data tables under ACLM governing: at first with them we can check if the data record underwent a change manipulation after *ts*; if no, the current status of the data record is the wanted, otherwise we can take steps as described in subsection III.D.1) to restore the data record's value at *ts*.

C. Compatibility with data structure change

First, adding or retiring a non-PK attribute would make no difference on restoring historical snapshots since the name and value of newly added or historical attributes had been recorded in attribute *chgfldname* of *chg_log* if they underwent a value change. Secondly, altering data type of a non-PK attribute would not mislead comprehension of the relevant record content (if any) logged by the change log, since the attribute's previous content was always recorded in a uniform type of character string (via *toString* translation) each time the attribute content underwent a value change. Besides, an alteration of data type without change content is of application-specific disposal which is none of the business of ACLM. Notice: normally retiring a data table field means that the field is no longer maintained but its previous values are still kept there.

Similarly we can elucidate that ACLM still performs its normal functions when data tables undergo a structure alteration in their PK attributes (attributes were added or retired, or their data type were changed) with the complementary PK structure specification from relation PK_spec (see Table II). So, ACLM is compatible with data structure or meta-data change.

VI. HANDLE PRIMARY KEY OF MULTI-ATTRIBUTES

The above ACLM solution has a precondition: all data tables governed must have a unitary-attribute PK. In practices, it is the 1NF [2] instead of a unitary-attribute PK that shall be the minimal requirement for relational schema designs, i.e., there is a PK (probably with multiple attributes) to exclude duplicate rows -- under 1NF at least all attributes together can uniquely fix on an instance of tuple. As to a data table without a unitary-attribute PK we shall introduce an artificial attribute to stand in as a unitary-attribute PK — named the Stand-In Unitary key, in short SIU key. Here, we propose two ways to set up an SIU key:

A. Map Multi-attributes into a unitary one

Let $PK=(F_1, F_2, \dots, F_k)$, which is an ordered tuple of names of all attributes from the PK, where the subscript numbers are defined by the $PK_attribute_seq$ in relation PK_spec regarding the $PK_attribute_name$ (see Table II). If $\forall j (\text{chr}(d) \notin \text{str}(F_j))$ and $\forall j (F_j \leftrightarrow \text{str}(F_j))$ we have

$$(F_1, F_2, \dots, F_k) \leftrightarrow \text{str}(F_1) + \sum_{j=2}^k (\text{chr}(d) + \text{str}(F_j)) = \text{SIU}.$$

Where, $\text{chr}(d)$ denotes the character whose decimal ASCII code is d providing that $\text{chr}(d)$ shall not appear in content of any involved relation attributes on high level MIS applications, “+” is the concatenating operator of character string, $\text{str}(X)$ is the function that converts the value of variable X into a character string, “ \leftrightarrow ” stands for a one-to-one mapping relation. Normally we choose $d=24$ for $\text{chr}(24)$ is a non-printable character for a cancel signal in hardware control. Further, we recommend to turn SIU into a fixed length by *Hash* (SIU) for a better space efficiency and well balanced performance of comparative computation. Often a hash algorithm named *SHA1* is adopted to map different SIU values into distinct strings of 40 hexadecimal characters.

B. Create a DBMS self-maintained field as SIU key

Such built SIU key is normally self-incremental, it can label distinctly data records existed, and grant each newly inserted data record with a unique identity. In semantics, a SIU key is equal to its original PK for all *Update* manipulations except those altering any existent PK value. Whenever a data record underwent first a *Delete* and subsequently an *insert* operation (an equivalence of the *Update* operation altering an existent PK value), it will be assigned a new SIU key value different from its previous ones.

In this sense, any audit record of N type does not link to a historical snapshot of the data record it audited with respect to the PK value of that data record, because the PK value of the newly-inserted data record should not appeared before (the

audit timestamp). As usual, the audit records of N type are used to exclude data records that are inserted after a given historical time in restoring the historical snapshot of a whole data table from its current status. But for restoring a specific data record with a given PK value (each member attribute value is given) at a given time it turns to be rather clumsy: we need to search the change log thoroughly for any *Audit_ID* value that is with each member attribute of the PK in change log records, and each member attribute from these records at least matched one time with their given value, this is because we don't know directly the SIU value at that given moment. But, if ACLM is merely oriented to the time property of data maintenance and historical snapshots of data table, this SIU key approach is to some extent simple and feasible.

VII. DISCUSSION AND CONCLUSION

The ACLM function is based on two basic conditions: 1) accurately logging all timestamps of maintenance operations that should be governed by ACLM and values of each involved attribute of the object data record just before each operation of change data; 2) being able to correlate together the values of all member attributes in the data tuple of snapshot. As illustrated in section III, the design and application of audit and change logs themselves have directly satisfy the second condition, and such a condition would be hold met ever since it was satisfied since both the audit and change logs permit only Insert-SQL maintenance manipulations; while the first condition is met by enabling DBMS trigger mechanism [5] that surely captures any event of maintenance operation.

Under ACLM we can flexibly program SQL scripts to recall any historical snapshot without difficulty. Compared with those powerful but extravagant tDBMS products, our ACLM-based solution is economical and exercisable (of DIY type) for MIS. This is because a) implementing ACLM is plain: three specific relational tables for logs and PK specification, a short script additive to usual RDBMS stored procedures of business logic, and a piece of SQL script (alone or embedded in existent SQL scripts) added in triggers of each data table audited, and b) the merits from ACLM: being compatible with changes to meta-data, and programmers can code in their usual way, e.g., implementing data maintenance via RDBMS stored procedures.

In fact, our ACLM approach does reflect a reality that actual tDBMS implementation is a workable evolution of RDBMS application rather than an innovation of nowadays RDBMS. The most valuable point thereby is, ACLM can be plainly deployed on current prevailing RDBMSs with less interference in routine MIS program practices, normal MIS users except who conduct audit would feel they are working with a familiar RDBMS for the current version data access as usual. We can conclude that the ACLM approach is a good option to replace current bitemporal database solutions for uni-temporal MIS applications in order to avoid unexpected or unnecessary cost and complexity.

ACKNOWLEDGMENT

The authors thank all members of their team and correlative colleagues for applying and validating ACLM in several e-

government MIS projects with requirements of uni-temporal database applications.

REFERENCES

- [1] C.J. Date, H. Darwen, and N. Lorentzos, *Temporal Data & the Relational Model*, 1st ed. San Francisco, USA: Morgan Kaufmann, 2002.
- [2] C. J. Date, "What first normal form really means," pp. 127–128. <http://www.dbdebunk.com/page/page/629796.htm>
- [3] D. Eastlake 3rd and P. Jones, *US Secure Hash Algorithm 1 (SHA1)*, RFC Editor, 2001. <http://tools.ietf.org/html/rfc3174>
- [4] N. Freed and N. Borenstein, *Multipurpose Internet Mail Extensions (MIME) - Part One: Format of Internet Message Bodies*, RFC 2045, Nov. 1996.
- [5] http://en.wikipedia.org/wiki/Database_trigger
- [6] C. S. Jensen, "Introduction to temporal database research," Aalborg University, Denmark, Tech. Rep: No.1, 2000, pp. 1-27.
- [7] J. Celko, *Joe Celko's SQL for Smarties: Advanced SQL Programming*, 4th ed. San Francisco, USA: Morgan Kaufmann, 2010. ISBN: 978-0-12-382022-8
- [8] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," in *Communications of the ACM*, vol 21(2), 1978, pp: 120–126. doi:10.1145/359340.359342
- [9] R.T. Snodgrass, *Developing Time-Oriented Database Applications in SQL*. San Francisco, USA: Morgan Kaufmann, 1999.
- [10] H. Guo, Y. Tang, X. Yang, X. Ye, "Improvement and extension to ATSQL2," in *Temporal Information Processing Technology and Its Application*, Y. Tang et al, Eds. Tsinghua University Press, Beijing and Springer-Verlag Berlin Heidelberg, 2010, pp. 245-259.
- [11] M. Kaufmann, P. M. Fischer, N. May, A. Tonder, and D. Kossmann, "TPC-BiH: a benchmark for bi-temporal databases," In *TPCTC*, 2013.
- [12] *TimeDB—A Temporal Relational DBMS*. <http://www.timeconsult.com/software/software.htm>

The Impact of Cognitive Tools on the Development of the Inquiry Skills of High School Students in Physics

Mohamed I. Mustafa,
Faculty of Education
University of Ottawa
Ottawa, Canada

Louis Trudel
Faculty of Education
University of Ottawa
Ottawa, Canada

Abstract—the purpose of the study was to compare the effectiveness of two teaching strategies that utilize two different cognitive tools on the development of students' inquiry skills in mechanics. The strategies were used to help students formulate Newton's 2nd law of motion. Two cognitive tools had been used: a computer simulation and manipulations of concrete objects in physics laboratory. A quasi-experimental method that employed the 2 Cognitive Tools × 2 Time of learning split-plot factorial design was applied in the study. The sample consisted of 54 Grade 11 students from two physics classes of the university preparation section in a high school of the province of Ontario (Canada). One class was assigned to interactive computer simulations (treatment) and the other to concrete objects in physics laboratory (control). Both tools were embedded in the general framework of the guided-inquiry cycle approach. The results showed that the interaction effect of the Cognitive Tools × Time of learning was not statistically significant. However, the results also showed a significant effect on the development of students' inquiry skills regardless of the type of cognitive tool they had used. Although the findings suggested that these two strategies are effective in developing students' inquiry skills in mechanics, students in the computer simulation group had shown larger gain in their inquiry skills test than their counterparts in the laboratory group.

Keywords—inquiry skills; teaching strategy; cognitive tools; high school physics; simulation; science laboratory

I. INTRODUCTION

A number of difficulties undermines learning science through inquiry, and prevents students from successful engagement in a meaningful investigation. One of the major barriers in inquiry-based learning is the students' lack of adequate level of inquiry skills [1,2,3], such as posing questions and identifying variables, formulating hypothesis, designing and conducting investigations, collecting and analyzing data. Due to an inadequate level of inquiry skills students often learn science through direct observations and problem solving tasks without spending efforts to experience natural phenomena, or construct abilities and knowledge to understand how natural world works [4,5]. Although most relevant research studies provided little guidance about how students acquire and develop these skills over time, some show that factors like students' self-efficacy, effective scaffolding, and collaborative learning environments, and utilization of

cognitive tools may have an impact on students' abilities and skills to progress through inquiry tasks [2,5].

Various types of cognitive tools have been identified and explored in terms of their potential usefulness to help students in constructing knowledge, and developing inquiry skills [1,6,7]. Great promises have been given to the visualization of natural phenomena via interactive computer simulation [8], and the manipulation of concrete objects in science laboratory [1].

On the one hand, interactive computer simulations provide support for nearly every aspect of scientific inquiry [5]. New computer simulations have evolved now to the point they can contribute to the development of students' learning through inquiry, and allow teachers to meet students' needs more fully [10]. Learners use various types of new simulations to collect, organize and analyze data, transform data into variety of representations, create virtual situations in which they can test hypothesis, and even synthesize and execute their own models [11, 5].

On the other hand, a body of research studies has considered inquiry-based laboratory as one of the ways students can understand the natural world of science and develop their inquiry skills [1,12]. When a laboratory investigation is constructed in a problem solving context, students will be required to plan a course of action, carry out the activity and collect necessary data, organize and interpret the data, and reach a conclusion. This type of laboratory investigation encompasses both the ways in which understanding is generated within the natural sciences and the approaches to solving problems, as well as attaining inquiry skills and understanding the process of scientific protocols.

Despite the potential advantages of computer simulations and concrete phenomena in enhancing inquiry-based learning, it is important to note that their uses have certain potential drawbacks as well, and therefore, must be taken into consideration. For example, computer-simulated experimentation software dictates the direction of inquiry by predefining variables. Even though simulations can be interactive, students cannot test alternative models or novel variables that are not programmed into the system, so that the opportunity to identify variables on their own is not available [13, 14]. Also, laboratory work has been criticized and claimed that it is unproductive and confusing. Very often students are

involved in technical activities (such as assembling the experimental setup) and few opportunities are given to students to present their interpretations and beliefs about natural phenomena [15]. These limitations could prevent students from having an authentic opportunity to develop adequate level of inquiry skills and consequently undermine their efforts to construct new knowledge and engage in inquiry-based learning [9].

Therefore, this study attempts to compare the impacts of two teaching strategies on the development of students' inquiry skills in the domain of physics. One strategy will integrate computer simulation while the other strategy will integrate concrete objects as cognitive tools. In this regard, the main objective is to find out which learning tool is more suitable to help students develop their inquiry skills and in which learning circumstances. Accordingly, the following research questions are designed to guide the study:

- 1) *What are the possible impacts of the visualization of Newton's second law of motion via interactive computer simulation on the development of students' inquiry skills?*
- 2) *What are the possible impacts of the manipulation of concrete objects to perform Newton's second law experiment on the development of students' inquiry skills?*
- 3) *How are the effects of the utilization of computer simulation and the manipulation of concrete objects compared?*

II. THEORETICAL FRAMEWORK

The theoretical framework of this study is based on the social constructivist approach of learning. Piaget [16] argued that learners are constantly constructing and reconstructing knowledge in their efforts to maintain a coherent system of interpretation [17]. Also, they are capable of performing at higher intellectual levels when they are guided by an experienced person and asked to work in collaborative situations [18]. We argue that learners, in a guided-inquiry environment, need to acquire an appropriate level of inquiry skills in order to accomplish inquiry tasks and phases successfully. One of the ways to achieve that relies on the opportunities given to students to learn with cognitive tools such as computer simulation or concrete objects in science laboratory.

When cognitive tools present natural phenomena as real-life problems, they provide feedback to students on their attempts to solve these problems. In the context of learning, feedback and reflection weigh students' performances while obtaining data, as well as their predictions about the outcomes of the investigation [19,6]. The argument is that, while students are investigating natural phenomena, they receive feedback not only from an experienced person (like teacher, or a capable peer), but also from their performances associated with cognitive tools. Accordingly, they may feel a need to adjust their performance in response to it. This is expected to raise students' awareness of their inquiry skills deficiencies and encourage them to think how to correct them.

Therefore, students tend to engage in inquiry investigations repeatedly to practice their inquiry skills and consolidate them in each engagement [7].

Furthermore, teacher guidance and peer interaction will present additional experiences that would lead students to complete their inquiry tasks and phases more efficiently.

Consequently, the integration of cognitive tools in guided-inquiry approach is expected to develop students' inquiry skills. Therefore, the main purpose of the study is to compare the effectiveness of two teaching strategies embedding two cognitive tools: visualization of natural phenomena via interactive computer simulation and manipulation of concrete objects in a physics laboratory on the development of students' inquiry skills in the context of Newton's 2nd law of motion.

III. TEACHING STRATEGIES

A. Teaching and Learning Objectives

The overall learning objectives are set for students to "demonstrate scientific investigation skills (related to both inquiry and research) in the four areas of skills: initiating and planning, performing and recording, analyzing and interpreting" [20, p.182]. Throughout the learning activities, students are expected to develop inquiry skills to be able to [20]:

- 1) *Identify variables such as forces, mass, time, etc.; and*
- 2) *Design and conduct an inquiry to analyze the effect of forces acting on objects in one dimension, using vector diagrams, free-body diagrams, and Newton's laws. During the inquiry investigation, students are expected to use their assigned cognitive tool (concrete objects or computer simulation) to analyze, in quantitative terms, the relationships between two variables when a third is kept constant (e.g., relationships between acceleration and applied force when the mass is kept constant). So in the end, students are expected to formulate Newton's Second Law in their own.*

The teaching objectives, therefore, will primarily be looking at ways to foster the development of students' inquiry skills and guide them to accomplish their inquiry tasks successfully.

B. The Cognitive Tools

The cognitive tools and materials are basically selected to support the development of students' inquiry skills.

- **Computer simulation:** The computer simulations are selected from the ones available on the Physics Education Technology (PhET) website (<http://phet.colorado.edu/index.php>). They are developed and tested by Physics Education Technology (PhET) project. Three interactive simulations have been selected from PhET website: Force in One Dimension [21], The Ramp [22], and Moving Man simulations [23]. As an example, the window of the first simulation is shown in figure 1.

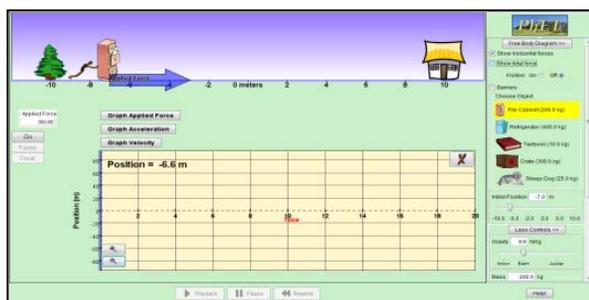


Fig. 1. Main window of Force in one dimension

- Concrete objects: The concrete objects are laboratory apparatus that are selected from the many physics activities available in grade 11 textbooks [24]. The concrete objects include dynamic cart, set of known masses, ticker-tape timer, banked ramp, and spring balance (fig. 2).



Fig. 2. A set of concrete objects used for Newton's second law activities

C. The Implementation of teaching strategies

Both strategies used the 5Es learning cycle- engage, explore, explain, elaborate, and evaluate- [25] in a similar way except in the implementation of the assigned cognitive tool. Before engaging students in the 5Es cycle, the teacher took the first day to organize students in small groups and train them with their assigned cognitive tool. The training session is designed as a learning activity, so students can monitor their progression and practice their abilities within a learning context. During the phases of 5Es cycle, both classes used identical instructional materials and inquiry tasks. Those tasks consist for each phase of the 5Es cycle:

- Engage: Reading a story-telling and respond to the associated questions. Posing questions to be investigated in the next phase.
- Explore: Designing and conducting an inquiry into the relationship between the acceleration of an object and its net force and mass.
- Explain: Analyzing the relationships between acceleration and applied forces.
- Elaborate: Conduct an inquiry that applies Newton's laws to analyze, in qualitative terms, the forces acting on an object, and use free-body diagrams to determine the net force and the acceleration of the object.

- Evaluate: The questions presented at the end of phase of the 5Es cycle are designed to evaluate students' progress.

The main difference between the simulation and the physics laboratory were mostly to be found firstly in the exploration phase of the 5Es cycle and secondly in the elaboration phase.

Firstly, in the exploration phase, the simulation provides controls that allow students to create virtual objects, change variables and control others and study the trend of the resulting type of motion in forms of graphical representations (fig. 3). In the other hand, in the physics laboratory, students were given concrete objects to manipulate and it allows students to observe the motion of real objects (Fig. 4).

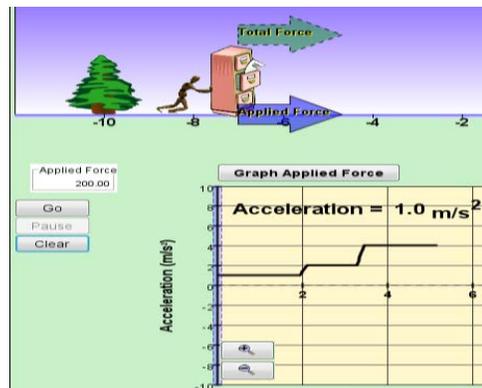


Fig. 3. Force in 1D simulation

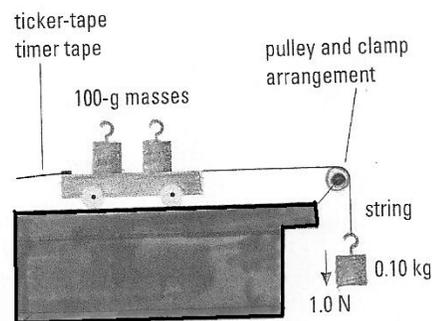


Fig. 4. An experimental setup used to simulate the motion of a cart in a horizontal plane

Secondly, in the elaboration phase, the simulation allows students to test the motion of an object in a frictionless inclined plane and provide them with free body diagrams to study various aspects of motion. In the other hand, the physics laboratory allows students to create different setups to test the motion of an object in a rough inclined plane, and as a consequence, they could observe directly the motion of real objects in an inclined plane.

With respect to the organisation and unfolding of the activities, the two teaching strategies differed notably. In the simulation strategy, students worked in small groups. They used *Force in one dimension* simulation to test their predictions by making cause and effect relations between force, mass, and acceleration. The simulation contains controls that allow

students to create objects with which they can associate net force, mass, and acceleration as variables. Next, students presented their own explanations that made sense of their observational data, and establish relationships between net force, mass, and acceleration to state Newton's 2nd law of motion in words and symbols. The teacher had guided the students through their data collection part and provided hints to keep them focused and on task. Small-group and whole-class discussions followed that have focused students on the conclusions they made to reach a consensus about the relationships between net force, acceleration, and mass.

In the laboratory strategy, students worked in 4-member groups. They designed an experimental setup (fig. 4) to analyze the motion of a cart and establish the relationship between the net force acted on the cart, its acceleration, and mass. The teacher guided students through the same scaffolded learning events that the teacher would have used with students in the simulation group.

IV. METHODOLOGY

The ideal method would be to compare, in a quasi-experimental design, the development of students' inquiry skills in two learning environments: one with an interactive computer-based simulation, and the other with concrete objects and materials in a physics laboratory. Accordingly, the following hypotheses were postulated and computed at the 0.05 level of significance:

Hypothesis 1: students taught via inquiry-based computer simulation will significantly develop their inquiry skills better than students taught via inquiry-based laboratory.

Hypothesis 2: students taught via inquiry-based laboratory will significantly develop their inquiry skills better than students taught via inquiry-based computer simulation.

Hypothesis 3: students taught via inquiry-based laboratory and inquiry-based computer simulation will significantly develop their inquiry skills.

With respect to the sample, 54 students have been conveniently selected from grade 11 – university preparation section from a high school in Ottawa. Without disturbing the school academic plan, this study was conducted during the first semester of school academic year 2010 – 2011. The content of the study was the regular physics course for grade 11 – university preparation section, developed by Ontario Ministry of Education, and facilitated by a high school physics teacher. The participating students came from two naturally formed classrooms (each of approximately 27 students). Both groups conducted "Force in one dimension", one of the experiments in Ontario Curriculum – Physics [20]. One group was randomly assigned to a computer-based environment, and the other one to a physics laboratory environment.

With respect to the research design, the independent variable was the learning tool factor with two levels: the simulation and the laboratory. The dependent variable consists of a test of integrated process skills (TIPS II). According to their authors, the test exhibits a total test reliability of 0.86 [26]. A two-way split-plot design was performed to examine if there was a significant statistical difference between the means of

students' scores in the pre- and post-TIPS (II) [26]. The number of questions answered correctly in the TIPS (II) served as the dependent variable in a 2 Learning Tool \times 2 Time split-plot Analysis of Variance (ANOVA).

The 2 levels of the learning Tool factor (between-subjects) consist of, as explained before, the interactive computer simulation and the physics laboratory; and the Time of test factor (within-subjects) will have two levels: pre- and post-Newton's second law activities. Additionally, another variable has been added to perform a one-way ANOVA comparing MDT scores, Mechanics Diagnostic Test [27], of participants in the two Learning Tools conditions to ensure no pre-existing differences in conceptual understanding, and therefore, was considered as a covariate factor.

V. RESULTS

A. Descriptive

The descriptive results of the MDT showed that students from both groups, in average, have similar levels of conceptual understanding in mechanics (Table 1).

TABLE I. MECHANICS DIAGNOSTIC TEST RESULTS BY GROUPS

Group	N	Mean	Std. Deviation
Laboratory	24	11.50	2.844
Interactive simulation	25	11.12	2.166
Total	49	11.31	2.502

The figure 5 describes two plots representing the means of students' scores in the pre- and post-skills test in laboratory group and simulation group. The graph shows an increase in the means in both groups. However, it appears that the line graph of the simulation group is steeper than that of the laboratory group. It seems that students in the simulation group obtained a slightly higher gain in TIPS scores than students in the laboratory group, despite the fact that the graph of the simulation group was lower than that of the laboratory group.

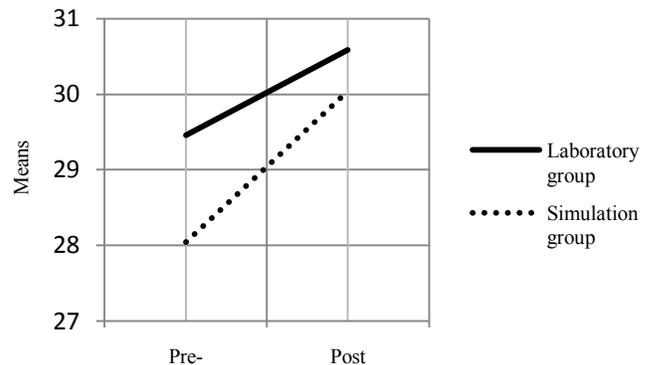


Fig. 5. Inquiry Skills Test (TIPS II)

B. Inferential statistics

The results of one-way ANOVA (Table. 2) supported the initial screening with the physics teacher and showed that students across the two groups were not significantly different

in their level of conceptual understanding ($F(1, 47) = 0.278$). Therefore, the students' MDT scores did not threaten the relationship between learning tool (independent factor) and students' scores of TIPS (II) (dependent factor).

TABLE II. TABLE 2 ONE-WAY ANOVA TEST

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	1.768	1	1.768	0.278*	0.600
Within Groups	298.640	47	6.354		
Total	300.408	48			

* $p < .05$

A 2 Learning Tool \times 2 Time split-plot repeated measures ANOVA analysis (Table 3) showed that the interaction effect was not significant ($F(1,47) = 1.349$). This suggested that the effect of learning activities over time on the development of student inquiry skills did not depend on the type of learning tools used during the activities. Hence, hypotheses 1 and 2 were rejected.

TABLE III. SPLIT-PLOT LEARNING TOOL \times TIME ANALYSIS OF VARIANCES

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Between subjects					
Intercept	85424.86	1	85424.86	3300.7	0.000
Learning Tool	23.56	1	23.56	0.910*	0.345
Error	1216.40	47	25.88		
Within subjects					
Time (pre-and post)	59.79	1	59.79	17.207**	0.000
Learning Tool \times Time	4.69	1	4.69	1.349***	0.251
Error (time)	163.31	47	3.48		

* $p < .05$, ** $p < .05$, *** $p < .05$

The repeated measures test (Table 3) also showed that there was no significant main effect for the Learning Tool ($F(1,47) = 0.910$). This suggests that the type of learning tool did not significantly influence the development of student's inquiry skills. However, the repeated measures test revealed a significant main effect for the within-subjects factor ($F(1,47) = 17.207$). This suggested that there was statistically significant difference between students' scores in the pre-TIPS (II) and the post TIPS (II), regardless of what type of learning tool they had used in Newton's second law activities. Therefore, hypothesis 3 was supported.

VI. DISCUSSION

The results of this research showed that in a guided inquiry-based learning environment, visualization via interactive computer simulations or manipulation of concrete objects in a physics lab can both be used as a learning tool to develop students' inquiry skills in mechanics. The positive effects of these intellectual learning tools (mind-tools) in supporting the development of students' inquiry skills are consistent with the cognitive tools that support structures that aim to compensate for students' knowledge or inquiry skills deficiencies [6,4].

Despite the fact that the two learning tools did not show a statistically significant difference, subjective observations of the classroom environment made by the researcher may give some insights on the effectiveness of the computer simulation and the manipulation of concrete objects. For example in their attempts to establish relationships between net force, mass, and acceleration, the students in computer simulation group not only had the chance to manipulate the values of net force and observe the resulting type of motion several times, but they also visualized the effect of specific components of the net force on the acceleration of an object, which they would not have identified if they were to manipulate concrete objects. In contrast, students in the physics laboratory were given opportunities to manipulate concrete objects, which provided real experiences with natural phenomena (motion of a dynamic cart) that allowed students to study what influenced the motion (acceleration) of a dynamic cart.

However, prior research has shown that the development of inquiry skills is typically difficult to achieve in a short time-effect of inquiry instruction, and it is always challenged by the abilities of students to self-direct (self-regulate) their inquiry investigation [28]. This view was noticed during this study as well. In this regard, different challenges have been recorded from the researcher's diary, and from students' booklets including observational errors, instrumental errors, long time on task, and lack of operational skills.

The major part of the above stated challenges, however, was compensated by a successful implementation of other components of the teaching strategy such as quality of Students' booklets, collaboration setting, and the physics teacher scaffolding.

VII. CONCLUSION

The results of the study showed that there were no significant interaction effects between the types of learning tool used during the activities and the development of the students' inquiry skills measured by the pre-TIPS (II) and the post TIPS (II). Students from both conditions significantly developed their inquiry skills over the time provided for the learning activities. From this, the students' level of conceptual understanding in mechanics did not influence their efforts to develop their inquiry skills.

The answer to the first research question concluded that computer simulations hold promises to assist students in identifying and establishing relationships between variables. Working with computer simulations allows students to work independently and save valuable class time for other learning activities. It should be noted, however, that the current simulations may not promote real experiences with natural phenomena.

The second research question concluded that, with careful guidance students can benefit from manipulation of concrete objects in a science laboratory. Students can challenge their ideas against natural phenomena and observe how it actually works, practice different components of inquiry skills, and realize how scientists work in real life. A successful laboratory experiment however, requires good preparation and longer time to perform, which may not be possible at all times.

All in all, if technology is used in balance with real experiences and is placed in its proper context, it can enrich the classroom by providing new and contrasting contexts in which to understand experiences [13].

Despite that, the study has certain limitations that should be taken into consideration when interpreting the results. The limitations pertain to the following:

- The study did not take students' diversity into account, which made the results less conclusive from the point of view of gender, physical abilities, and cultural background.
- The small sample size limits the study's generalizability therefore, the results are subjected to the conditions and learning circumstances that occurred during this study.
- The teaching strategy embedded two learning tools to tackle the five inquiry skills. The results were not specific enough to identify which particular inquiry skill has led the overall improvement.

VIII. FUTURE WORK

The results from this study present further research opportunities including how the development of inquiry skills might support students' efforts to understand physics concepts and principles; what aspects of teaching strategies afford opportunities for the development of inquiry skills; what kind assessment tools are required on inquiry investigation; and how new teaching strategies imparting learning science as inquiry. Furthermore, the study present research opportunities regarding the methods that should be used to prepare teachers to teach science as inquiry, and what skills teachers should acquire to be able to guide students through inquiry learning.

REFERENCES

- [1] A. Hofstein, "Providing high school chemistry students with opportunities to develop learning skills in an inquiry type laboratory: a case study", *International Journal of Science Education*, vol. 26, no 1, pp. 47-62, 2004
- [2] D. Ketelhut, "The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment", *Journal of Science Education and Technology*, vol. 16, no 1, pp. 99-111, 2007
- [3] D. Kuhn, and M. Pease, "What needs to develop in the development of inquiry skills?", *Cognition and Instruction*, vol. 26, pp. 512-559, 2008.
- [4] S. Manlove, A. Lazonder and T. de Jong, "Regulative support for collaborative scientific inquiry learning", *Journal of Computer Assisted Learning*, vol. 22, pp. 87-98, 2006.
- [5] S. Khan, "New Pedagogies on Teaching Science with Computer Simulation", *Journal of Science Education and Technology*, vol. 20, no 3, pp.215-232, 2011.
- [6] D. Jonassen, *Computer in the Classroom: Mindtools for Critical Thinking*. Englewood Cliffs, NJ: Merrill, 1996
- [7] D. Kuhn, and D. Dean, "Is developing scientific thinking all about learning to control variables?", *Psychological Science*, vol. 16, no 11, pp. 866-870, 2005.
- [8] S. Abdullah, and M. Abbas, "The effects of inquiry-based computer simulation with cooperative learning on scientific thinking and conceptual understanding", *Malaysian Online Journal of Instructional Technology (MOJIT)*, vol. 3, no 2, pp. 1-16, 2006.
- [9] P. Brown, S. Abell, A. Demir, and F. Schmidt, "College Science Teachers' Views of Classroom Inquiry", *Science Education*, vol. 90, no 5, pp. 784-802, September (2006).
- [10] S. Foti, and G. Ring, G. (2008). "Using a simulation-based learning environment to enhance learning and instruction in a middle school science classroom", *Journal of Computers in Mathematics and Science Teaching*, vol. 27, no 1, pp. 103-120, 2008.
- [11] P. Kubicek, "Inquiry-based learning, the natural of science, and computer technology: New possibilities in science education", *Canadian Journal of Learning and Technology*, vol. 31, no 1, Winter (2005).
- [12] M. Kipnis, and A. Hofstein, "The inquiry laboratory as a source for development of metacognitive skills", *International Journal of Science and Mathematics Education*, vol. 6, pp. 601-627, 2007.
- [13] T. Machet, D. Lowe, and C. Gütl, "On the potential for using immersive virtual environments to support laboratory experiment contextualisation", *European Journal of Engineering Education*, vol. 37, no 6, pp.527-540, December (2012).
- [14] B. Dalgarno, and M. Lee, "What are the learning affordances of 3-D virtual environments?", *British Journal of Educational technology*, vol. 41, no 1, pp. 10-32, January (2010).
- [15] P. Knipp, "Broadcasting a lab measurement over existing conductor networks", *The Physics Teacher*, vol. 47, no 7, pp.471-473, 2009.
- [16] J. Piaget, "Cognitive development in children: Development and learning", *Journal of Research in Science Teaching*, vol. 2, pp. 176-186, 1964.
- [17] B. Davis, D. Sumara, and R. Luce-Kapler, *Engaging minds: Changing teaching in complex times*, 2nd Edition, NY: Routledge, 2008.
- [18] L. Vygotsky, *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press, Wadsworth Publication Company, 1978.
- [19] J. Dewey, *Experience and Education*. Kappa, Indiana: Delta Pi, 1938.
- [20] Ministry of Education, *The Ontario Curriculum Grades 11 and 12 (Revised)*. Retrieved from Ontario Ministry of Education's website at www.edu.gov.on.ca, 2009
- [21] PhET University of Colorado (2009). *Forces in 1 Dimension*. Interactive Simulations. Retrieved August 15, 2010, from <http://phet.colorado.edu>.
- [22] PhET University of Colorado (2009). *The Ramp*. Interactive Simulations. Retrieved August 15, 2010, from <http://phet.colorado.edu>.
- [23] PhET University of Colorado (2010). *The Moving Man*. Interactive Simulations. Retrieved August 15, 2010, from <http://phet.colorado.edu>.
- [24] M. DiGiuseppe, "Nelson Physics 11: University Preparation". Scarborough, ON: Nelson Education, 2012.
- [25] J. Renner, "Investigation in natural science: Biology teachers' guide". Report: ED260927, 177 pp., 1985
- [26] J.C. Burns, J. R. Okey, and K.C. Wise, K. C., "Development of an integrated process skills test: TIPS (II)", *Journal of Research in Science Teaching*, vol. 22, no 2, pp.169-177, 1985.
- [27] I. Halloun, "Common sense concepts about motion", *American Journal of Physics*, vol. 53, no 11, pp. 1056-1065, 1985.
- [28] E. Scanlon, S. Anastopoulou, L. Kerawalla, and P. Mulholland, "How technology resources can be used to represent personal inquiry and support students' understanding of it across contexts", vol. 27, no 6, pp.516-529, 2011.

A System for Multimodal Context-Awareness

Georgios Galatas
Computer Sci. and Eng. Dept.
University of Texas at Arlington
Arlington, TX, USA IIT, NCSR Demokritos
Athens, Greece

Fillia Makedon
Computer Sci. and Eng. Dept.
University of Texas at Arlington
Arlington, TX,
USA

Abstract—in this paper we present the improvement of our novel localization system, by introducing radio-frequency identification (RFID) which adds person identification capabilities and increases multi-person localization robustness. Our system aims at achieving multi-modal context-awareness in an assistive, ambient intelligence environment. The unintrusive devices used are RFID and 3-D audio-visual information from 2 Kinect sensors deployed at various locations of a simulated apartment to continuously track and identify its occupants, thus enabling activity monitoring. More specifically, we use skeletal tracking conducted on the depth images and sound source localization conducted on the audio signals captured by the Kinect sensors to accurately localize and track multiple people. RFID information is used mainly for identification purposes but also for rough location estimation, enabling mapping of the location information from the Kinect sensors to the identification events of the RFID. Our system was evaluated in a real world scenario and attained promising results exhibiting high accuracy, therefore showing the great prospect of using the RFID and Kinect sensors jointly to solve the simultaneous identification and localization problem.

Keywords—Multimodal; Context-awareness; Microsoft Kinect; RFID; Localization; Identification

I. INTRODUCTION

An assistive ambient intelligence environment is a smart space that aids the inhabitants with its embedded technology. The proliferation of ambient intelligent environments has triggered research related to applications, such as monitoring Assistive Daily Living (ADL), fall detection, risk prevention and surveillance [1, 2]. For achieving these goals, activity recognition performed in a natural and unintrusive way is of utmost importance. The most fundamental step towards activity monitoring and ultimately context-awareness is successful multi-person identification and localization. By utilizing the location of the person in a domestic setting, related activities can be derived. Accurate person localization plays an essential role in all the aforementioned applications and has been dealt with using many different approaches. Nevertheless, when used domestically, most current implementations can be considered as invasive. Our novel system uses information from multiple sensors in order to ensure reliable and unintrusive localization of the inhabitants.

A. Localization:

Applications that rely on localization such as surveillance and monitoring of ADL commonly use video cameras as an affordable and abundant source of information. Many

approaches based on either a single camera or multiple cameras have been proposed in the literature.

In single camera setups, discriminative appearance affinity models [3] and level-set segmentation [4] have been used for tracking, while other approaches based on tracking-by-detection exist [5, 6]. In multi-camera setups, stereo-vision is employed in order to introduce depth perception. In [7], color histograms of the person-shaped blobs are used to disambiguate between people, when they are very close to each other. The system tracks multiple people standing, walking, sitting, entering and leaving in real-time. In [8] two techniques were used to determine the location of a person in 3-D space. These were 1) best-hypothesis heuristic tracking and 2) probabilistic multi-hypothesis tracking to derive the 3-D location of people. The results show similar tracking performance for both approaches. However, the simplistic probabilistic approach produces more false alarms, which may be improved by using a sophisticated probabilistic model.

Solving the problem using only cameras is very challenging for a large space with many people. The reason is that localization requires wide coverage to capture and map the respective locations of many people simultaneously, but identification requires zooming into a person's face. In surveillance applications, cameras are typically mounted on tall poles and configured such that they could provide maximum coverage. Nevertheless, video feeds from such settings may not be sufficient to provide accurate information about a person's face or other biometric features. In addition, the segmentation and tracking problems can be very challenging, thus hindering the system's reliability in a camera-only setup. Furthermore, despite the fact that the use of cameras and computer vision techniques are very promising, extensive use of video cameras in a domestic setting can be considered a violation of privacy [9]. Therefore, our main focus is to achieve the same goal of identifying and localizing multiple people in an assistive environment in a less intrusive manner.

B. Identification:

RFID (Radio-frequency Identification) systems are frequently being used to track medicine and patients in large hospitals in order to verify the correct medicine reaches the correct patients [10]. RFID sensors have become very popular, as they are cheap, easy to use and provide accurate identification information wirelessly [11]. Although RFID is very effective in identifying objects, it may not be as effective in surveillance applications, since people are required to wear an RFID tag so that the events related to the tag are detected. As a result, such systems may not be able to detect intruders or

This material is based upon work supported by the National Science Foundation under Grants No. NSF-CNS 1035913, NSF-CNS 0923494.

anyone not wearing a tag. However, it constitutes a viable solution for recognizing activities in a smart environment, since its inhabitants can very easily carry a passive RFID tag with them.

Multimodal person identification has become a significant area of research in recent pervasive assistive applications. Some of these applications use existing biometric identification methods, such as face recognition and speaker identification [9, 12, 13] to identify multiple people in smart environments [14]. Nevertheless, these approaches do not convey the location information of the person.

C. Simultaneous Identification and Localization:

Locating multiple users simultaneously while identifying each one is considered to be the first step to create a context-aware application, such as activity and human behavior recognition. RFID technology has also been used to solve the problem of simultaneous identification and localization. Although radio signal propagation suffers from various problems, such as multipath, line of sight path, diffraction or reflection etc. even in an indoor environment [15], several indoor-based localization algorithms have been proposed in the literature, which, according to [16] can be classified into three categories: 1) distance estimation, 2) scene analysis and 3) proximity. Among them, distance estimation algorithms use different range measurement techniques, such as Received Signal Strength, Time of Arrival, Time Difference of Arrival, Received Signal Phase etc. and apply triangulation to estimate a location. On the other hand, the scene analysis approaches first measure fingerprints of an environment and then, try to match the target's range measurements with the appropriate set of fingerprints for estimating the location. Finally, the proximity-based algorithms determine a target's location by mapping it to the location of an antenna that receives the strongest signal.

Overall, RFID technology possesses a promising solution to identify and localize multiple objects with attached RFID tags. Existing well-known systems, such as LANDMARC [17] use active RFID tags and exploit the signal strength property to correctly localize an object. Passive RFID tags have also been used in the past to identify and locate multiple objects. In [18], the authors have utilized the percentage of tag counts at different power attenuation levels in order to approximate the distance between a reader and a tagged object. Another, indirect way of deriving the location information of an object is to record the location of the reader as the location of an object. But, the location accuracy and precision of such a system heavily depends on the level of deployment of readers and antennas in the space [19, 20].

However, RFID still lacks sufficient localization accuracy especially for the minimal number of deployed antennas and tags in a domestic environment. Simply using RFID to obtain the location of an object can lead to many false readings, e.g., an RFID antenna may miss a tag depending on the tag's position and the antenna's orientation.

In an attempt to improve accuracy, multi-modal person localization has become a significant research area in recent applications. Thus, for a very dynamic environment,

information collected from multiple sources, such as video cameras, microphone arrays, sensors etc. are all combined together such that the system can achieve better identification and localization accuracy. Techniques, such as Hidden Markov Models, K-nearest neighbors etc. can be applied to captured audio-visual signals to extract higher-level semantic information, such as identification and location in real time. A system that combines face and audio based identification along with motion detection, person tracking and audio based localization has been proposed in the literature [21]. Such a system applies state-of-the-art methods to process results from each individual modality and uses particle filtering to fuse both modalities for providing robust identification and localization.

Methods that combine localization using cameras with identification using wearable sensors or accelerometers are also proposed in the literature. Since most of the recent mobile phones contain accelerometers and magnetometers attached to them, mobile phones are considered to be very convenient and fulfill all of the above requirements. In [1] the authors combined an existing CCTV based system with sensors (accelerometers and magnetometers) embedded to a person's mobile phone as a solution. According to this method, the camera captures the location of each person, which is transmitted wirelessly to the mobile phone carried by the respective person. After receiving the location information, the mobile phone resolves the most probable location by matching them with the measurements from its own sensors. The identification process is very easy in this case, as each person is labeled with his/her mobile phone's unique ID.

The deployment of wireless sensor network (WSN) is another common approach nowadays to monitor and localize persons in assistive environments [22, 23]. RFID systems and WSNs can be combined together not only for identifying and localizing objects, but also for real-time monitoring [24]. To identify and localize in open areas, researchers of [2] derived a calibration method for a joint RFID-camera system based on the area of overlap between the field of view (FoV) of a camera and the field of sense (FoS) of RFID sensors.

In our approach we build in prior work [29] utilizing the identification capabilities of RFID and combining them with precise 3D tracking from the Kinect to create an accurate identification and localization solution. The latter is an active sensor, able to accurately measure the position of the person in the 3-D space. Skeletal tracking is carried out using the Kinect sensor's 3D depth images and sound source localization is conducted utilizing microphone arrays of 2 such sensors, to deduce accurate location information. At the same time, the video information is not captured, making this approach less intrusive than using video cameras. RFID is used mainly for discerning between users and also for providing a rough estimate of their location utilizing the RSSI. Our goal is to map the location of multiple people in an ambient intelligence environment at a detailed level that will allow inference of conducted activities (figure 1).

In the following sections we will present the architecture and operation of our system for person identification and localization, the experimental setup and finally our concluding remarks.

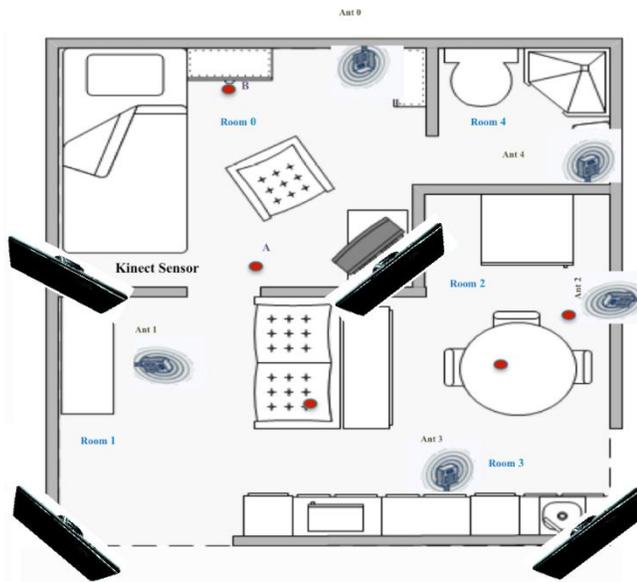


Fig. 1. Example apartment layout with RFID antennas and Kinect sensors deployment.

II. SYSTEM OVERVIEW

A. Hardware

The Microsoft Kinect (figure 2) is a novel device mainly used for gesture recognition. It is based on the PrimeSensor design [25] and it incorporates a color camera, a depth sensor and a microphone array. Depth images are acquired using the structured light technique. According to this method, a laser beam passes through a grating, and is split into different beams. The beams are then reflected from an object in the device's field of view (FOV) and captured by an infra-red sensor, making it possible to calculate the distance of the object using triangulation [26]. The range of the depth sensor is 2.3-20 ft. (restricted to 13 ft. by the SDK) The microphone array is comprised of 4 microphones, enabling sound source localization. For our application, we implemented the least intrusive setup possible by capturing data only from the depth sensor and the microphone array, without capturing the actual color video data.

The RFID system we have used is the commercially available Alien 9900+ developer kit, which includes a reader with two circularly polarized antennas. The tags used in our experiment are EPC Class 1 Generation 2 supported by the 9900 readers. Figure 3 shows an example tag and antenna design from Alien. As the antennas are circularly polarized, the tag orientation is not an issue for our experiment. However, for an indoor environment, the antenna read range for the passive RFID tags varies from 20 to 30 ft. Such a read range is sufficient to detect the presence of a person carrying a tag in the simulated rooms of our Heracleia Assistive Apartment, given the tags are within the FOS of the antennas.

B. System Architecture

The architecture of our system is modular, comprising of 3 main components as shown in figure 4.

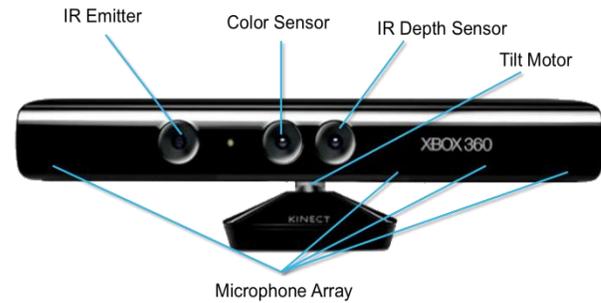


Fig. 2. The MS Kinect and the location of its various sensors and components.

Communication between the modules is based around the Joint Architecture for Unmanned Systems (JAUS) [27], originally developed by the U.S. Department of Defense, to govern the way that unmanned systems are designed. The user datagram protocol (UDP) is used for inter-module communications, which increases the level of interoperability, allowing new software modules to be easily integrated in the system or existing modules to be installed on different systems. Input is provided by the RFID reader and the 2 Kinect devices. One of them is considered as primary, capturing both a stream of depth images and audio, while the secondary captures only audio for performing sound localization. Interfacing with the Kinect is carried out using the MS software development kit (SDK) v1.0 [28]. The 3 modules 1) skeletal tracking based localization, 2) audio localization and 3)RFID tracking are described in detail in the following paragraphs.



Fig. 3. RFID transceiver equipment used and example Alien RFID Tag and Antenna Designs.

1) Skeletal Tracking Based Localization Module

Skeletal tracking is used in our system in order to detect and track a person in the FOV of the sensor, as s/he moves in the smart space and it was implemented using the MS Kinect SDK. This module has been explained in our previous work [29] and therefore only briefly described.

When a person is detected to be moving, her/his center of mass is determined and a skeletal model is fitted. The detected skeleton has a unique identifier for a specific session and is defined by the 3-D coordinates of its 20 joints $\langle X_{di}, Y_{di}, Z_{di} \rangle$, expressed in meters.

Each joint can be at any of the three associated states: 1) tracked, 2) not-tracked and 3) inferred. Furthermore, two kinds of filters are applied to the joint coordinates due to the nature of the captured data, 1) high frequency jitter and 2) temporary spikes rejection. Localization using such skeletal tracking is very accurate and unintrusive since we only utilize the coordinates calculated from the depth sensor feed.

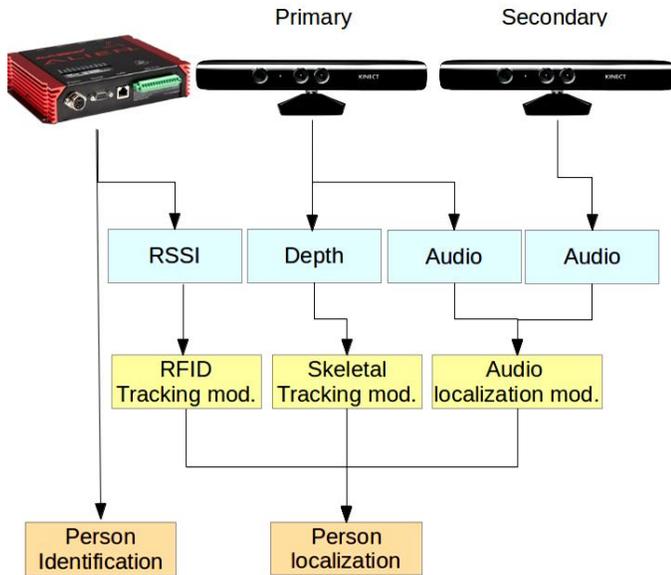


Fig. 4. System architecture showing the devices used (RFID and Kinect sensors), as well as the different types of information captured and processing pipelines for the identification and localization tasks.

2) Audio Localization Module

The audio localization module, acts as an auxiliary form of information input. The operation of the module has been described in detail in [29] and it is presented here epigrammatically. The Kinect incorporates a microphone array, comprised of 24-bit ADCs driven by 4 microphones. The frequency response of the microphones is tuned appropriately for speech and their directionality is isotropic for these frequencies. Sound source localization is applied to the audio signal in order to determine the angle of the sound source in relation to the device and acquire the audio signal from that particular direction. The returned values are the sound source angle (in degrees) in relation to the axis that is perpendicular to the device, and a confidence level of the reported angle.

Although the angle of the sound source can be acquired, this information is inadequate in estimating the source's distance. Thus, a second Kinect is introduced, used only for sound source localization (figure 5). The additional information provided by this unit can be used for accurate 2-D localization through triangulation. We denote θ_A, θ_B the angles between the wall and the axes perpendicular to devices A and B respectively and assuming there is a sound source S detected by the two devices, let the corresponding detected angles be $\phi_A, \phi_B \in (-50, 50)$. We consider the triangle that is created, with A, S and B as its vertices such that the altitude of the triangle passing from vertex S, divides L into a and b so that $a+b=L$. Let the length of the altitude (in our case the distance of

the audio source/person from the wall) be X_s . Since $L=a+b$, the final solution to the system of equations is given by:

$$X_s = \frac{\tan(\theta_A - \phi_A) \cdot \tan(\theta_B + \phi_B) \cdot L}{\tan(\theta_A - \phi_A) + \tan(\theta_B + \phi_B)}$$

$$a = \frac{X_s}{\tan(\theta_A - \phi_A)}$$

$$b = \frac{X_s}{\tan(\theta_B + \phi_B)}$$

This method allows for the calculation of the precise 2-D position of the audio source in the room.

Some additional restrictions concerning this setup were that the sound source angles are taken into account only when the sound level exceeds 50dB, the confidence for both estimated sound source angles is more than 50% and that there is a solution for the equation system and that this solution falls within the boundaries of the room.

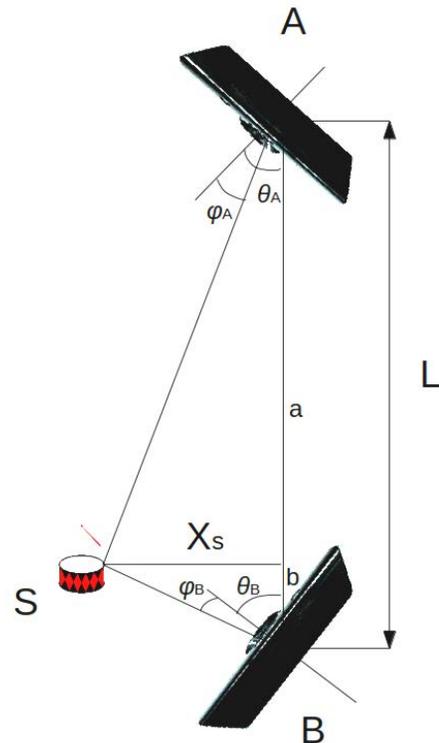


Fig. 5. Audio localization configuration utilizing the microphone arrays of 2 Kinect sensors.

3) RFID Based Localization Module

The RFID system that we used was comprised of two antennas and a tag reader. Its main role was to identify the person in its field of sense (FOS), but also to provide a rough estimate of her/his location using the received signal strength indicator (RSSI) from each antenna. The mapping between the RSSI values and the actual position of the tag is accomplished through a calibration process that accounts for both the

directionality of the antennas and the specific layout of the room. Multiple people are identified using their unique RFID tag and tracked as long as they remain in the FOS of the system. Skeletal tracking alone may not be able to discern between different people since a new tracking id is issued each time a person is lost from the FOV of the Kinect and then re-enters. Therefore, we improved our system's accuracy by matching the new RFID tag with the new tracking id as soon as an individual enters the room. This technique allows identification of each individual detected by the skeletal tracker. In the case where an unmatched tag id or skeletal id appears e.g. if a person was not detected upon entrance by either sensor, they are matched when they both appear in the same sector. Finally, when no skeleton is detected in the FOV, but a tag is still being detected, audio localization is utilized in order to increase accuracy (e.g. when only one antenna reads the tag).

Localization is based on a training phase during which statistical regression is applied on pre-specified position signatures (RSSI in our case) in order to build a classifier. More specifically, we divide the entire room into multiple sectors, as shown in figure 6. Next, we collect the RSSI signatures of the detected tags in these different sectors using the antennas. Given the measurement from the Kinect sensor for any particular person, if the measured location falls within that specific sector, then we map that particular person to the location described by the Kinect sensor. As afore-mentioned, in both approaches we use the sound from the microphone array as another modality besides skeletal tracking to resolve ambiguities in mapping.

III. SYSTEM OPERATION

The main function of our system is person localization utilizing information from all three modules. The main source of location information is the skeletal tracking module. More specifically, this module detects a person as soon as s/he enters the FOV of the sensor and tracks her/him while moving in the room. The accuracy and robustness of the tracker is exceptional due to the nature of the depth sensor, so the person is tracked while standing, walking or even sitting.

We consider the location of the person as the average of the 3-D coordinates of all the tracked joints, expressed as $\langle \bar{X}_d, \bar{Y}_d, \bar{Z}_d \rangle$, where:

$$\bar{X}_d = \frac{1}{20} \sum_{i=1}^{20} X_{di} \quad \text{the mean distance from the sensor's plane.}$$

$$\bar{Y}_d = \frac{1}{20} \sum_{i=1}^{20} Y_{di} \quad \text{the mean deviation from the sensor's axis.}$$

$$\bar{Z}_d = \frac{1}{20} \sum_{i=1}^{20} Z_{di} \quad \text{the mean distance from the floor.}$$

Another source of location information is the audio localization module. It should be noted that the audio localization module is capable of estimating the location of the

person in 2 dimensions expressed by $\langle X_s, a \rangle$, not accounting for height, as described in the previous section.

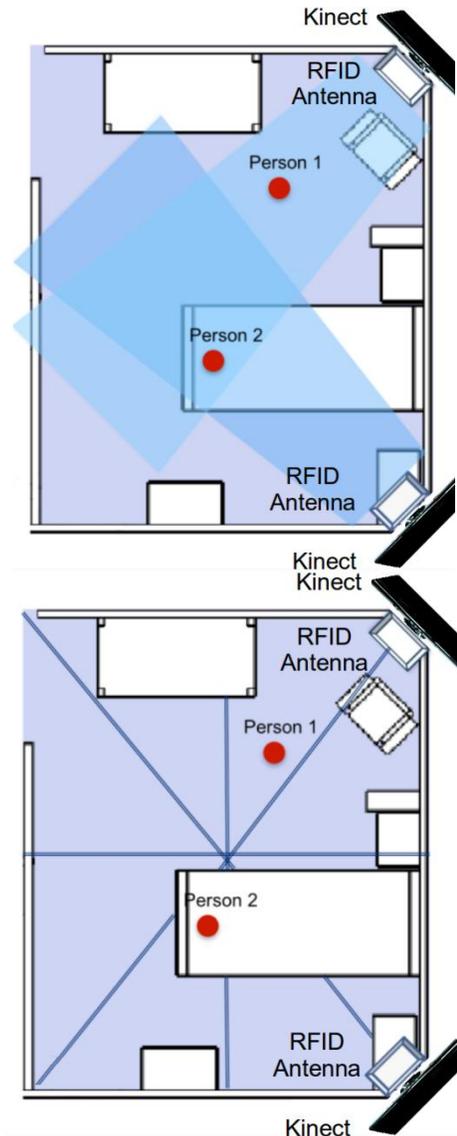


Fig. 6. Deployment setup combining the RFID RSSI and recognized sectors and the Kinect distance information for person localization.

In order to determine the final estimated location of the person we consider the available localization information from all three modules hierarchically, according to our experimental results presented in the next section. So, in the case where one of the modules does not return any coordinates, then the other module's coordinates are considered. The order in which we determine the location of each person is: 1) Skeletal tracker, 2) RFID, 3) Sound source localization. If skeletal tracking information becomes unavailable (e.g. if the person is outside the FOV of the depth sensor), then we rely on RFID. Similarly, if both skeletal tracking and RFID information are unavailable (e.g. tag undetected by 1 antenna), then sound source localization is used. In addition, we experimented by calculating the average location for each person.

More specifically, when a location estimate is available from both the RFID and the skeletal tracker, the average of each of the 2-D coordinates is calculated after proper transformation to match the 2 coordinate systems, while the third coordinate equals that of the skeletal tracking module. For our application, the detected activity is bound to the estimated location of the person. Therefore, if a person is standing by an appliance such as the oven or refrigerator we infer that s/he is using this particular appliance.

IV. EXPERIMENTAL SETUP

An extensive set of evaluation experiments were conducted in order to fine-tune the parameters of the setup at our simulated apartment (figure 7). As mentioned earlier, two Kinect devices and two RFID antennas were used, mounted at the opposite sides of one of the walls, facing the entrance. The distance between the two devices was 175.5 inches. The axis perpendicular to the sensors' axes pointed at 45 degrees towards the interior of the apartment, maximizing both the FOS, FOV and microphone coverage.



Fig. 7. An aspect of the Heracleia assistive apartment.

All modules were installed on the same computer, although our system's implementation permits the use of separate computers for each one of the modules. For our experiments we partitioned the space in 8 different sectors, intersecting at the center of the room. The estimated location of the person was considered accurate when the coordinates fell within the boundaries of the corresponding sector. For our application, the detected activity is bound to the estimated sector.

In our experimental setup, we have deployed two antennas at the two corners of the bedroom. We have simulated an experiment for identifying and localizing up to 4 people, limited only to part of the apartment, although the system can be extended to more rooms by adding more Kinect sensors in the apartment. During the experiment, each person wears an RFID tag around her/his neck.

We conducted extensive experiments in our realistic domestic setup. Four individuals participated in our experiments, with one, two or four occupying the apartment simultaneously. Subjects were asked to move in the apartment in 10 sessions and perform 4 activities, namely walk and sit in a chair, at a desk or on a bed. In table 1 we report results for

both the identification and localization tasks after 10-fold cross validation. For both tasks, accuracy degraded for more occupants, due to the people interacting and the resulting occlusions. Identification accuracy using RFID was at very high levels, considering single antenna misdetections. Localization accuracy denotes the percentage of correctly estimated locations for all individuals present in the room and also accounts for misidentifications and mismatches between the detected tag sector and skeletal id location. The accuracy attained using the Kinect was over 90%, and constituted the most accurate source for person location information. The accuracy achieved using RFID was over 80% and 75.9% using sound (only 1 speaker).

V. CONCLUSIONS

In this paper we presented the introduction of RFID technology to our existing novel person localization system improving its location estimation robustness and adding identification capabilities. Our system combines the tracking capabilities of the Kinect sensor with identification information from existing RFID technology. 3 types of data were used to solve the localization problem, namely the RSSI, 3D depth and audio information. Accurate position estimation for each person was carried out using the depth sensor and microphone arrays of the Kinect devices as inputs, by means of skeletal tracking and sound source localization respectively. The system was deployed in a simulated apartment and during the experiments conducted, it achieved high localization and identification accuracy for the 4-person localization scenario. More specifically, identification and localization accuracy always remained over 90% even in the 4 person scenario, when using information from RFID and the Kinect respectively. After confirming the effectiveness of our design we plan to extend it by utilizing depth and audio information from additional Kinect devices for increased robustness and coverage.

TABLE I. EXPERIMENTAL RESULTS FOR THE IDENTIFICATION AND LOCALIZATION TASKS FOR ALL THREE MODULES.

Task/Source	4 people	2 people	1 person
Identification/RFID	92.5%	97.1%	100%
Localization/Kinect	90.3%	93.8%	98%
Localization/RFID	82.1%	87.2%	85.4%
Localization/Sound	75.9% (1 speaker)		

ACKNOWLEDGMENTS

We would like to thank Shahina Ferdous for her invaluable help.

REFERENCES

- [1] T. Teixeira, D. Jung and A. Savvides, "Tasking networked cctv cameras and mobile phones to identify and localize multiple people", In Proc. Ubicomp, pp. 213-222, 2010.
- [2] R. Cucchiara, M. Fornaciari, A. Prati and P. Santinelli, "Mutual calibration of camera motes and RFIDs for people localization and identification", In Proc. ICDCS, 2010.

- [3] C. H. Kuo and R. Nevatia, "How does Person Identity Recognition Help Multi-Person Tracking?", In Proc. CVPR, pp. 1217-1224, 2011.
- [4] D. Mitzel, E. Horbert, A. Ess and B. Leibe, "Multi-person tracking with sparse detection and continuous segmentation", In Proc. ECCV, pp. 397-410, 2010.
- [5] M. Andriluka, S. Roth and B. Schiele, "People Tracking-by-Detection and People Detection by Tracking", In Proc. CVPR, pp. 1-8, 2008.
- [6] B. Wu, R. Nevatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet Part Detectors", In Proc. IJCV, pp. 247-266, 2007.
- [7] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. H. and S. Shafer, "Multi-camera multi-person tracking for easy living", In Proc. IEEE IWVS, pp. 3-10, 2000.
- [8] D. Focken and R. Stiefelbogen, "Towards Vision-Based 3-D People Tracking in a Smart Room", In Proc. IEEE ICMI, pp. 400-405, 2002.
- [9] Y. Chang, R. Yan, D. Chen and J. Yang, "People identification with limited labels in privacy-protected video", In Proc. ICME, pp. 1005-1008, 2006.
- [10] B.S. Ashar and A. Ferriter, "Radiofrequency identification technology in health care: benefits and potential risks", JAMA: The Journal of the American Medical Association, 298:2305-2307, Nov. 2007.
- [11] K. Finkenzerler, "RFID handbook: fundamentals and applications in contactless smart cards and identification", Wiley, New York, NY, USA, 2003.
- [12] T. J. Hazen, E. Weinstein, B. Heisele, A. Park, and J. Ming, "Multimodal Face and Speaker Identification for Mobile Devices," Book Chapter in Face Biometrics for Personal Identification: Multi-Sensory Multi-Modal Systems, R. I. Hammoud, B. R. Abidi, and M. A. Abidi (Editors), pp. 123-138, Springer, 2007.
- [13] N. Fox, R. Gross, P. Chazal, J. Cohn, and R. Reilly, "Person Identification Using Automatic Integration of Speech, Lip and Face Experts," In ACM SIGMM 2003 Multimedia Biometrics Methods and Applications Workshop, Berkeley, CA, 2003, pp. 25-32.
- [14] H. K. Ekenel, M. Fischer, Q. Jin, and R. Stiefelbogen, "Multi-modal person identification in a smart environment," CVPR Biometrics Workshop, Minneapolis, USA, June 2007.
- [15] D. Hahnel, W. Burgard, D. Fox, K. Fishkin and M. Philipose, "Mapping and localization with RFID technology", In Proc. IEEE ICRA, pp.1015-1020, 2004.
- [16] M. Bouet and A. L. D. Santos, "RFID tags: Position principles and localization techniques", Wireless Day, pp. 1-5, 2008.
- [17] L. Ni, Y. Liu, Y. Cho Lau and A. Patil, "LANDMARC: Indoor location sensing using active RFID", Wireless Networks, pages 701-710, 2004.
- [18] P. Wilson, D. Prashanth and H. Aghajan, "Utilizing RFID signaling scheme for localization of stationary objects and speed estimation of mobile objects", In Proc. IEEE International Conference on RFID, pp. 94-99, 2007.
- [19] D. Hahnel, W. Burgard, D. Fox, K. Fishkin and M. Philipose, "Mapping and localization with RFID Technology", TR IRS-TR-03-014, Intel Research, 2003.
- [20] D. Joho, C. Plagemann and W. Burgard, "Modeling RFID signal strength and tag detection for localization and mapping", In Proc. IEEE ICRA, pp. 3160-3165, 2009.
- [21] M. M. Trivedi, K. S. Huang and I. Mikic, "Dynamic context capture and distributed video arrays for intelligent spaces", In IEEE Transactions on Systems: Man and Cybernetics - Part A: Systems and Humans, volume 35, January 2005.
- [22] L. Klingbeil and T. Wark, "A wireless sensor network for real-time indoor localization and motion monitoring", In Proc. International Conference on Information Processing in Sensor Networks, pp. 39-50, 2008.
- [23] J. A. Stankovic, et al, "Wireless sensor networks for in-home healthcare: potential and challenges", In Proc. HCMDSS, 2005.
- [24] L. Zhang and Z. Wang, "Integration of RFID into wireless sensor networks: Architectures, opportunities", In Proc. GCCW, pp. 463-469, 2006.
- [25] "The PrimeSensor Reference Design", [online] available at: <http://www.primesense.com/?p=514>
- [26] C. Liebe, C. Padgett, J. Chapsky, D. Wilson, K. Brown, S. Jerebets, H. Goldberg and J. Schroeder, "Spacecraft hazard avoidance utilizing structured light", In Proc. IEEE Aerospace Conference, pp.10, 2006.
- [27] S. Rowe and C. Wagner, "An Introduction to the Joint Architecture for Unmanned Systems (JAUS)", Technical Report, Cybernet Systems Corporation.
- [28] "The Microsoft Kinect SDK", [online] available at: <http://msdn.microsoft.com/en-us/library/hh855347.aspx>
- [29] G. Galatas, S. Ferdous and F. Makedon, "Multi-modal person localization and emergency detection using the Kinect", International Journal of Advanced Research in Artificial Intelligence, vol. 2 issue 1 pages 41-46.

AUTHORS PROFILE



Georgios Galatas is pursuing his PhD in Computer Engineering at the Heracleia Human-Centered Laboratory of the Computer Science and Engineering department of the University of Texas at Arlington. He is also a research fellow of the Institute of Informatics and Telecommunications of the National Center for Scientific Research "Demokritos". He received his combined bachelor's and master's degree from the Electrical and Computer Engineering department of the University of Patras in 2008. His research interests include Computer Vision, Digital Image Processing, Automatic Speech Recognition and Person Localization. He has co-authored several peer reviewed papers published in technical conferences and journals.



Fyllia Makedon is Distinguished Professor and Department Head of Computer Science and Engineering at the University of Texas at Arlington. She received her PhD in Computer Science from Northwestern University in 1982. Between 1991-2006, she was professor of computer science at Dartmouth College where she founded and directed the Dartmouth Experimental Visualization Laboratory (DEVLAB). Prof. Makedon has received many NSF research awards in the areas of trust management, data mining, parallel computing, visualization and knowledge management. She is author of over 300 peer-reviewed research publications. She directs the Heracleia Human-Centered Laboratory that develops assistive technologies for human monitoring and smart health. She is member of several journal editorial boards and chair of the annual PETRA conference (www.petrae.org).

Improving the Security of the Medical Images

Ahmed Mahmood
School of Engineering
University of Guelph
Guelph, ON, CANADA

Charlie Obimbo
School of Computer Science
University of Guelph
Guelph, ON, CANADA

Tarfa Hamed
School of Computer Science
University of Guelph
Guelph, ON, CANADA

Robert Dony
School of Engineering
University of Guelph
Guelph, ON, CANADA

Abstract—Applying security to the transmitted medical images is important to protect the privacy of patients. Secure transmission requires cryptography, and watermarking to achieve confidentiality, and data integrity. Improving cryptography part needs to use an encryption algorithm that stands for a long time against different attacks. The proposed method is based on number theory and uses Chinese remainder theorem as a backbone. This approach achieves high level of security and stands against different attacks for a long time. On watermarking part, the medical image is divided into two regions: a region of interest (ROI) and a region of background (ROB). The pixel values of the ROI contain the important information so this region must not experience any change. The proposed watermarking technique is based on dividing the medical image in to blocks and inserting the watermark to the ROI by shifting the blocks. Then, an equivalent number of blocks in the ROB are removed. This approach can be considered as lossless since it does not affect on the ROI, also it does not increase the image size. In addition, it can stand against some watermarking attacks such cropping, and noise.

Keywords—Medical Imaging Security; Telemedicine Security; Chinese remainder theorem; Watermarking

I. INTRODUCTION

Until recently the sole responsibility of keeping patients' records in confidence was with the Physicians. This meant that the Physician was not to disclose any medical information revealed by a patient or discovered by a physician in connection with the treatment of a patient to any unauthorized person [1]. However, with the advent of recent computer technology, and it's permeation into the Medical field through E-health [2], Telemedicine [3-6], to name but a few, the challenges of confidentiality arising from the storage and transmission of medical data cannot be left to physicians alone.

Indeed, transferring medical data such as radiological results from a medical database center to another one without applying security techniques means low level of privacy for patients. Medical information transmission has increased with the use of telemedicine. Telemedicine is important because it enables consultations by remote specialists, loss-free and immediate availability of individual patient information, and improved communication between partners in a health care system [7].

Security of medical information imposes three mandatory characteristics: confidentiality, reliability and availability. Confidentiality means that only the entitled users have access to the information and this can be achieved using encryption. Reliability has two aspects; i) Integrity: the information must not been modified by unauthorized people, and, ii) Authentication: a proof that the information belongs indeed to the correct patient and is issued from the correct source and one of the techniques to achieve this is watermarking. Availability is the ability of an information system to be used by the entitled users in the normal scheduled conditions of access and exercise. For storage and transmission, encryption is a very efficient tool, but once the sensitive data is decrypted, the information is not protected anymore. Once the images are in the open (plain-text) form, the major threat is the violation of the access rights and of the daily logs by the intruder.

Watermarking is made to introduce identifiers, which, by construction, are inseparable from the document they are embedded in. They may be seen as ultimate ramparts against usurpation and fabrication. Medical tradition is very strict with the quality of biomedical images, in that it is often not allowed to alter in any way the bit field representing the image (nondestructive) [8]. Watermarking technique is based on the data modification principle. Therefore, the watermarking method must be reversible, in that the original pixel values must be exactly recovered. This limits significantly the capacity and the number of possible methods. It also constrains to have dedicated routines to automatically suppress and introduce the mark in order to prevent the transmission of unprotected documents. However, applying watermarking blindly is not acceptable in the medical imaging field where any modification in the high information area of the image is not acceptable. Dividing the medical image into two regions can solve this problem: a region of interest (ROI) and a region of background (ROB). The pixel values of the ROI contain the important information so this region must not experience any change. On the other hand, the ROB can provide a suitable place to embed the watermarking data.

In this paper next section provides background about medical imaging encryption (MIE) and its current methods in Section II. Section II also illustrates medical images watermarking (MIW) and the current approaches in the

medical field. Section 3 presents the proposed approaches in the encryption and watermarking areas. Image analyzing methods that are used to measure the performance of the proposed algorithm such as histograms, correlation coefficients, and average intensity difference are described in section IV. Finally, the conclusions and the future work are presented in last section.

II. MEDICAL IMAGE ENCRYPTION & WATERMARKING

A. Medical Image Encryption

Medical images may be encrypted to ensure privacy and integrity. As known, the strength of many encryption algorithms lies in the encryption key and its length, which is a major design issue. The encryption scheme is said to be computationally secure if it meets the following criteria:

- 1) *the cost of breaking the cipher exceeds the actual value of the encrypted information;*
- 2) *the time required to break the cipher exceeds the useful lifetime of the information.*

Image encryption, using CRT is performed by Thien and Lin [9] in 2002, using lossless and lossy forms. In 2006 Meher and Patra [10] used CRT secret image sharing in a naïve way. There have been some attempts to improve the performance of these methods [11-13].

1) Encryption Types

Stream ciphers and block ciphers are the main parts of symmetric algorithms. For stream cipher methods, the procedure is to encrypt one single bit of plaintext at a time, while the block ciphers take a number of bits and encrypt them as a single unit. Stream ciphers typically execute at a higher speed than block ciphers and have a lower complexity. However, stream ciphers can be vulnerable to serious security problems if used incorrectly [14].

2) Encryption Properties

Important properties for designing encryption schemes for medical data applications are presented below [15] [16]:

- The data size of medical images is large due to lossless compression. The encryption/decryption speed of some existing ciphers is not fast enough. This is especially true of software implementation that uses the naïve approach. Hence, the size of the data to be encrypted is an important consideration in the design of encryption schemes for medical imaging.
- Compressibility: when compression is applied after encryption, the randomness of the cipher text will considerably decrease the amount of compression achieved. As a result, one approach is to encrypt the content after compression; however after compression the entire compressed content needs to be encrypted. For this reason a stage within compression needs to be identified where partial encryption can be performed without affecting the compression. Consequently, there is a tradeoff needed between compression and encryption.
- The avalanche property: a good encryption algorithm should have the avalanche property where a small

change in either plaintext or the key should result in a huge change in the ciphertext.

- Security and usability: medical images may stored for a long time; therefore, the encryption algorithm should result in a ciphertext that can stand against different attacks.

3) Encryption Attacks

Attacks can be classified into two basic categories. In the first type, the attacker has some knowledge of the algorithm and/or a sample of a plaintext-ciphertext pair. In the second type, the attacker has no knowledge of the algorithm; this is known as brute-force attack when every possible key on a piece of ciphertext is tried until its plaintext is obtained [17]. The different types of cryptanalytic attacks are explained in the following paragraphs.

- Ciphertext - Only Attack.

In this attack the attacker is able to get part of the ciphertext when the attacker has some medical data that are encrypted by the same encryption algorithm. The attacker's job is to recover the plain- text of as many images as possible, or to deduce the key (or keys) used to recover the images [17] [18] [19]. An example of this attack is the jigsaw puzzle attack. The attacker first divides a cipher image into many small pieces; then, the attacker tries to break these pieces simultaneously.

As each piece is very small compared with the whole cipher image, the time needed to break each piece is less than that needed to break the entire cipher image [17].

- Known - Plaintext Attack.

This attack occurs when attacker is able to have access to some cipher images and their original images. This may help in determining the key or a part of the key [17] [18] [19].

- Chosen- Plaintext Attack.

An attacker is able to select some medical images and get the relating cipher images. This occurs when the attacker not only has access to the cipher images and the original medical images, but the attacker also has the ability to choose the image parts that get encrypted. This is more powerful than a known-plaintext attack because the cryptanalyst can choose specific image blocks to encrypt [14] [20] [17].

- Chosen - Ciphertext Attack.

In this attack, the attacker is able to get several cipher images and original images [14] [20] [17]. When an attacker is able to modify the choice between the two types of images based on the results of previous encryption.

The DICOM medical images are usually secured using classical encryption algorithms such as advanced encryption standard (AES) or triple data encryption standard (3DES) [21]. Encrypting a medical image using AES or 3DES encryption algorithms provide a high level of security, but require long processing time.

For example, encrypting an MRI brain image with dimensions of 512×512 pixels takes 521.67 s using a computer that runs on a CPU Intel Core2 Quad Q6700 [22].

B. Medical Image Watermarking

Watermarking is the process of embedding small sensitive data such as copyright and ownership identification in images; it has become a necessary component of multimedia applications that are subject to illegal use [23]. In addition, it is used for data authentication purposes to detect any changes in a medical image. The robust watermarking algorithm should be able to retrieve sensitive data after applying different image processing such as translation, resizing, and cropping, as well as different types of distortions such as filtering, and contrast. These issues are very important when creating a new design for a strong watermarking method.

The three types of watermarking methods for medical images are

- 1) *minimum distortion*,
- 2) *lossless watermarking*, and
- 3) *segmentation* [24].

The first method consists in using classical watermarking methods while minimizing the distortion. In this case, the watermark replaces some image details, such as the least significant bit of the image. However, embedding different watermarks may cause degradation in the watermarked medical images and this degradation in the image quality is measured with a number of metrics, such as MSE and PSNR [25]. Lossless or reversible watermarking represents the second type: the watermark can be removed from the image once the embedded data is read, allowing retrieval of the original image. This approach provides authentication without proof of ownership. The third type is implemented by segmenting the medical image using image segmentation techniques into two regions where the first is known as region of interest (ROI) and the second as region of background (ROB); the watermarked data are embedded within the ROB in order not to compromise the diagnosis capability. Image segmentation techniques that are necessary for deciding the suitable areas for using the watermarking technique.

The two main categories of digital watermarking are

- 1) *visible* and
- 2) *invisible*.

A visible watermark is similar to stamp a watermark on paper. It is seen in many digital applications such as the logos of television channels, or the data of medical images as shown in Figure 1. On the other hand, invisible watermarking is often used to identify copyright data, for example, an author or a distributor. There are different classifications of invisible watermarking algorithms. Watermarking approaches can be distinguished in terms of the watermarking host such as text, audio, images, and video. In addition, they can be classified types according to whether the extraction of the original signal is non-blind, semi-blind, or blind. In non-blind schemes, both the original image and the secret key are needed, while in semi-blind schemes, both the secret key and the watermark are needed; blind schemes need only the secret key.

The main classes of watermarking techniques are spatial domain, and transform domain. The watermark in the spatial domain technique represents the first class where it is

embedded by changing the pixel values of the original image, while in the transform domain technique, the data are



Fig. 1. Visible Watermarking of a Medical Image

embedded by modulating the transform domain signal coefficients. In [26] Li uses the moment-preserving threshold, which is a pixel-based segmentation to separate the ROI from the ROB for mammogram medical images. However, ultrasound and brain images most of the time create confusion for the ROI and ROB so some researchers have tried to offer a solution. For example in [27] Cao suggests adding a digital envelope so the important data can be embedded in it. This solution increases the size of the transmitted data.

C. Watermarking Properties

Properties for an efficient watermarking system are application dependent; one of the challenges in this area is that these properties compete with each other. None of the digital watermarking techniques have yet to meet all of these properties.

- **Robustness.** Digital images commonly are subject to many types of distortions, such as filtering, resizing, and cropping. These distortions are still very common and represent an open issue with respect to the robustness of watermarking. However, the mark should be discovered if these distortions occurred.
- **Capacity.** The capacity of the hidden data is another important issue where the watermarking algorithm should embed a predefined number of bits that can be hidden in the host signal. This number will depend on the application and there is no general rule for this. In general, the number of bits that can be inserted in the data is limited and in the LSB method; it is between (0.125 - 0.25) of the total size.
- **Invisibility.** There are two types of invisibility due to the implementation method: perceptual invisibility, and statistical invisibility. In perceptual invisibility the

Identify applicable sponsor/s here. If no sponsors, delete this text box (sponsors).

watermark is hidden in such a way that it is hardly noticed. An unauthorized person should not be able to detect the watermark by means of statistical methods. For example, the availability of a large number of digital works watermarked with the same code should not allow the extraction of the embedded mark by applying statistically based attacks. A possible solution is to use a content-dependent watermark.

D. Watermarking Attacks

The main challenges in the watermarking research area are a lack of standards and benchmarking, and a lack of comprehensive mathematical theory. The two basic types of watermarking are visible and invisible watermarking.

An image watermark may survive many attacks. The attacks related to invisible watermarking are similar to stego-attacks. Attacks on visible watermarks include an analysis of lighting and shadows, localized analysis of noise, histogram, and looking for discontinuities. Watermark attack types based on the aim of the attacker are as follows [28] [29].

Possible attacks on MIW are grouped into the followings.

- *Passive Attacks.* The attacker tries to determine whether a watermark is present but the removal of the watermark is not a goal.
- *Active Attacks.* Active attacks can be divided into three types as follows:
 - *Robustness Attack.* In this attack the attacker attempts to remove or destroy the watermark, so that the watermark detector is unable to detect watermark, and key issue in proof of ownership, fingerprinting, copy control. Types of this attack include geometrical, filtering, and noise.
 - *Collusion Attacks.* The attacker uses several copies of watermarked data (images, video etc.) to find the watermark and to construct a copy with no watermark. This is serious for fingerprinting applications.

Forgery Attacks. The attacker tries to embed a valid watermark. This has serious implications in authentication

III. PROPOSED ALGORITHM

The proposed algorithm has two parts. The first part is to perform the watermarking part and the second part is to implement strong encryption using the Chinese remainder theorem (CRT). Figure (2) presents the steps of the proposed algorithm. The watermarked data is the username of the person that requested for the medical images and a serial number between the two parties to achieve authenticity and avoid any forgery one.

This watermark that contains characters and numbers are converted to the hexadecimal representation then it is encrypted using Caesar shift cipher method. The watermark insertion place is selected near to the edge of the medical image. This method does not affect the ROI because the insertion of the watermark is in ROB, therefore this method can be considered as a lossless method.

The encryption part starts by selecting two relatively prime numbers one of them should be 256 so the encrypted pixel value does not exceed 255 which is the maximum value for an eight bit pixel. In order to have a good understanding about the CRT implementation is presented the following equations.

Consider $n \geq 2$, and m_1, m_2, \dots, m_n are positive relatively prime integers. Let the integer b_i denote the remainder of x modulo m_i for $1 \leq i \leq n$. The CRT is represented by the

$$\begin{cases} x \equiv b_1 \pmod{m_1} \\ x \equiv b_2 \pmod{m_2} \\ x \equiv b_3 \pmod{m_3} \end{cases} \quad (1)$$

following system that has a unique solution x .

This by Obimbo's notation may be written as:

$$x \equiv (b_1, b_2, b_3)S(m_1, m_2, m_3) \quad (2)$$

The solution of x in equation 2 can be computed in a number of ways [30][31][33][34]. One of the approaches to solve these equations is Charlie's method. This method is selected because it has higher performance and faster than other methods [33]. The method starts by computing the value of x that satisfies the equations iteratively as shown below:

$$x_n = m_n \quad (3)$$

$$x_{n-1} = x_n + l_n \times k_n \quad (4)$$

$$\text{where, } l_{n-1} = m_{n-1} \times m_n \quad (5)$$

$$\text{and } k_n = l_{n-1} \pmod{m_{n-1}} \quad (6)$$

IV. IMAGE ANALYZING METHODS

In order to verify the security and the performance of a new algorithm, this algorithm should be analyzed and tested according to the image features. Some of the used keywords are defined.

The image mean can be defined as the average pixel value of an image, and for grey-scale medical images it is equal to the average brightness or intensity, while the image variance gives an estimate of the spread of pixel values around the image mean.

A. Histogram

The encrypted image histogram should be close to the uniform distribution to avoid statistical attacks [32]. The histogram of an image shows the number of occurrences for each grey level in the medical image. Mathematically, the histogram is a discrete function and its grey levels are in the range $[0, L - 1]$ as in the following equation:

$$\text{hist}(r_k) = \frac{n_k}{N} \quad (7)$$

Where r_k is the k th grey level, and n_k is the number of pixels in the image with that grey level. N is the total number of pixels in the image. It may be noted that $k = 0, 1, \dots, L - 1$.

The histogram gives a global description of the image, so having a narrow histogram of the image means that the image is poorly visible because the difference in grey levels present in the image is generally low [35]. In the same way, a widely distributed histogram means that almost all the grey levels are present in the image, and thus the overall contrast and visibility increases.

B. Entropy

Entropy is a statistical measure of disorder and randomness. The entropy $En(d)$ of data d is measured as [34]:

$$En(d) = \sum_{i=1}^L p(m_i) \log \frac{1}{p(m_i)} \quad (8)$$

Where L is the total number of pixels and $p(m_i)$ represents the probability occurrence of a pixel with value m_i . When the entropy of the encrypted image is close to $\log L$ bits, its histogram is considered sufficiently uniform.

C. Difference Between Original & Watermarked Images

This measurement is very useful to show the effect of the watermark on the image. If the input image of a system is $f(x,y)$, and the watermarked image of that system is $g(x,y)$, then the error function $e(x,y)$ can be defined as the difference between the input and the watermarked images [36].

This difference value between the two images represents the effect of watermarking, as expressed in the following equations:

$$e(x, y) = f(x, y) - g(x, y) \quad (9)$$

The mean square error Ems (or MSE) formula is:

$$Ems = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} e(x,y)^2 \quad (10)$$

And the peak signal to noise ratio (PSNR) formula is described below:

$$PSNR = 10 \log_{10} \frac{255^2}{Ems} \quad (11)$$

The algorithm used to place the encrypted watermark is illustrated below in Figure 2.

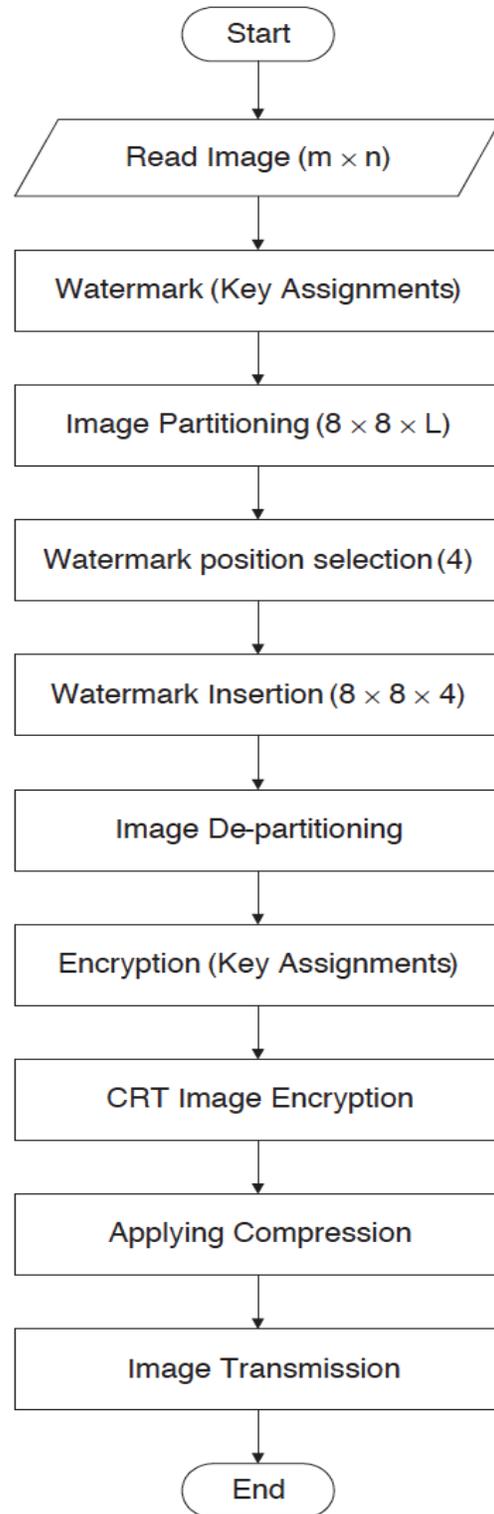


Fig. 2. Algorithm for Watermark Encryption & Placement

V. RESULTS AND DISCUSSIONS

Watermarked images are shown below in Figures 3 and 4

where the watermark cannot be noticed visually.

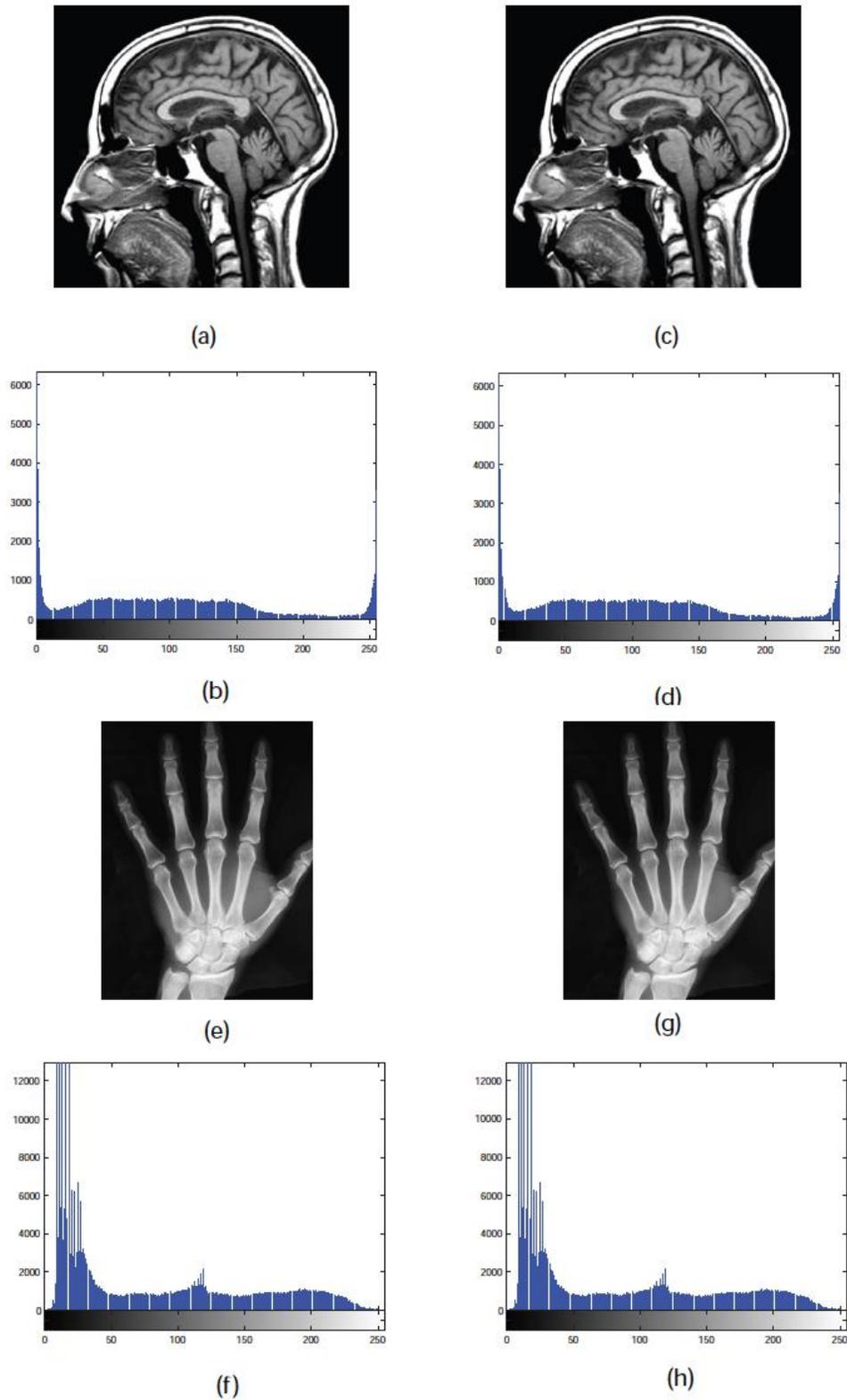


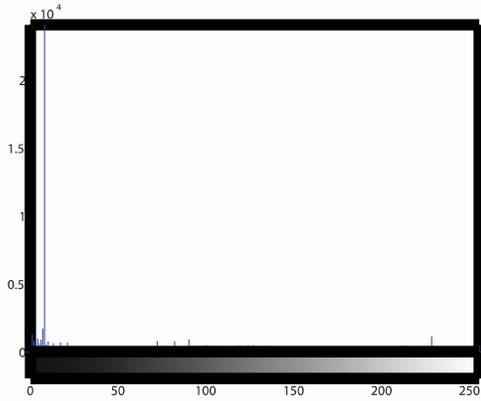
Fig. 3. Watermarked Images with Their Histograms



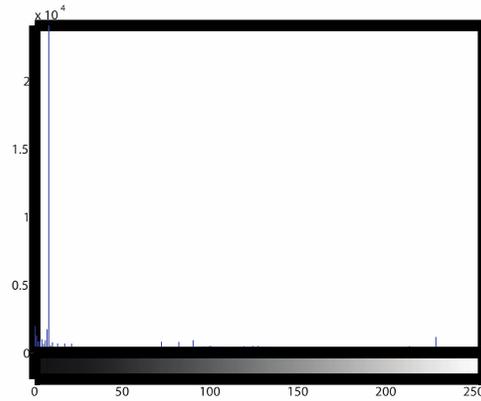
(a)



(c)



(b)



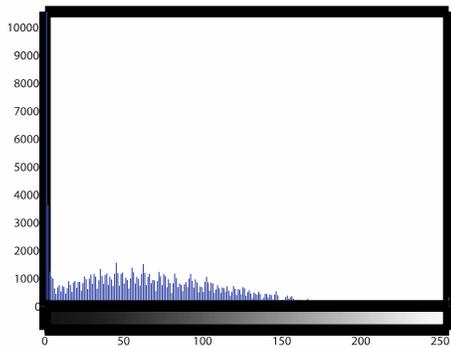
(d)



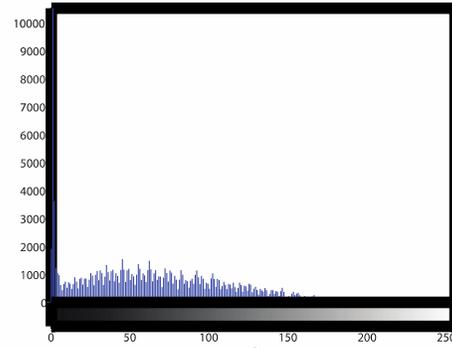
(e)



(g)



(f)



(h)

Fig. 4. Watermarked Images with Their Histogram

The Ems and PSNR calculation are shown in Table 1. It can be seen that the insertion of the watermark had a small effect on the Ems and PSNR values.

TABLE I. EMS AND PSNR FOR THE WATERMARKED IMAGES

<i>Image</i>	<i>EMS</i>	<i>PSNR</i>
CT Ankle	0.0688	59.7538
MRI Brain	0.0739	99.4457
X-ray Hand	0.1895	55.3553
Ultra sound	0.0331	82.9366

On the other hand, the CRT method did not show good performance to secure the medical images and this is related to characteristics of the medical images.

Table 2 shows the results of CRT encryption using two bits. The increase of bits improves the performance where the entropy was better for the four bits CRT. However, the eight bits CRT entropy was lower than the four bits value.

TABLE II. ENTROPY FOR THE ENCRYPTED IMAGES

<i>Image</i>	<i>Entropy after Watermark</i>	<i>Entropy after Encryption</i>
CT Ankle	3.6834	4.3804
MRI Brain	6.2493	6.4751
X-ray Hand	6.4386	6.9569
Ultra sound	5.7701	6.3961

The results of CRT encryption using base 2 are shown below in Figure 5.

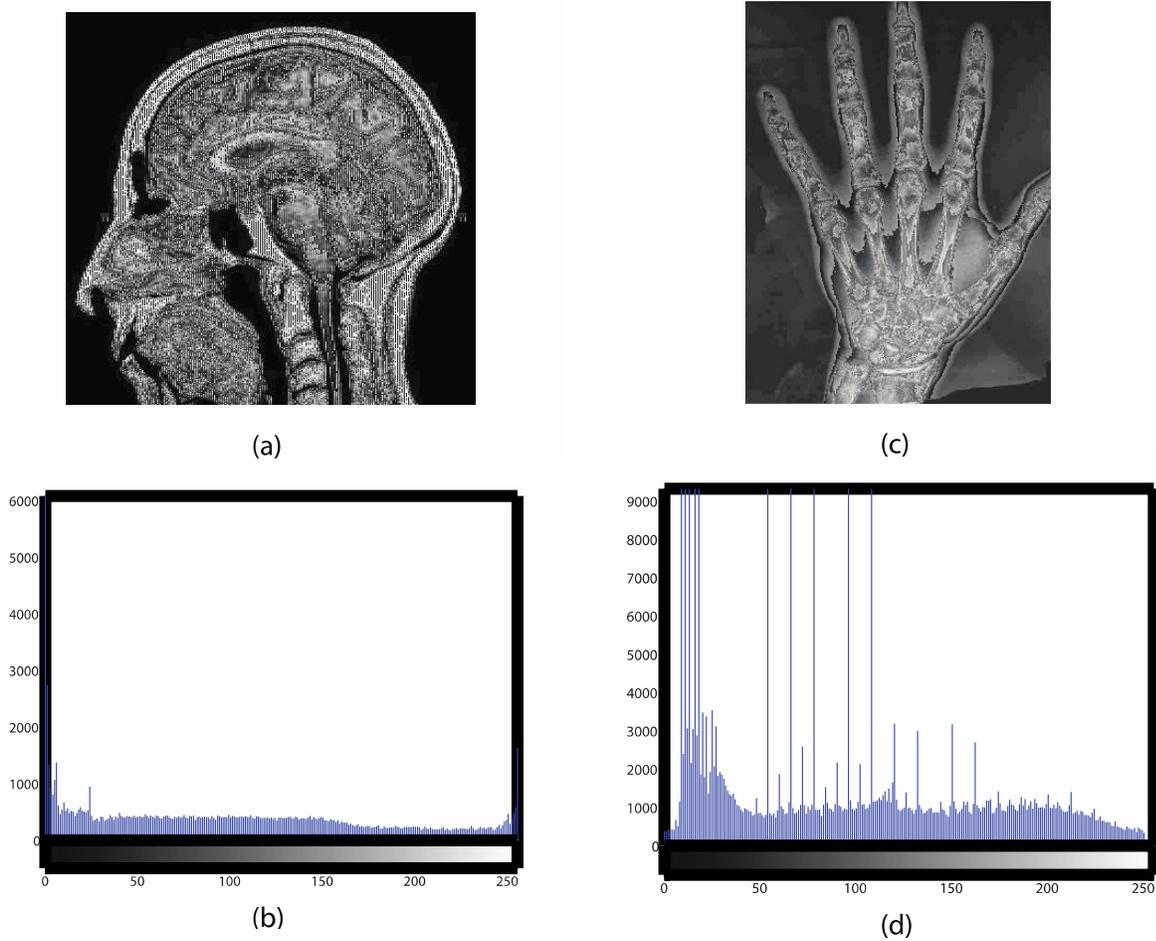


Fig. 5. Encrypted Images with Their Histograms

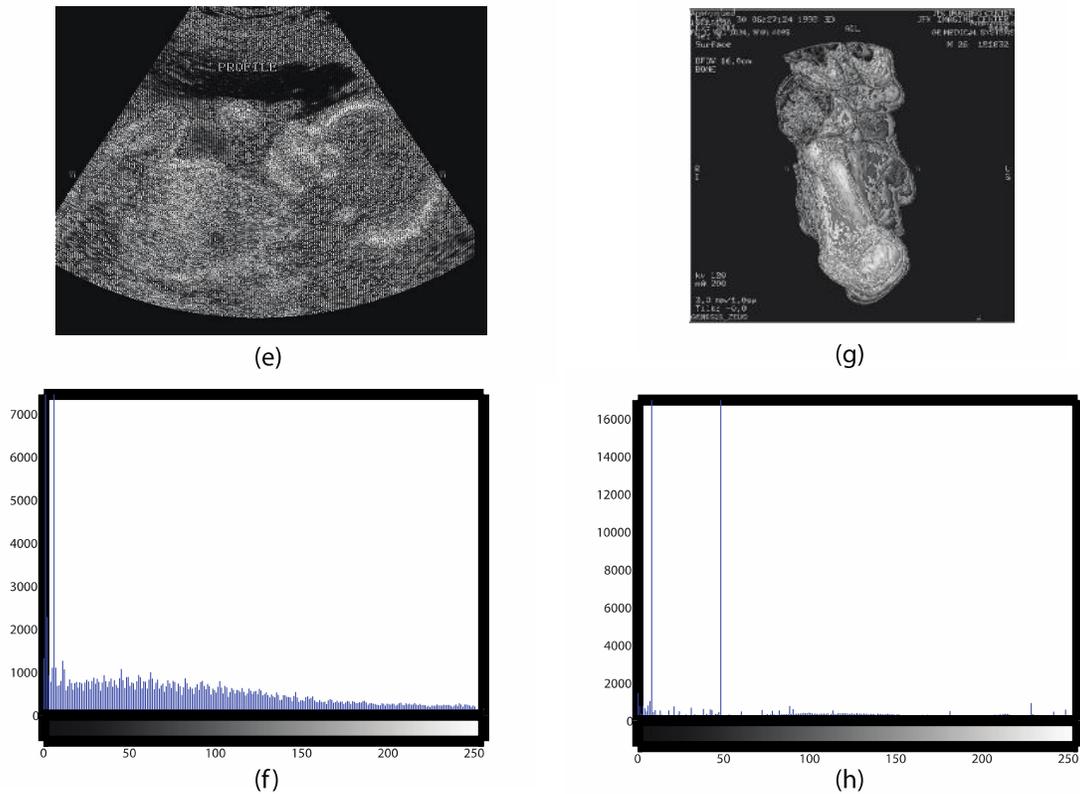


Fig. 5. (cont.): Encrypted Images with Their Histograms

The algorithm used was found to be considerably faster than the one for encryption for both AES and triple-DES. Figure 6 shows the processing time of proposed algorithm compared with AES and 3DES encryption algorithms. AES

and 3DES are the adopted algorithms for securing medical images. The processing time was obtained using a MATLAB 7.10 code in a computer runs on a CPU Intel i7 820.

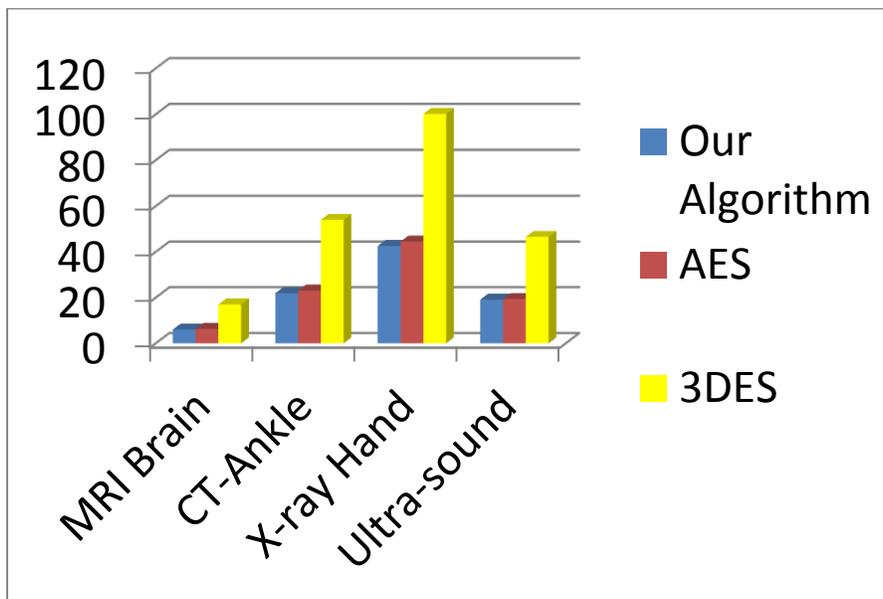


Fig. 6. Encryption Time of Different Algorithms

VI. CONCLUSION

Security techniques become more complex as the speed of computers increases. Providing a secure algorithm to protect medical images is complicated because of the special concerns of the medical community. First, the algorithm must recover the exact image with no change in any pixel value. Second, the algorithm should have a short processing time with high security that can stand for a long time.

Since the use of CRT as an encryption method for Medical images is novel, it can still be improved to work better with in this area. The proposed security algorithm tries to obtain better performance by reducing the encryption processing time compared with the current methods such as AES. At the same time, the time needed to break this method is still very big. In addition, the watermarking approach does not affect on the ROI of the medical image, therefore it can be considered as lossless

REFERENCES

- [1] Patient Confidentiality. American Medical Society. Available from: <http://www.ama-assn.org/ama/pub/physician-resources/legal-topics/patient-physician-relationship-topics/patient-confidentiality.page>. Retrieved 2013-23-09.
- [2] "National E-Health Transition Authority: About Us". National E-Health Transition Authority. 2013. Retrieved 2013-23-09.
- [3] Berman, Matthew; Fenaughty, Andrea (June 2005). "Technology and managed care: patient benefits of telemedicine in a rural health care network". *Health Economics* 14 (6). Wiley. p. 559-573. doi:10.1002/hec.952.
- [4] Van't Haaff, Corey (March/April 2009). "Virtually On-sight". *Just for Canadian Doctors*. p. 22.
- [5] Saylor, Michael (2012). *The Mobile Wave: How Mobile Intelligence Will Change Everything*. Perseus Books/Vanguard Press. p. 153.
- [6] Conde, Jose G.; De, Suvranu; Hall, Richard W.; Johansen, Edward; Meglan, Dwight; Peng, Grace C. Y. (January/February 2010). "Telehealth Innovations in Health Education and Training". *Telemedicine and e-Health* 16 (1). p. 103-106. doi:10.1089/tmj.2009.0152.
- [7] W. Yu-zeng, G. Shi-chao, F. Yu-jun, and F. Zhi- quan, "Research the Compression and Transmis- sion Technology of Medical Image Base on the Remote Consultation," in *The 2nd International Conference on Bioinformatics and Biomedical En- gineering (ICBBE)*, Shanghai, China, 2008, pp. 2142–2145.
- [8] Y. Tomioka, N. Aida, K. Kakehi, K. Nagami, H. Juzoji, and I. Nakajima, "Recent survey on patent applications for medical communications and telemedicine in Japan, USA and Europe," in *Proceedings of the 7th International Workshop on Enterprise Networking and Computing in Health*, pp 79 – 82.
- [9] C. Thien and J. Lin, "Secret image sharing," *Computers & Graphics*, vol. 26, no. 5, pp. 765–770, 2002.
- [10] P. Meher and J. Patra, "A new approach to secure distributed storage, sharing and dissemination of digital image," *Island of Kos, Greece*, 2006, pp. 373–376.
- [11] V. Jagannathan, A. Mahadevan, R. Hariharan, and E. Srinivasan, "Number theory based image com- pression encryption and application to image mul- tiplexing," *Chennai, India*, 2007, pp. 59 – 64.
- [12] S. J. Shyu and Y.-R. Chen, "Threshold secret image sharing by Chinese Remainder Theorem," *Piscataway, NJ, USA*, 2008, pp. 1332 – 1337.
- [13] S. Hou and T. Uehara, "A Content Providing System With Privacy Protection," *Nanjing, Jiangsu, China*, 2010, pp. 912 – 917.
- [14] W. Stallings, *Cryptography and Network Security: Principles and Practice*. Boston, MA: Prentice Hall, 2011.
- [15] A. Massoudi, F. Lefebvre, C. De Vleeschouwer, B. Macq, and J. Quisquater, "Overview On Selective Encryption Of Image And Video: Challenges And Perspectives," *EURASIP Journal on Information Security*, vol. 2008, pp. 1–18, 2008.
- [16] N. Kulkarni, B. Raman, and I. Gupta, "Multimedia Encryption: A Brief Overview," *Recent Advances In Multimedia Signal Processing And Communications*, pp. 417–449, 2009.
- [17] B. Schneier, *Applied Cryptography: Protocols, Algorithms, And Source Code In C*. New York, NY: John Wiley & Sons, Inc., 1996.
- [18] S. Li, C. Li, K.-T. Lo, and G. Chen, "Cryptanalyzing - An Encryption Scheme Based On Blind Source Separation," *IEEE Transactions On Circuits And Systems-I: Fundamental Theory And Applications*, vol. 55, no. 4, pp. 1055 – 1063, 2008.
- [19] S. Li, C. Li, Lo, and G. Chen, "Cryptanalysis Of An Image Scrambling Scheme Without Bandwidth Expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 3, pp. 338– 349, 2008.
- [20] S. Vaudenay, *A Classical Introduction To Cryptography: Applications For Communications Security*. Boston, MA: Springer, 2006.
- [21] A. B. Mahmood and R. D. Dony, "Segmentation based encryption method for medical images," in *Proceedings of the 6th International Conference on Internet Technology and Secured Transactions*, Abu Dhabi, United Arab Emirates, 2011, pp. 596 – 601.
- [22] Y. Zhou, K. Panetta, and S. Agaian, "A lossless encryption method for medical images using edge maps," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*, Minneapolis, MN, 2009, pp. 3707 – 3710.
- [23] I. Cox, L. Matthew, A. Jeffrey et al., *Digital Watermarking and Steganography*. Burlington, MA: Morgan Kaufmann Publishers), 2007.
- [24] G. Coatrieux, L. Lecornu, B. Sankur, and C. Roux, "A Review Of Image Watermarking Applications In Healthcare," in *International Conference IEEE Engineering in Medicine and Biology Society*, New York, NY, 2006, pp. 4691 – 4694.
- [25] B. Planitz and A. Maeder, "Medical Image Watermarking: A Study On Image Degradation," in *Proceedings of the Workshop on Digital Image Computing: Techniques and Applications*, 2005, pp. 3–8.
- [26] C.-T. Li, Y. Li, and C.-H. Wei, "Protection of Digital Mammograms on PACSs Using Data Hiding Techniques," *International Journal of Digital Crime and Forensics*, vol. 1, no. 1, pp. 75 – 88, 2009.
- [27] F. Cao, H. Huang, and X. Zhou, "Medical Image Security In A HIPAA Mandated PACS Environment," *Computerized Medical Imaging and Graphics*, vol. 27, no. 2-3, pp. 185–196, 2003.
- [28] L. Perez-Freire, P. Comesana, J. Troncoso- Pastoriza, and F. Perez-Gonzalez, "Watermarking Security: A Survey," in *LNCS Transactions on Data Hiding and Multimedia Security*, Heidelberg, Germany, 2006, pp. 41–72.
- [29] T. Kalker et al., "Considerations on watermarking security," in *Proceedings of the IEEE Workshop on Multimedia Signal Processing*, Cannes, France, 2001, pp. 201–206.
- [30] Donald Knuth. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89684-2. Section 4.3.2 (pp. 286–291), exercise 4.6.2–3 (page 456).
- [31] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 31.5: The Chinese remainder theorem, pp. 873–876.
- [32] Laurence E. Sigler (trans.) (2002). *Fibonacci's Liber Abaci*. Springer-Verlag. pp. 402–403.
- [33] C. Obimbo, "An Algorithm To Solve The Chinese Remainder Problem," in *The 2001 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'2001)*, Las Vegas, Nevada, USA, 2001, pp. 2246–2250.
- [34] K. Wong, "Image Encryption Using Chaotic Maps," *Intelligent Computing Based on Chaos*, pp 333–354, 2009.
- [35] Y. Mao and G. Chen, "Chaos-based image encryption," *Handbook of Geometric Computing*, pp. 231–265, 2005.
- [36] R. Gonzalez and R. Woods, "Digital Image Processing," Upper Saddle River, N.J., 2008.

Mining Positive and Negative Association Rules Using FII-Tree

T Ramakrishnu

Dept. of Computer Science and Engineering
National Institute of Technology
Warangal, AP, India

R B V Sbramanyam

Dept. of Computer Science and Engineering
National Institute of Technology
Warangal, AP, India

Abstract—Positive and negative association rules are important to find useful information hidden in large datasets, especially negative association rules can reflect mutually exclusive correlation among items. Association rule mining among frequent items has been extensively studied in data mining research. However, in recent years, there has been an increasing demand for mining the infrequent items. In this paper, we propose a tree based approach to store both frequent and infrequent itemsets to mine both the positive and negative association rules from frequent and infrequent itemsets. It minimizes I/O overhead by scanning the database only once. The performance study shows that the proposed method is an efficient than the previously proposed method.

Keywords—data mining; association rule; frequent itemset; positive association rule; negative association rule

I. INTRODUCTION

Association rule mining is a data mining task that discovers associations among items in a transactional database. Association rules have been extensively studied in the literature since Agrawal et al. first introduced it in [1, 2]. A typical example of association rule mining application is the market basket analysis. Much effort has been devoted and algorithms proposed for efficiently discovering association rules [2, 3, 4, 5, 6,16].

Association rules provide a convenient and effective way to identify and represent certain dependencies between attributes in a database[1]. Association rule mining includes positive and negative association rule mining[9,11,12,17]. In the traditional approach to find association rules, one merely thinks in terms of positive association rules: especially when determining the degree of support and confidence[2].

The study of the negative association rule is a new active research field in recent years. It still focuses on the transactional databases, and has made a number of important research results[7,8,10,18]. Brin M. et al. referred to the relevance of the two sets firstly [1]. Savasere O. et al. described a strong negative association rules model [2]. Xindong Wu et al. proposed a PR model [3], and gave an algorithm that can mine positive and negative association rules simultaneously.

Ling Zhou et al. [14] and Junfeng Ding et al. [15] proposed methods to mine association rules from infrequent itemsets. Mining positive association rules from frequent itemsets and negative association rules from infrequent itemsets with some interesting measures are described in [9]. Honglei Zhu et al.

[12] mine both positive and negative association rules from frequent and infrequent itemsets respectively with the differential support and confidence.

In this paper we examine the problem of mining positive and negative association rules from frequent and infrequent itemsets.

The rest of this paper is organized as follows. Section 2 briefly presents the relevant concepts and definitions. In Section 3, the existing strategies for mining both positive and negative association rules are reviewed. The proposed algorithm is presented in Section 4. Section 5, illustrate the computational results. The concluding remarks are finally made in Section 6.

II. CONCEPTS AND DEFINITIONS

Let $I = \{i_1, i_2, i_3, \dots, i_n\}$ be a finite set of items and DB be a transactional database. Support of the itemset $X \subseteq I$ is [1]:

$$Supp(X) = \frac{\text{No. of transactions contains } X}{\text{Total No. of Transactions in DB}} \quad (1)$$

Definition 1: If the support of itemset X is greater than or equal to user defined minimum support (ms) threshold, X is called frequent itemset otherwise infrequent Itemset [5].

A. Positive Association Rule:

A (positive) association rule is of the form: $X \Rightarrow Y$, with $X, Y \subseteq I, X \cap Y = \emptyset$ [1][9]. Support and confidence of $X \Rightarrow Y$ are defined as[2]:

$$Supp(X \Rightarrow Y) = Supp(X \cup Y) \quad (2)$$

$$Conf(X \Rightarrow Y) = \frac{Supp(X \cup Y)}{Supp(X)} \quad (3)$$

An interesting positive association rule has support and confidence greater than user given thresholds *minimum support* (ms) and *minimum confidence* (mc) respectively.

B. Negative Association Rule:

A negative association rule is an implication of the form $X \Rightarrow \neg Y$ (or $\neg X \Rightarrow Y$), where $X \subseteq I, Y \subseteq I$, and $X \cap Y = \emptyset$ [9]. The rule $\neg X \Rightarrow \neg Y$ is equivalent to a positive association rule in the form of $Y \Rightarrow X$. From [12], we extracted the following formulas:

$$Supp(\neg X) = 1 - Supp(X) \quad (4)$$

$$Supp(\neg X \cup Y) = Supp(X) - Supp(XUY) \quad (5)$$

$$Supp(X \cup \neg Y) = Supp(Y) - Supp(XUY) \quad (6)$$

$$Supp(\neg X \cup \neg Y) = Supp(X) - Supp(Y) + Supp(XUY) \quad (7)$$

$$Conf(X \Rightarrow \neg Y) = \frac{Supp(X) - Supp(X \cup Y)}{Supp(X)} \quad (8)$$

$$Conf(\neg X \Rightarrow Y) = \frac{Supp(X) - Supp(X \cup Y)}{1 - Supp(X)} \quad (9)$$

$$Conf(\neg X \Rightarrow \neg Y) = \frac{1 - Supp(X) - Supp(Y) + Supp(X \cup Y)}{1 - Supp(X)} \quad (10)$$

$$Corr(X, Y) = \frac{Supp(X \cup Y)}{Supp(X)Supp(Y)} \quad (11)$$

Definition 2: The bit vector (BV) of an item i is in form $BV = (b_1, b_2, b_3 \dots b_m)$, $b_k \in [0, 1]$, If $i \in T_k$, and then $b_k=1$, otherwise $b_k=0$, where $k=1, 2 \dots m$. The size of BV is equal to the number of items in i , and the support of an itemset is equal to the number 1s in the bit vector.

III. RELATED WORK

Several algorithms have been proposed for mining association rules, negative association rules. But only few algorithms have been proposed for mining both positive and negative association rules concurrently. Wu et al [9] presented an Apriori-based framework for mining both positive and negative ARs based on rule dependency measures and an additional threshold *minimum interest (mi)*. A rule $X \Rightarrow \neg Y$ (or $\neg X \Rightarrow Y$) is only considered as a valid negative AR, if both X and Y are frequent and the *interest (X, $\neg Y$)* $\geq mi$ (or *interest ($\neg X, Y$)* $\geq mi$).

The most common frame-work in the association rule generation is the "Support-Confidence" one. In [11], authors considered another frame-work called correlation analysis that adds to the support-confidence. They combined the two phases (mining frequent itemsets and generating strong association rules) and generated the relevant rules while analyzing the correlations within each candidate itemset. Their algorithm avoids evaluating item combinations redundantly. For each candidate itemset, they computed all possible combinations of items are outputted to analyze their correlations. At the end, they keep only those rules generated from item combinations with strong correlation. If the correlation is positive, a positive rule is discovered. If the correlation is negative, two negative rules are discovered.

Honglei Zhu et al. [12] proposed for the purpose of simultaneously generating positive ARs from frequent itemsets and negative ARs from infrequent itemsets with differential minimum support and differential minimum confidence. An innovative approach has proposed in [13]. In this, authors dividing the itemset space into four parts for mining positive and negative association rules. In [14], the authors proposed a method to mine association rules from infrequent itemsets.

IV. PROBLEM DESCRIPTION AND PROPOSED METHOD

Most of the methods proposed for mining positive and negative association rules, maintains both frequent and infrequent itemsets and hence suffer from scalability. To maintain the execution time within user's expectations, it is necessary to design an efficient approach to mine both positive and negative association rules.

Problem Statement: Given a database of transactions DB and user-defined minimum support (ms) value, minimum confidence (mc) values, the problem is to extract all interesting positive and negative Boolean association rules.

We propose Frequent and Infrequent Itemset tree (FII-tree) as a data structure to hold requisite itemsets and also a method to extract all the positive and negative association rules.

The proposed process consists of two phases.

Phase 1: In this phase, we construct an FII-tree which can hold all frequent and infrequent itemsets. The root of the tree, labeled with "null". Each non-root node in a tree has generic form $\langle I, c \rangle$, where I is an itemset and c is the support of I . Infrequent 1-itemsets are ignored. However, all frequent 1-itemsets are placed in level 1 in lexicographic order, all frequent and infrequent 2-itemsets are in level 2, and so on, all frequent and infrequent k -itemsets are placed in level k . The highest level of the FII-tree is L , where L is equal to the number of frequent 1-itemsets. For FII-tree, two indices, one is *FreqIndex* for all frequent itemset lists and the second one is *InfreqIndex* for all infrequent itemset lists are maintained separately for easy accessibility of frequent and infrequent itemsets. The step by step process of Creation of FII-tree is given below:

- 1) Scan the database DB once, and store in an item based vectors BV, then find frequent 1-itemsets based on definition2.
- 2) Insert frequent 1-items one by one in the tree, and assign to an index *FreqIndex*
- 3) Generate candidate k -itemsets C_k ($k=2, 3 \dots$) from frequent $(k-1)$ -itemsets. For each item X in a candidate k -itemsets C_k
 - a) If $supp(X) \geq min_supp$ and $Corr(X) > 1$ then assign X to frequent k -itemset list (FL_k) otherwise assign X to infrequent k -itemset list (IFL_k). Calculate support of X by performing bitwise AND operation between bit vectors (BV) (if $x_1, x_2 \in X$ then the $supp(X) = x_1 \wedge x_2$).
 - b) Assign the FL_k and IFL_k to $FreqIndex_k$ and $InfreqIndex_k$ respectively (where $k=2, 3, 4 \dots$).
- 4) Repeat step3 until to generate largest itemset.

Phase 2: Mining Positive and Negative Association Rules:

- 5) Read frequent k -itemsets ($k=1, 2, 3 \dots$) from FII-tree and generate Positive Association Rules based on given threshold values.
- 6) Read frequent k -itemsets ($k=1, 2, 3 \dots$) from FII-tree and generate Positive Association Rules based on given threshold values.

Example: Transactional Database (DB):

TABLE I. TRANSACTIONAL DATABASE

TID	Items	TID	Items
1	A B C D E	6	A B C
2	A B C	7	A B C
3	A B D	8	B C E
4	B C D	9	B C D
5	C D E	10	C D

The Transactional Database (DB) consists of 5 items and 10 transactions. The Item based bit vectors are:

BV1 = 1110011000
 BV2 = 1111011110
 BV3 = 1101111111
 BV4 = 1011100011
 BV5 = 1000100100

Based on definition2 the support count of an itemset is the number of 1s in each bit vector. For example $Supp(A)=5$, $Supp(B)=8$, $Supp(C)=9$, $Supp(D)=6$ and $Supp(E)=3$. The minimum support (ms) is 5 (50%), then the frequent 1-itemsets are: {A, B, C, D}.

Insert frequent 1-itmsets in FII-tree. The candidate 2-itemsets are AB, AC, AD, BC, BD and CD. Next perform the bitwise AND (^) operation between each pair of frequent itemset.

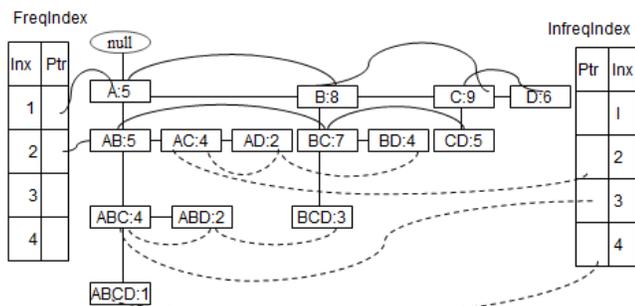


Fig. 1. FII-tree

For example, $R=BV_1 \wedge BV_2 = (1110011000) \wedge (1111011110) = 1110011000$, the number of 1s in the resultant bit vector R is 5, so the $Supp(AB)=5$, $Supp(AC)=4$, $Supp(AD)=2$, $Supp(BC)=7$, $Supp(BD)=4$ and $Supp(CD)=5$. Thus the frequent 2-itemsets are {AB, BC, and CD} as their support counts are not less than 50%, and the infrequent 2-itemsets are {AC, AD, and BD} as their support counts are less than 50%.

Insert frequent and infrequent 2-itemsets in FII-tree. For three itemsets perform the bitwise AND operation between three bit vectors as given below: $R = (BV_1 \wedge BV_2 \wedge BV_3) = (1110011000) \wedge (1111011110) \wedge (1101111111) = 1100011000$, the number of 1's in the resultant bit vector is 4, so the $Supp(ABC)=4$, $Supp(ABD)=2$, and $Supp(BCD)=3$. Thus the infrequent 3-itemsets are {ABC, ABD, and BCD} as their supports are less than 50% and there are no frequent 3-itemsets

and 4-itemsets, $Supp(ABCD)=01$, this is the largest itemset for the given transactional database (DB), then the algorithm stop processing. The FII-tree for the above example is shown in fig1.

In the Fig. 1, solid straight lines are the links between the items in the same levels and links between different levels of the tree. All arcs are the links between frequent itemsets and the dashed arcs are the links between infrequent items. Different levels of frequent items are linked to frequent index called *FreqIndex* and infrequent items are linked to infrequent index called *InfreqIndex*.

Algorithm 1:

Input: MinSup, MinConf.
 Output: FrequentItemset, InfrequentItemset, PositiveAR and NegativeAR.

Phase 1:

- 1) Scan database DB once and find frequent 1-itemset, $freq_1$
- 2) Insert $freq_1$ in FII-tree: *Insert_FII-tree ()* //calls algorithm2
- 3) Generate candidate k itemset C_k ($k=2, 3, \dots$) from $freq_1$ and Insert $freq_k$ in FII-tree: *Insert_FII-tree()*

Phase 2:

- // *FreqIndex* frequent itemset index
- // *InfreqIndex* infrequent itemset index.
- 4) Generate Positive Association Rules from $FreqIndex_k$ ($k=1, 2, \dots$)
- If $Corr(X, Y) > 1$ then
- If $Conf(X, Y) > MinConf$ then
- $PositiveAR \leftarrow PositiveAR \cup \{X \rightarrow Y\}$
- 5) Generate Negative Association Rules from $InfreqIndex_k$ ($k=2, 3, \dots$)
- For each infrequent item i in *InfreqIndex*
- For each expression X Y, $X \cup Y = i$ and $X \cap Y = \emptyset$
- {
- Step A:
- If $Corr(X, Y) < 1$ then
- If $Supp(X, \neg Y) > MinSupp$ and $Conf(X, \neg Y) > minconf$ then
- $NegativeAR \leftarrow NegativeAR \cup (X \rightarrow \neg Y)$
- Step B:
- If $Corr(X, Y) < 1$ then
- If $Supp(\neg X, Y) > MinSupp$ and
- $Conf(\neg X, Y) > minconf$ then
- $NegativeAR \leftarrow NegativeAR \cup (\neg X \rightarrow Y)$
- }
- 6) $AR \leftarrow PositiveAR \cup NegativeAR.$

Algorithm 2:

Input: MinSup, Candidate Itemset C_k , k
 Output: Frequent Inferquent Index Tree (FII-tree)

Inseret_FII-tree ()
 { // k=1 for frequent 1-itemsets
 //k= 2, 3... for frequent 2, 3..
 1: if $k==1$ then
 {

```

For all items in  $freq_1$ 
  Insert  $i, i \in freq_1$  in a node then add node to the tree.
   $FreqIndex_1 \leftarrow FreqL_1$ 
  //FreqL1 is frequent itemset list1
}
2: else{
  for all items in  $C_k$ 
  {
    //FreqLk is k frequent itemset list
    //InFreqLk is k infrequent itemset list
    Insert  $i, i \in C_k$  in a node then add node to the tree.
    if  $supp(i) > MinSup$  then
       $FreqL_k \leftarrow i$ 
    else  $InFreqL_k \leftarrow i$ 
     $FreqIndex_k \leftarrow FreqL_k$ 
     $InfreqIndex_k \leftarrow InFreqL_k$ 
  }
}

```

V. EXPERIMENTAL RESULTS

We conduct experiments on a different transaction size and differing number of transactions in a database to compare our approach with the PNAR [12]. The execution time with different minimum supports for the dataset T50I30D200K is shown in the Fig. 2.

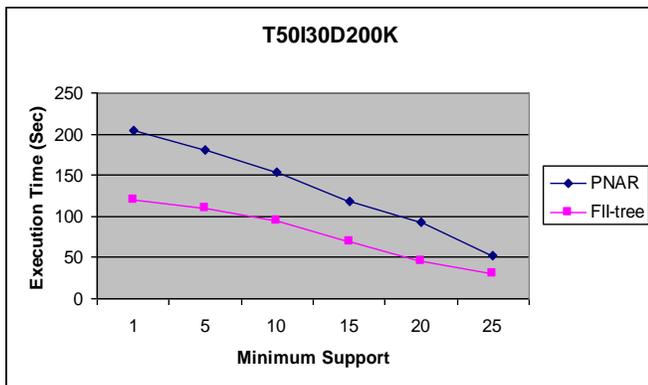


Fig. 2. Execution time with different minimum supports

The execution time with different dataset sizes (number of transactions) for the fixed minimum support 0.5 is shown in Fig. 3. It can be observed that both the methods generate equal number of positive and negative association rules, but the proposed approach reduce the execution time over the existing method.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have designed a new tree structure to store both the frequent and infrequent itemsets for mining both positive and negative association rules. In the proposed method the database is scanned only once for mining positive and negative association rules, so it reduces the number of I/O operations. Another flexibility of the structure is, if any new frequent 1-itemsets are mined by reducing the user threshold value (minimum support (ms)), the proposed method allows appending of new items to the tree without reconstructing from scratch.

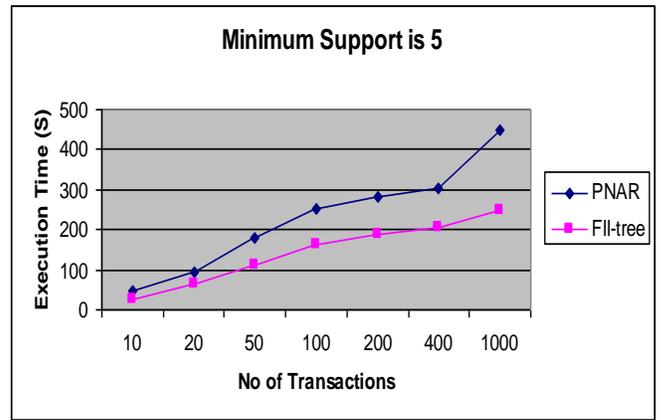


Fig. 3. Execution time with different dataset size

Recently, there have been some interesting studies about mining frequent patterns in databases which allow adding new data or deleting old data. The maintenance of the already mined frequent patterns when updating databases is an interesting topic for future research.

REFERENCES

- [1] R. Agrawal, T. Imielinski, A. Swami, "Mining association rules between sets of items in large databases", In Proceedings of ACM SIGMOD International Conference on Management of Data, New York, May 1993, pp. 207-216.
- [2] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules", In Proceedings of 20th International Conference on VLDB, Chile, May 1993, pp. 207-216.
- [3] J. Han and Y. Fu, "Mining multiple-level association rules in large databases", IEEE Trans. on Knowledge and Data Engineering, Vol. 11, No 5, September 1999, pp. 798-805.
- [4] M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen and A. I. Verkamo, "Finding interesting rules from large sets of discovered association rules", In proceedings of 3rd International Conference on Information and Knowledge Management, Maryland, November 1994, pp 401-407.
- [5] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables", ACM SIGMOD Record, Vol. 25, No. 2, June 1996, pp. 1-12.
- [6] Brin, S., Motwani, R., and Silverstein, C., "Beyond Market: Generalizing association rules to correlations", In Proceedings of ACM SIGMOD International Conference on Management of Data, New York, May 1997, pp. 265-276.
- [7] Savasere, A., Omiecinski, E., and Navathe, S., "Mining for strong negative associations in a large database of customer transactions", In Proceedings of 14th International Conference on Data Engineering, Orlando, February 1998, pp. 494-502.
- [8] X. Yuan, B.P. Buckles, Z. Yuan and J. Zhang, "Mining negative association rules", In Proceedings of 7th International Symposium on Computers and Communication, Italy, July 2002, pp. 623-629.
- [9] X. Wu, C. Zhang and S. Zhang, "Efficient mining of both positive and negative association rules", ACM Trans. on Information Systems, Vol. 22 No. 3, July 2004, pp. 381-405.
- [10] S. Wang, H. Song, and Y. Lu, "Study on negative association rules", Trans. of Beijing Institute of Technology, Vol. 24, No. 11, November 2004, pp. 978-981.
- [11] M.L. Antonie and O.R. Za'iane, "Mining positive and negative association rules: an approach for confined rules", In Proceedings of International Conference on Principles and Practice of Knowledge Discovery in Databases, Italy, September 2004, pp. 27-38.
- [12] Honglei Zhu, Zhigang Xu, "An Effective Algorithm for Mining Positive and Negative Association Rules", In Proceedings of International

- Conference on Computer Science and Software Engineering, Hubei, December 2008, pp. 455-458.
- [13] Chris Cornelis, peng Yan, Xing Zhang, Guoqing Chen, "Mining Positive and Negative Association Rules from Large Databases", In Proceedings of IEEE Conference on Cybernetics and Intelligent Systems, Bangkok, June 2006, pp. 1-6.
- [14] Ling Zhou, Stephen Yau, "Efficient association rule mining among both frequent and infrequent items", *Computers and Mathematics with Applications*, Vol. 54, No.6, September 2007, pp. 737-749.
- [15] Junfeng Ding, Stephen S.T. Yau, "TCOM, an innovative data structure for mining association rules among infrequent items", *Computers and Mathematics with Applications*, Vol. 57, No. 2, January 2009, pp. 290-301.
- [16] Tamanna Siddiqui, M Afshar Alam and Sapna Jain, "Discovery of scalable association rules from large set of multidimensional Quantitative datasets", *Journal of Advances in Information Technology*, Vol. 3, No. 1, February 2012, pp. 69-76.
- [17] Nikky Suryawanshi, Susheel Jain and Anurag Jain, "A review of negative and positive association rule mining with multiple constraint and correlation factor", *International Journal of Emerging Technology and Advanced Engineering*, Vol. 2, No. 12, December 2012, pp. 778-781.
- [18] Li-Min, Shu-Jing Lin and Don-Lin Yang, "Efficient mining of generalized negative association rules", In proceedings of IEEE International Conference on Granular Computing, USA, August 2010, pp. 471-476.

Link-Budget Design and Analysis showing Impulse-based UWB Performance Trade-Off flexibility as Integrator Solution for Different Wireless Short-Range Infrastructures

M.S Jawad, Othman A. R, Z. Zakaria and A.A.Isa.

Faculty of Electronics and Computer Engineering
(FKEKK), Univeristi Teknikal Malaysia Melaka (UTeM),
Hang Tuah Jaya, Durrian Tunggal, Melaka, Malaysia,
76100.

W. Ismail

Kampus Kejuruteraan, Universiti Sains Malaysia, Seri
Ampangan, 14300 Nibong Tebal, Seberang Perai Selatan,
Pulau, Pinang, Malaysia.

Abstract—Future wireless indoor scenarios are expected to be complex requiring wireless nodes to adaptive responding to dynamic changes according to channel conditions. Interacting with neighboring nodes to achieve optimized performance in term of data-rates, distance and BER performance are the main concerns of designing future wireless solutions. IR-UWB came to the picture as the missing Puzzle to achieve these requirements and gluing the different wireless indoor existing infrastructures in global platform solutions. This paper shows the flexibility of IR-UWB in signal design at the physical layer level as cross-layer architecture of optimized performance. A detailed performance analysis presented in this paper as a mathematical model of the proposed wireless solution and described in proposed link budget design template. The performance evaluation is carried-on to show the proposed system as a good candidate in different wireless scenarios for different data-rate requirements, distance for specific requirement BER. Simulations results and well as experimental statistical analysis of the received signal under different channel models and conditions are carried-out as a proof of concept of the proposed system.

Keywords—Ultra Wideband; Time Hopping-pulse position Modulation; Radio Frequency Identification; Wireless Sensors Networks; RAKE Receiver; Bit Error rate

I. INTRODUCTION

The impulse based IR-UWB technology has attracted a great deal of research recently, due to its enormous advantages which can be utilized for wireless indoor applications and infrastructure such as RFID and WSN networks. Thanks to its signal' unique physical characteristics which can be described as a very short a train of pulses in ultra-wide bandwidth, making this type of signal a very low duty cycle in comparison with classical narrowband signal. This will add a good advantage of IR-UWB as reduction in power consumption, since in these discontinuous emitted pulses the emitter will consume much less.[1] This type of continuous emission technique can be used with a large bandwidth to achieve a high data [2]. The IR-UWB as a discontinuous emission technique in a wide bandwidth can be seen as a noise-like signal difficult for any intruder to intercept or detect making this signal secure at the media with good performances as the

ability to co-existence with other narrowband technologies and resist to interface. Also, a high resolution for localization can be achieved using this technique of low-duty cycle pulses in wide bandwidth allowing the tracking of objects with less than 1 cm accuracy. As this technology considered carrier less technology with no need for carrier signal as modulation done by altering the pulses positions, amplitude or polarity so the IR-UWB emitter can be designed without RF stage. With no VCO (Voltage Control Oscillator), mixer and other RF components this mostly digital architecture will consume less power and allow simple receiver structure [3-4]. All these advantages can be utilized for wireless indoor sensing and identification applications combining the functionalities of detecting, identifying and tracking the objects in one global platform. Since RFID and WSNs communication systems share in common some system design constrains such as the limitations of power consumption and hardware complexity[5]. IR-UWB can be a good integrator for both of them. Fortunately many IR-UWB signal parameters can be adjusted and tuned allowing what's called a reconfigurable IR-UWB radio[4]. A link budget analysis model is proposed to show system parameters selection and then the radio performance to understand the trade-off flexibility for different data- rates, ranges (between the transmitter and receiver) and BER performance. The proposed model can be utilized for different high-rate, low-rate, and short to medium wireless indoor scenarios. The structure of the paper is as follows. Section2 consider a literature review studies to show the utilization of IR-UWB for different high-rate and low-rate WPAN with the focus on wireless sensing and identification. Section 3 models mathematically the IR-UWB transmission chain to fully understand the signal generation parameters. Section 4 shows a suitable IR-UWB receiver structure and its performance in AWGN channel model, then in section 5 the link budget with different signal physical parameters is modeled and analyzed for a given IR-UWB power compliant with the FCC rules [6] and with some simplified hypotheses to show the maximum distances of propagation when a pre-determined probability of error must be guaranteed for a given data-rate at the receiver. Finally, the results remarks are concluded with suggestions for future researches.

II. LITERATURE STUDIES

Many researches in literature proposing the utilization of IR-UWB in WPAN (Wireless Personal Area Network) communication for high data-rate, low data-rate and from short to medium ranges applications [7]. (Pezzin, M., B. Denis, et al) they propose utilization of IR-UWB in low complexity and low data rate platform for location and tracking services [8]. Achieving security by making the wireless signal secure physically at the medium rather than achieving security by securing the contents of data using some cryptography algorithm can be utilized for RFID system using IR-UWB technique and this security approach can reduce the hardware complexity significantly especially for passive RFID tags by eliminating the logic gates (at the hardware level) needed to cipher or decipher blocks of data and this has been proposed recently in some researches [9-10]. Also Time Hopping multi-Pulse Position Modulation (UWB-TH-MPPM) is been proposed for high data rate wireless short ranges applications[2].

Many proposals been introduced in literature with the aim of integrating both RFID and WSN networks to combine their capabilities of detecting the objects of interest and tracking these objects while sensing, collecting and processing the data [11-14] and some researches utilizing the great advantages of IR-UWB as integrator for both RFID and WSN systems [15]. IR-UWB technology also has the potentials to support identifications, Location, sensing and connectivity in ubiquitous computing environments and in cognitive radio scenarios since this technology consider good candidate satisfying many requirements of cognitive radio[16-17]. The IR-UWB reconfiguration flexibility in many aspects is been studied with the link budget analysis in [4]. As a conclusion from all these previous studies, it clearly that the IR-UWB unique characteristics with the reconfiguration capabilities of IR-UWB can be utilized for many wireless indoor applications to show the trade-off flexibility for different data-rates, ranges and BER requirements and the best way to analyze this flexibility is by proposing link budget analysis model which this paper investigate.

III. IR-UWB MODEL AND PARAMETERS

The time hopping as spectrum spreading technique allowing multiple access performance with the M-ary Pulse Position modulation is considered due to its simplicity and suitability for simple RFID tags or wireless nodes sensors. The signal generated with time hopping (TH) with M-ary with pulse position modulation for K-th user can be mathematically modeled as

$$S^K(t) = \sum_{j=-\infty}^{\infty} A_{d[j/N_s]}^{(k)} p(t - T_f - C_j^{(k)} - \delta_d(k)_{[j/N_s]}) \dots \dots \dots (1) [18]$$

Understanding the signal parameters with their typical values is important to show the flexibility and reconfiguration capabilities of IR-UWB at the physical layer. These parameters can be described as the following:

P(t) represents the pulse shaping which is the second derivative of Gaussian pulse with pulse width T_p , T_f is the frame time and each frame is divided into N_h time slots with

duration time $T_c \cdot C_j^{(k)}$ is the pulse shaping patterns, which is pseudorandom numbers with period T_c , it represent the time hopping sequence for K_{Th} users, here additional shift is required to avoid catastrophic collisions caused by the multiple access interference (MAI). The d parameter is the sequence for the data stream generated by the K_{Th} user after channel coding. δ is the shift introduced for pulse position modulation utilized by M-ary PPM. If it is assumed that the signal amplitude $A^{(v)} = 1$, for the M-ary PPM signal, then equation (1) can be written as $S^{(v)}(t) = \sum_{j=-\infty}^{\infty} p(t - jT_f - c_j^{(k)}T_c - \delta d_{[j/N_s]}^{(k)})$. The following figures show the simulation of the generated signal then the PPM modulation for transmitting 1010 code.

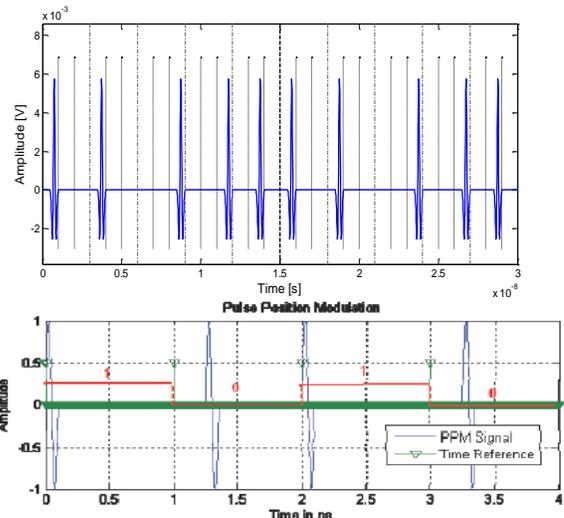


Fig. 1. a:Simulation of TH-PPM signal generation.

Fig. 1. b: TH-PPM signal coding

A. IR-UWB Receiver structure

Optimum receiver structure in AWGN channel

For M-ary PPM, assuming the time shift ξ introduced by PPM modulation is larger than the pulse duration T_m , as shown in the Figure-3- where the structure of the optimum receiver in this case, shows that the output of the signal correlators will be considered as decision variable and can be expressed as the following :-

$$Z_0 = \alpha S_{m0} + n_0$$

$$\vdots$$

$$Z_{M-1} = \alpha S_{m(M-1)} + n_{M-1}$$

Where $S_{mK} = \sqrt{E_{TX}} \int_0^{T_s} P_0(t - m\xi) P_0(t - k\xi) dt$

$P_0(t)$ is the energy- normalized waveform of the basic pulse, E_{TX} is the transmitted energy per pulse, ξ shift introduced by PPM modulation. [8] [7]

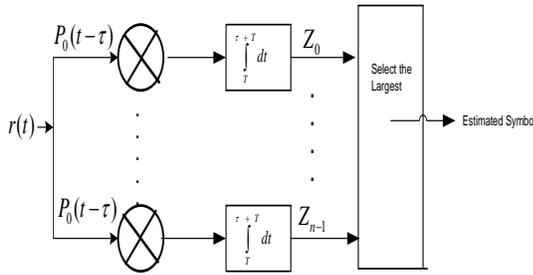


Fig. 2. TH-PPM receiver structure for AWGN channel

When M possible waveforms are generated, the average error probability on symbol P_{re} is the probability of misdetection one of the symbols, the error occurs at the receiver side if at least one $M-1$ output Z_K is larger than Z_0 , the probability of this even to occur can be expressed as the following :-

$$Pr_e = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left[1 - \left(\frac{1}{2\pi} \int_{-\infty}^y e^{-\frac{x^2}{2}} dx \right)^{M-1} \right] e^{\left[\frac{-1}{2} \left(y - \sqrt{\frac{2E_{Rx}}{N_0}} \right)^2 \right]} dy \dots\dots\dots (2) [18]$$

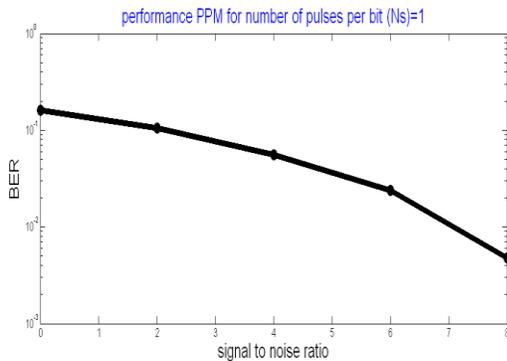


Fig. 3. TH-PPM receiver performance

B. Receiver Structure in multipath fading environments

To optimize correlation in multipath environment, additional correlators should be associated for different replicas of the same transmitted waveform, this scheme what is called rake receiver was invented by (Prince and Green, 1958)

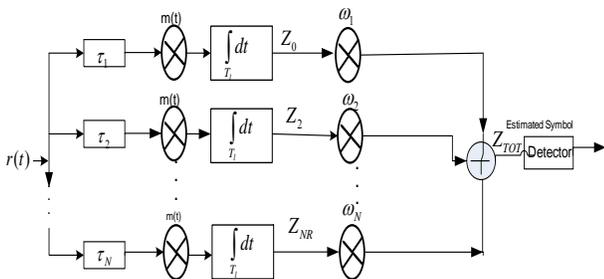


Fig. 4. RAKE receiver structure in Multi-path fading channel

As shown in Figure 4, the RAKE receiver consists of parallel bank of N_R correlators, where each correlator is associated to one of the different replicas of the transmitted pulse, so that the correlator mask $m_j(t)$ on the time j -th branch of the RAKE is aligned in time with the j -th delayed replica of the transmitted pulse $m_j(t) = m(t - \tau_j)$, where $m(t)$ is the correlation mask, and τ_j is the propagation delay of the j -th path, at the end the output of the correlators feeding the combiner. The weighting factors $(\omega_1, \dots, \omega_{N_r})$, can be combined depending on the diversity method implemented in the receiver. In MRC case, to output of each branch is multiplied by the weighting factor, which is proportional to the signal amplitude on that branch

IV. LINK BUDGET ANALYSIS

In reality UWB radio signals, coexist with other radio signals which causes interferences with other UWB or narrowband communication systems .it's also subjected to channel conditions changes especially in severe multipath environments. The problem is for any UWB system optimization should be within the regulated values of intrinsic UWB principle.

As such, UWB should be flexible to adapted to environments, whether this refer to channel condition changes or interference patterns while maintaining compatibility with the regulations on emitted UWB signal . The UWB signal format offers interesting and appealing properties by high numbers of tunable parameters and this is especially due to the impulsive nature of the pulses. Signal characteristics can be tuned efficiently by playing with a variety of parameters. These parameters include the transmission factor such as the number of pulses representing one bit, coding factors such as periodicity and cardinality of TH. Modulation factors related to the PPM shift, shaping factors related to specific pulse shape.

A. Adjusting emitted radiation

The power limitations of UWB signal set by FFC emission masks determine the maximum allowed transmitted power. In order to understand how to read an emission mask, it is important at the beginning to define the effective isotropic radiated power (EIRP) for a given range of operating frequencies. EIRP can be understood as the maximum power that the transmitter can transfer to the transmitter antenna, so it can be defined as the product of the available power at the transmitter W_{TX} , expressed in Watts and the gain of the transmitter antenna G_{AT} .

$$EIRP = W_{TX} G_{AT} \dots\dots\dots (3)$$

The condition here is the maximum power transfer from the output impedance of the transmitter Z_{TX} to the input impedance of the antenna Z_{AT} is verified that is, $Z_{AT} = Z_{TX}$. EIRP is measured in dBm that is $10 \log_{10} EIRP$ mWatts.

So in brief, emission mask impose limits on the PSD of the emitted signal and can be expressed as dBm/Hz or dBm/MHz.

commonly provided in practice, the emission mask expressed in term of power value for a given frequencies which indicate the maximum allowed radiated power with measured bandwidth (mb) around center frequency f_c .

B. IR-UWB Link Budget Analysis over AWGN Channel

Given allowed power and under simplified hypotheses, we can evaluate the maximum distance over which the propagation can reach, with a predetermined level of error should be guaranteed at the receiver for a given data rate.

The analysis start by observing the received energy E over a finite time interval, allowing the decision at the receiver and in case of propagation over AWGN channel, the received energy is based on the signal term E_r and the thermal noise E_{noise} , from the signal to noise ratio SNR can be defined as:

$$SNR = \frac{E_r}{E_{noise}} \geq M \cdot SNR_{spec} \dots\dots (4)$$

Where SNR_{spec} the required SNR, M is the system margin
So the required energy at the receiver can be expressed:

$$E_r = M \cdot SNR_{spec} \cdot E_{noise} \dots\dots (5)$$

The thermal noise energy can be defined as:

$$E_{noise} = \frac{1}{2} N_o = \frac{1}{2} K Temp_s \dots\dots (6)$$

Where N_o is Single-sided thermal noise spectral power density, K is the Boltzman constant and $Temp_s$ the spot noise temperature.

$$Temp_s = Temp_A + (F(f) - 1)Temp_o \dots\dots (7)$$

Where $Temp_A$ is the receiving antenna temperature, $F(f)$ is the spot noise figure and $Temp_o$ is the standard temperature (290°k)

E_r Can be defined as received signal power P_r over a finite time interval T_b , so from equation (2) E_r can be expressed as
 $E_r = M \cdot SNR_{spec} \cdot \frac{1}{2} K (Temp_A + (F(f) - 1)Temp_o) \dots\dots (8)$

Which is equivalent to

$$P_r T_b = M \cdot SNR_{spec} \cdot \frac{1}{2} K (Temp_A + (F(f) - 1)Temp_o) \dots\dots (9)$$

The receiver power $P_r = 2 \int_{f_L}^{f_H} \frac{P_s(f)}{A_{FS}} df = 2 \int_{f_L}^{f_H} \frac{P_s(f)}{\frac{(4\pi)^2 D^2 f^2}{G_T G_R C^2}} df \dots\dots (10)$

$P_s(f)$ Is the double sided power spectral density, $A_{FS}(f)$ is the free space attenuation.

So equation (6) can be expressed as:

$$2T_b \int_{f_L}^{f_H} \frac{P_s(f)}{\frac{(4\pi)^2 D^2 f^2}{G_T G_R C^2}} df =$$

$$M \cdot SNR_{spec} \cdot \frac{1}{2} K (Temp_A + (F(f) - 1)Temp_o) \dots\dots (11)$$

From equation (8) we can derive the squared value of the maximum distance D that can be covered by transmission.

$$D^2 = \frac{\frac{G_T G_R C^2}{(4\pi)^2} 2T_b \int_{f_L}^{f_H} \frac{P_s(f)}{f^2} df}{M_S SNR_{Spec} \frac{1}{2} K (Temp_A + (F(f) - 1)Temp_o)} \dots\dots (12)$$

From equation (9) we can derive the maximum distance of propagation of an UWB point to point link given P_{smin} and the given target SNR_{Spec}

For the sake of simplicity it can be assumed that $Temp_A = Temp_o$, in this assumption, the noise temperature at the receiving antenna is considered the standard temperature. This assumption is practical for terrestrial links. Equation (9) becomes:

$$D^2 = \frac{\frac{G_T G_R C^2}{(4\pi)^2} 2T_b \int_{f_L}^{f_H} \frac{P_s(f)}{f^2} df}{M_S SNR_{Spec} \frac{1}{2} K Temp_o} \dots\dots (13)$$

In the worst case for $P_s(f)$, it can be substituted by its lower case value P_{min} within its bandwidth, so equation (10) can be then:

$$D^2 = \frac{\frac{G_T G_R C^2}{(4\pi)^2} 2T_b P_{SMin} \int_{f_L}^{f_H} \frac{P_s(f)}{f^2} df}{M_S SNR_{Spec} \frac{1}{2} K Temp_o} \dots\dots (14)$$

And it can be expressed as in equation (12):

$$D = \sqrt{\frac{\frac{G_T G_R C^2}{(4\pi)^2} 2T_b P * S_{min} \frac{f_H - f_L}{f_H f_L}}{M_S SNR_{Spec} \frac{1}{2} K Temp_o}} \dots\dots (15)$$

Usually the term, probability of symbol error P_{re} is commonly used for system specifications, rather than SNR. Modulation scheme effecting the relation between SNR and P_{re} and that can be easily expressed in term of noise in AWGN. In case of optimum AWGN receiver, the performance can be evaluated for various modulation types as the following:

Given a target P_{re} , we want to evaluate the required SNR_{Spec} , we assume the system margin is zero for ($M_s = 0$) for the sake of simplicity. referring to some resources of digital communication dealing with these relations for different modulations schemes (lee and Messecrschmitt (1994), Proakis (1995) and summary in Guvence and Arslan (2003)), we can summarize the link budget analysis for M-PAM and M-PPM as the following:

First: for M-ary PAM

$$Pr_e = \left\{ 1 - \frac{1}{M} \right\} erf c(y) \dots\dots (16)$$

$$\text{Where } y^2 = \frac{SNR_{Spec}}{\frac{2}{3}(M^2 - 1)} = \frac{P_r T_b \text{Log}_2 M}{\frac{1}{2} k F Temp_o \frac{2}{3}(M^2 - 1)} = \frac{E_b}{N_o} \frac{3 \text{log}_2 M}{No(M^2 - 1)} \dots\dots (17)$$

With SNR is sufficiently high ($\frac{E_b}{N_o} > 4.43 \text{ dB}$) then the upper bound for M-ary orthogonal PPM for $\frac{E_b}{N_o} > 4.43 \text{ dB}$ can be expressed as :

$$Pr_e < e^{-\text{log} M (\frac{E_b}{N_o} - 2 \text{log}_e 2) / 2} \dots\dots (18)$$

Equation (15) is valid only for high SNR values and for SNR vales lower than 4.43 dB, the lower bound can be

$$\text{expressed as } Pr_e < 2e^{-\text{log}_2 M (\sqrt{\frac{E_b}{N_o}} - \sqrt{\text{log}_e 2})^2} \dots\dots (19)$$

Figure (5) evaluate the Pr_e for an M-ary PAM signal and the upper bound of the Pr_e for M-ary PPM .

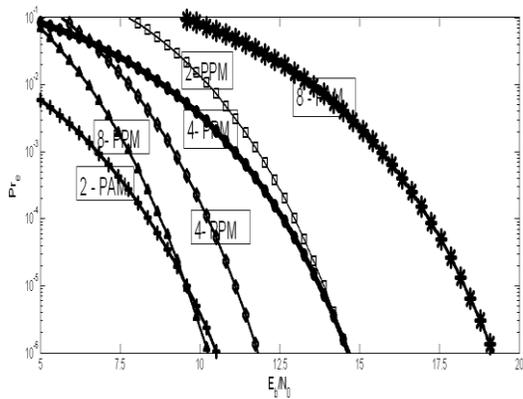


Fig. 5. comparison for probabilities of symbol error for M-PAM and PPM according to equations 18 & 19

The notice here is that the performance improves as M increases for the case of M-PPM signal, whereas in case of M-PAM the opposite is true.

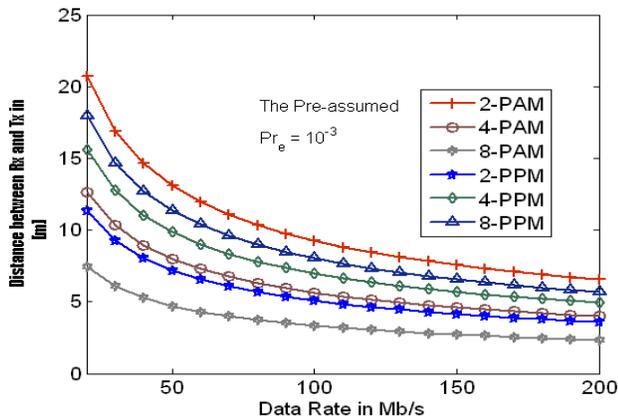


Fig. 6. Trade-off Flexibility for the distance between transmitter and receiver as a function of data-rate between M-PAM and M-PPM signal

V. REAL-TIME SIMULATION OF IR-UWB COMMUNICATION SYSTEM

Simulink fixed-point simulation model been carried-out showing the real-time point-to-point IR-UWB communication behavior for different signal parameters propagating over multi-path IEEE 803.15.3a channel model that can simulate the signal behavior over different LOS (lines of Sight) channels types and NLOS (Non-line of Sight) with different distances ranging from CM1 to CM2 channel, then simulation of the received signal will show the effect of the multipath fading reflection on the channel. At the end the BER performance evaluation of the transmitted signal at the receiver side will show in real-time the signal behavior and its trade-off flexibility of the performance for different sets of the signal parameters (which directly indicating the data-rates) and the distances from different multi-path channel types.

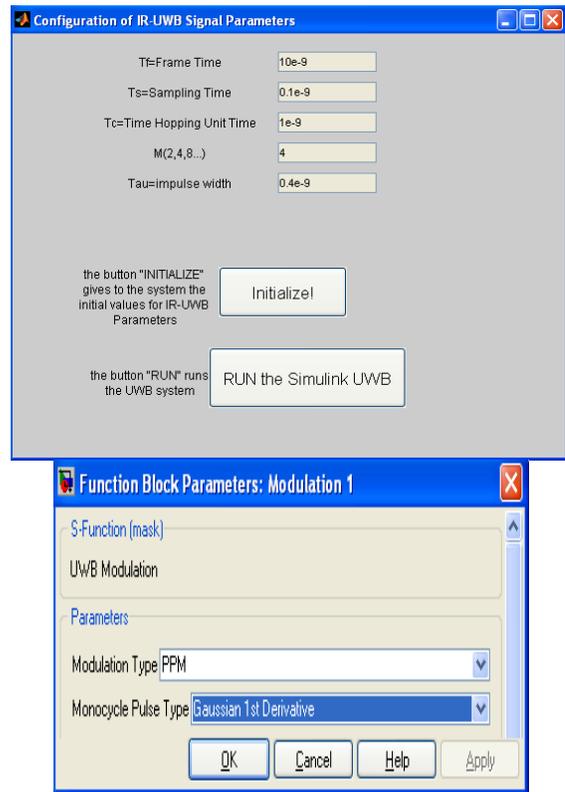


Fig. 7. Parameters Settings for TH-PPM Signal

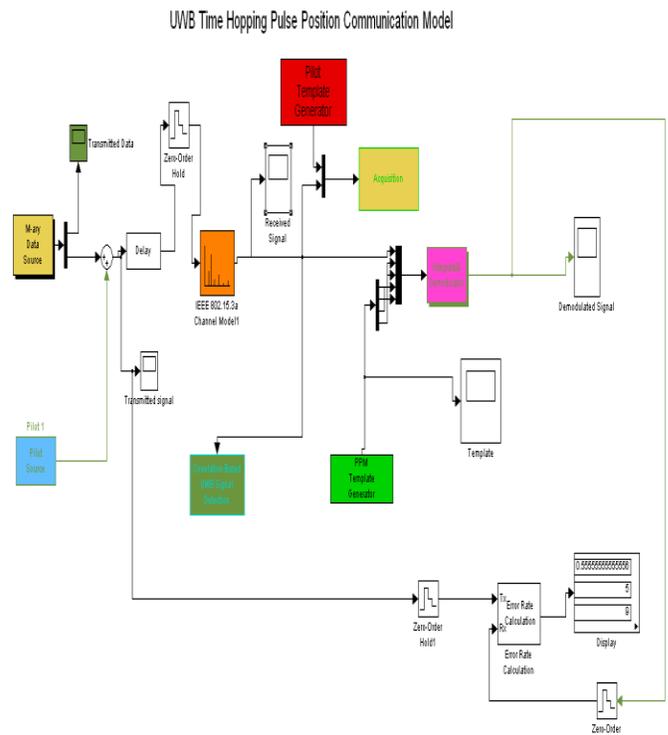


Fig. 8. IR-UWB Communication Model with BER Performance Evaluation

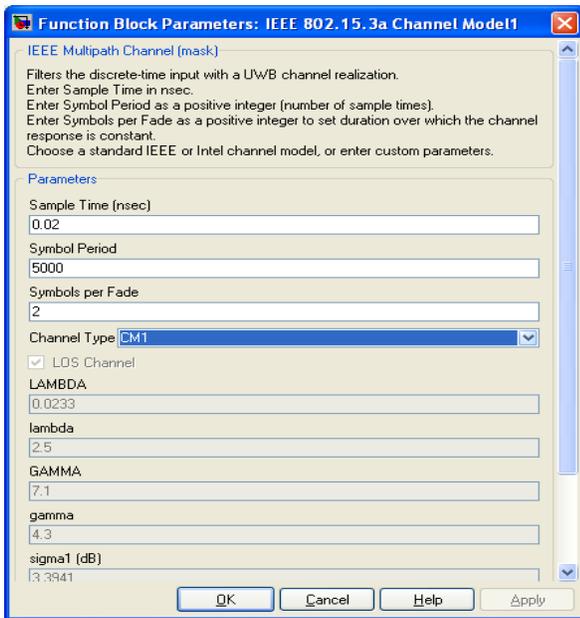


Fig. 9. Configurations of IEEE 803.15.3a Channel Model

Table I shows the system parameters in real-time implementation as well as the BER results. All values are obtained from the Simulink workspace during and after system simulation. The IR-TH PPM system is designed for specific data rates of predefined BER requirements propagating over four IEEE 802.153a channel model scenarios (C1 to C4), as shown from figure 4.19. The required BER should be within the range of 1×10^{-2} s to 4×10^{-2} s. The design of the system is according to the settings of these parameters, including signal generation, shaping (at the data-source level), and modulation (at the signal-pilot level), to achieve the required performance of the radio link, given that the BER result is 2×10^{-2} s.

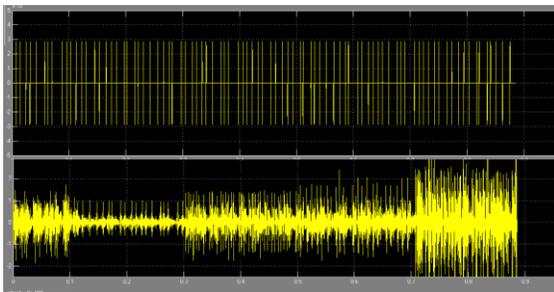


Fig. 10. the upper part is the TH-PPM Transmitted signal propagating over IEEE 802.15.3a channel model with configurations as stated in figures 7 & 9. The lower part is the received signal with the multi-path channel model effects.

TABLE I. IR-UWB SYSTEM PARAMETERS AND EVALUATION IN REAL TIME

Data source	
M-ary	2
Symbol time	100.000 ns
Link selector number [0 reserved for pilot]	1
Promised data-loading state	[11000]
Modulation	
Frame time	10.000 ns
TH unit time	1.000 ns
PPM unit time	0.156 ns
Impulse width	0.2877 ns
Symbol time in terms of frame time T/Tf	10
PN sequence period	31
Average signal power	-40.0 dBm
Pulse energy	1.000000e-015 J
Initial state of PN code shift register [m = 5]	[10001]
Bits used to generate hopping code	3
Channel	
IEEE 802.15.3a channel model	
UWB signal detection threshold	95%
Acquisition threshold	70%
False alarm threshold	30%
Initial state of receiver PN code shift register	[00010]
Simulation in real time	
Sampling time (uwb.Ts) = 0.0500 ns; UWB signal detected; time = 4.550 ns; coefficient = 97.515	
Acquisition coefficient = 10.092%	
Acquisition coefficient = -3.805%	
Acquisition coefficient = 100.319%	
PN hopping code acquisition	
Acquired frame time = 590.000 ns	
Tracking coefficient = 101.779%	
Tracking coefficient = 100.507%	
Data synchronization; demodulation begins.	
Demodulation start time = 810.000 n	
BER command	Bercal
BER result	0.020000000000000 s

VI. FUNCTIONAL COMPARISON OF TH-PPM COMMUNICATION USING P220 EVALUATION KIT

This section provides a meaningful comparison among different TH-UWB approaches and their performance metrics. A functional comparison approach is adopted. First, a number of different application scenarios are considered. Then, performance metrics pertaining to these application scenarios are used for comparison. Finally, conclusions describing the performances according to indoor and outdoor environments and recommendations for applications are drawn.

The deployment in WPAN communication scenarios is considered for both indoor and outdoor environments, wherein data rate varies from very high to low for different coverage ranges. These scenarios are applicable in wireless body networks, in which sensor nodes are all placed in the field of the human body and require radiating extremely high-data-rate UWB pulses for imaging, as well as in UWB radar applications for monitoring patient movements in intensive care units, which can be classified as high rate with short-range WPAN (few meters). Low data rates with medium range are applicable for RFID systems and WSNs, whereas high data rates with medium range are applicable in digital multimedia smart environments (including homes, clinics, and libraries), where nodes send high-data-rate multimedia contents.

A. Propagation Environments

This subsection presents the results of the measurements conducted in selected indoor and outdoor environments under the following conditions.

- The outdoor environment is an open grassy plain under LOS condition for transmitter and receiver positioning, as shown in Figure 11.
- The indoor environment is under NLOS conditions, wherein numerous thick and thin walls, electrical equipment, and the distances between the transmitters are considered. The layout of this test site is shown in Figure 11, where T1 represents fixed transmitter location; and R1, R2, R3, and R4 denote different receiver distances from the transmitter location. The estimated ranges between the four positions of the transmitter and the receiver are: 30.48m, 38, 1 m, 45.72 m and 60.96 m.



Fig. 11. Experimental layout of indoor/outdoor propagation

B. Experimental Results and Discussion

Figures 12 and 13 show variations of bit energy to effective noise density ratio E_b/N_{eff} as a function of distance and data rates between the transmitter and the receiver based

on the measurements for four transmission rates in two conducted sets of measurements (indoor and outdoor) for short-to-medium-range communications. Based on the result, E_b/N_{eff} decreases as distance increases of increased data rates for both outdoor and indoor experimental environments. For example, at 9.6 Mbps the E_b/N_{eff} is less than 5 dB which is the lowest compared to other lower data rates with the same distance. This is because as the data rates become higher, more data can be transmitted which makes more noise can disturb into the system data.

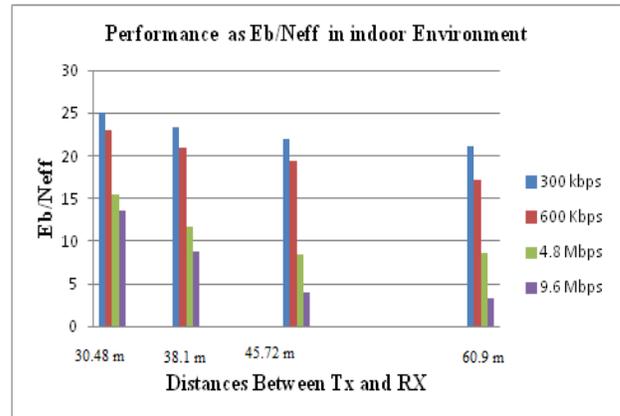


Fig. 12. Performance evaluations with regard to E_b/N_{eff} in indoor environment

In comparison with indoor performance, outdoor outperform indoor in general as shown in Figure 13 for many cases for example; E_b/N_{eff} for 300 Kbps for all distances, 4.8 Mbps in 38.1 m and 60.96 m also 9.6 at 45.72 m. This can be explained as the pulses in indoor environment are more exposed to multipath fading effects which resulting in pulses power loss.

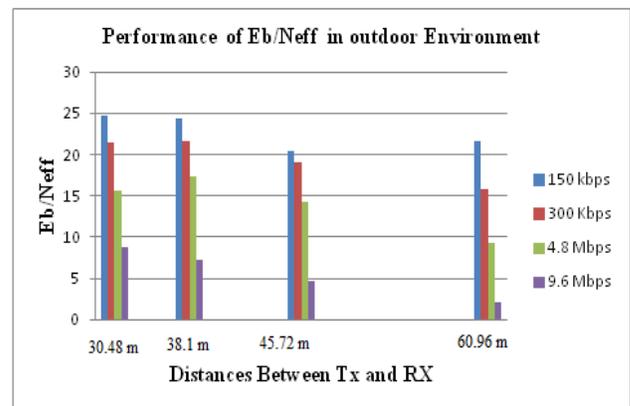


Fig. 13. Performance evaluations with regard to E_b/N_{eff} in outdoor environment

BER obtained at different data rates for indoor and outdoor measurements are respectively plotted in Figures 14 and 15. It shows that in the extreme distance case of 60.96 m in indoor environment, the BER is higher for all the data rates and for the extreme data rate of 9.6 Mbps at the distance of 54.72 m the BER also is with higher value for indoor compared to outdoor performance. Results in general show that BER increases as distance and data rate increase and this is in

agreement with the theoretical analysis and previous MATLAB simulation. Similarly, the performance of IR-UWB in outdoor scenarios with regard to BER in various distances and data rates outperforms that in indoor scenarios especially for the cases of high data rates with longer distances.

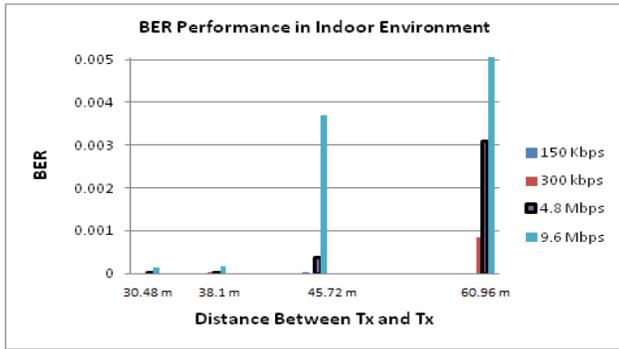


Fig. 14. Performance evaluations with regard to BER in indoor environment

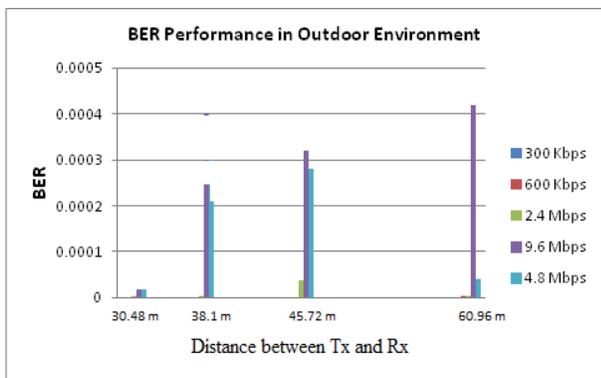


Fig. 15. Performance evaluations with regard to BER in outdoor environment

Figures 16 and 17 plot the experimental relationship between BER and E_b/N_{eff} for outdoor and indoor environments, respectively. Based on the curves, the performance generally matches with the theory. However, results differ significantly in several cases of the experiment in outdoor environments at data rates between 4.8 Mbps and 9.6 Mbps (Figure 17). Such difference can be related to severe multipath degradation in several locations where conditions change, as well as to the limitation and sensitivity of the evaluation tool-kit.

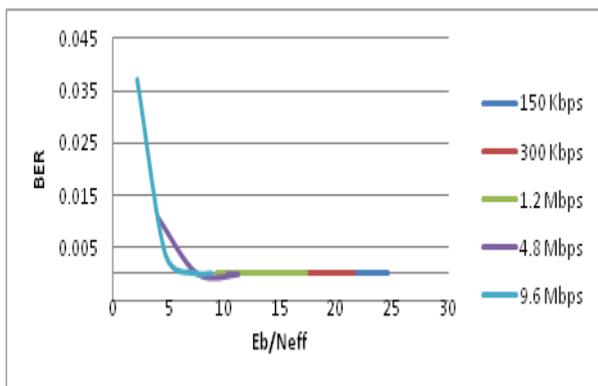


Fig. 16. Comparison of performance evaluation between BER and E_b/N_{eff} for indoor environment

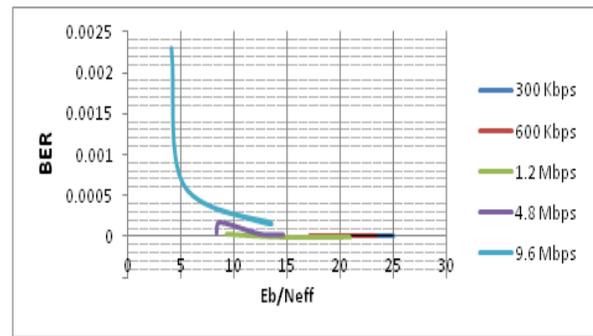


Fig. 17. Comparison of performance evaluation between BER and E_b/N_{eff} for outdoor environment

The achieved link throughput is effective evaluation criteria for wireless solution designer. The tunable impulse based UWB proposed solution is evaluated experimentally for the actual achieved data rate in different indoor distances as shown in Figure 18. It can be considered as a map to check the performance of the designed solution at certain distance. In general, the throughput decreases as distance increases but, the throughput reduction is more obvious for high data rates in extreme distance.

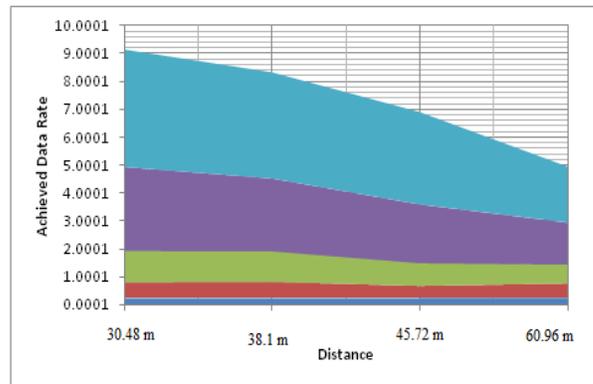


Fig. 18. Throughputs (achieved data rate) as a function of distance between Tx and Rx in indoor radio environment

C. Performance Statistical Evaluation of Experimentally Collected Data

For the two test scenarios (outdoor and indoor environments), results show the effects on performance of varying the distance on the achieved data rate for BER and the E_b/N_{eff} . The BER, as a function of E_b/N_{eff} for each scenario, is plotted. This evaluates practically the proposed solution to validate the theory analysis as well as the simulation efforts.

The following conditions are observed from the performance evaluation:

- The throughput achieved for low rates of less than 1.2 Mbps is more stable in indoor environments. This finding can be explained by UWB pulse bounces and the reflection at the receiver side.
- The throughput for indoor environment is efficient at low-to-medium data rates ranging from 300 Kbps to 1.2 Mbps. By contrast, for high data rates ranging from 4.8 Mbps to 9.6 Mbps, the throughput is not supportive

during the transmission of the output power level of -12.2 dBm (approximately 50% packet drops). At an extreme distance of 61 meters, the throughput decreases significantly, as higher data rate is more exposed to interferences and multipath effects. Also, the signal power attenuated due to the extended ranges.

- During the transmission of data at a rate of 150 kbps, BER performance is perfect, except in the extreme case wherein the distance is 61 meters. Results of the first experiment show that BER rate increases as distance increases, thus exhibiting the same trend as Eb/Neff varying according to distance.

VII. CONCLUSIONS AND FUTURE WORKS

IR-UWB link budget analysis can be performed as a function showing the trade-off flexibility between BER, Data rates and coverage ranges for specific IR-UWB parameters. This can be utilized for different wireless indoor scenarios and as integrator for different wireless short-ranges infrastructures.

The IR-UWB cross-layer design architecture is needed since the cost function is proposed and developed at the MAC layer level to reconfigure the physical layer parameters for optimal performance according to pre-fixed requirements (pre-defined Data rates, BER or ranges).

Future research can be done to optimize this solution using genetic algorithm as the cost function can be considered fitness function. The initial population started with IR-UWB certain parameters then the crossover and mutation operations are performed to approach the optimal parameters selection for optimal performance according to the fitness function (cost function) and for Pre-defined BER requirements.

REFERENCES

- [1] UWB Communication Systems A Comprehensive Overview. EURASIP Book Series on Signal Processing and Communications, ed. T.K. Edited by: Maria-Gabriella Di Benedetto, Andreas F. Molisch, and C.P. Ian Oppermann, and Domenico Porcino. Vol. Volume 5, 410 Park Avenue, 15th Floor, #287 pmb, New York, NY 10022, USA: Hindawi Publishing Corporation.
- [2] Wang, Y.-q., et al., UWB multi-pulse position modulation for high data-rate wireless application. The Journal of China Universities of Posts and Telecommunications, 2006. 13(4): p. 19-23.
- [3] Lecoindre, A., D. Dragomirecu, and R. Plana. Software defined radio layer for IR-UWB systems in Wireless Sensor Network context. in Circuits and Systems, 2007. NEWCAS 2007. IEEE Northeast Workshop on. 2007.
- [4] Aubin LECOINTRE, D.D., Robert PLANA, Design and Hardware Implementation of a Reconfigurable Mostly Digital IR-UWB Radio. ROMANIAN JOURNAL OF INFORMATION, 2008. Volume 11, Number 4: p. 295-318.
- [5] Liu, H., et al., Taxonomy and Challenges of the Integration of RFID and Wireless Sensor Networks. IEEE network., 2008. 22(6): p. 26.
- [6] Martin, R.F. Ultra-wideband (UWB) rules and design compliance issues. in Electromagnetic Compatibility, 2003 IEEE International Symposium on. 2003.
- [7] Fernandes, J.R. and D. Wentzloff. Recent advances in IR-UWB transceivers: An overview. in Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on. 2010.
- [8] Pezzin, M., et al. A LDR IR-UWB Demonstration Platform for Services based on Location and Tracking. in Mobile and Wireless Communications Summit, 2007. 16th IST. 2007.
- [9] Ha, D.S. and P.R. Schaumont. Replacing Cryptography with Ultra Wideband (UWB) Modulation in Secure RFID. in RFID, 2007. IEEE International Conference on. 2007.
- [10] Pengyuan Yu, P.S.a.D.H.V.T., Securing RFID with Ultra-wideband Modulation: Electrical and Computer Engineering Department. Blacksburg, VA 24061
- [11] Pereira, D.P., et al. Model to integration of RFID into Wireless Sensor Network for Tracking and Monitoring Animals. in Computational Science and Engineering, 2008. CSE '08. 11th IEEE International Conference on. 2008.
- [12] Ze, L., S. Haiying, and B. Alsaify. Integrating RFID with Wireless Sensor Networks for Inhabitant, Environment and Health Monitoring. in Parallel and Distributed Systems, 2008. ICPADS '08. 14th IEEE International Conference on. 2008.
- [13] L opez, T.S., et al., Integrating Wireless Sensors and RFID Tags into Energy-Efficient and Dynamic Context Networks. The Computer Journal, 2009. 52(2): p. 240-267.
- [14] Beiwei, Z., H. Kunyuan, and Z. Yunlong. Network architecture and energy analysis of the integration of RFID and Wireless Sensor Network. in Control and Decision Conference (CCDC), 2010 Chinese. 2010.
- [15] Baghaei-Nejad, M., et al. A Novel Passive Tag with Asymmetric Wireless Link for RFID and WSN Applications. in Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on. 2007.
- [16] Kadhim, D.J., L. Wei, and C. Wenqing. Ultra Wideband Cognitive Network Objective Issues. in Future Computer and Communication, 2009. FCC '09. International Conference on. 2009.
- [17] Arslan, H. and M. Şahin, UWB-Based Cognitive Radio Networks, in Cognitive Wireless Communication Networks. 2007. p. 213-230.
- [18] Di Benedetto, M.-G. and G. Giancola, Understanding ultra wide band radio fundamentals. Prentice Hall communications engineering and emerging technologies series. 2004, Upper Saddle River, N.J.: Prentice Hall PTR.

Comparative Study of the Software Metrics for the complexity and Maintainability of Software Development

Dr Sonal Chawla
Associate Professor
Department of Computer Applications
Panjab University Chandigarh

Gagandeep Kaur
Assistant Professor
Department of Information Technology
GGDSD College Chandigarh

Abstract—Software metrics is one of the well-known topics of research in software engineering. Metrics are used to improve the quality and validity of software systems. Research in this area focus mainly on static metrics obtained by static analysis of the software. However modern software systems without object oriented design are incomplete. Every system has its own complexity which should be measured to improve the quality of the system. This paper describes the different types of metrics along with the static code metrics and Object oriented metrics. Then the metrics are summarized on the basis of relevance in finding the complexity and hence help in better maintainability of the software code, retaining the quality and making it cost effective.

Keywords—Static metrics; OO metrics; MOOD

I. INTRODUCTION

Software Metrics are used to increase the quality of software since decades. For the better software development, measurement plays a very critical role for software engineering to make it a true engineering discipline. Hardware as well as software became complex day by day, so manageability is a major concern. Past were the days when only traditional metrics were used to improve the quality and technical decisions regarding softwares.

Modern systems are impossible without OO design as object-oriented programming plays a very critical role for effective and efficient software development. Software engineers developed many ways to maintain software quality and developed softwares using object-oriented programming to solve the common problems. Object-oriented design contains all the properties and quality of software that is related to any large or small project [1].

It is a degree through which a system object can hold a particular attribute or characteristics. Object-oriented is a classifying approach that is capable to classify the problem in terms of object and it may provide many paybacks on reliability, adaptability, reusability and decomposition of problem into easily understood objects and providing some future modifications [2].

II. OBJECTIVE

The software quality engineering metrics are used for quality planning, process improvement, quality control,

reliability estimation and analysis of customer satisfaction data. They are used to increase the efficiency of software development life cycle. For example if the number of defects are less, the effectiveness of the Development and the Testing team is improving. To make the modern application software reliable and maintainable large numbers of metrics are used these days. This paper is an attempt to understand the impact of static and OO metrics values on the complexity and maintainability of the code. In the first section, static metrics are discussed preceded by OO metrics because characteristics of object oriented design like abstraction, inheritance, modularity and polymorphism cannot be represented using traditional metrics as they play an important role in modern software applications.

Only object oriented metrics allow the modifications to reduce the cost effectiveness, time consumption and improve the quality. Additionally there is an attempt to discard the obscure metrics and use the simple ones because easy and simple ones are appreciated in software applications and also they are easy and simple to collect.

Also size measures and complexity alone cannot provide accuracy in maintaining the applications and they alone are inappropriate for predicting the defects, so other important OO metrics are used to for reducing the complexity and easier maintainability of modern applications. Moreover modern applications are incomplete without OO design.

III. STATIC CODE METRICS

Static metrics are derived from the measurement on static analysis of the software code. It is performed without executing any of the code. Static analysis is better to understand the security problems within the program code and can easily identify nearly 85% of the flaws in the programming code.

A. Source Lines Of Code (Sloc)

Source lines of code (SLOC) is a software metric that calculate the size of a computer program by counting the number of lines in source code of program. Main types of SLOC measures are: physical SLOC (LOC) and logical SLOC (LLOC). Physical SLOC is the total count up of lines in the program's source code together with comment lines. Logical SLOC measures the number of executable statements.

B. Comment Percentage (Cp)

The CP is defined as a ratio of the number of comment lines to the number of non-blank LOC [3]. Software development life cycle is normally long. In any stage of the life cycle, comments will help developers and maintainers to better understand the programs. Higher comment percentages will increase understandability and maintainability [4]. It is suggested to maintain at least 8% on comment percentage to enhance the understandability [5].

C. Halstead Metrics

Halstead Metrics are used to measure the complexity of a program by using operands and operators. Halstead metrics is used to interpret the source code as a sequence of tokens that can be operands and operators and counted as

- number of unique (distinct) operators (n1)
- number of unique (distinct) operands (n2)
- total number of operators (N1)
- total number of operands (N2).

The number of unique operators and operands (n1 and n2) as well as the total number of operators and operands (N1 and N2) are calculated by collecting the frequencies of each operator and operand token of the source program. Though Halstead Metrics are traditional metrics but they are used to measure the modern programs like C, C++ and Java. These metrics are used to calculate the errors, programs size, volume and testing time.

D. McCabe's Cyclomatic Complexity

Cyclomatic complexity is a software metric that is used to measure the complexity of a program and was measured by McCabe in 1976. It directly measures the number of free paths through the source code of program. Cyclomatic complexity is calculated using the formula.

$$\text{Cyclomatic Complexity} = E - N + P$$

Where E is the number of edges of the graph; N is the number of nodes of the graph; P is the number of connected components. These metrics are used for control quality of software products.

IV. OBJECT ORIENTED METRICS

Dynamic metrics are derived from the measurement on dynamic analysis of the software code. They are based on studying the code behavior during execution. Earlier major work was focused on static metrics but now more attention has given to Dynamic metrics as they study the code at run time. Object-oriented programs can use Halstead Metrics but some essential factors like inheritance coupling remain uncovered using these metrics.

The CK metrics suite is designed for measuring object-oriented programs [6]. The suite includes six metrics discussed as follows.

A. Chidamber And Kemerer (Ck) Metrics Suite

Chidamber and Kemerer (CK) are the most well known object-oriented suite of measurements for Object-Oriented software. They have defined six metrics for the OO design.

a) Weighted Method Per Class (Wmc)

It is defined as the sum of the complexities of individual class. A class with more member functions than its peers is considered to be more complex and therefore more error prone [7]. As the children will inherit all the methods defined in a class, the potential impact on children will be as greater according to the number of methods in a class.

b) Depth Of Inheritance Tree (Dit)

The depth of a class in object oriented programming can be found with the inheritance. Hierarchy is the maximum extent from the node to the root of the tree. The higher the level of inheritance is greater is the value of DIT.

c) Number Of Children (Noc)

Number of immediate subclasses of a class is called its NOC. Greater number of children of a class means more reusability as inheritance is the form of reusability.

d) Coupling Between Object Class (Cbo)

It is defined as the count of the classes to which this class is coupled. Coupling is defined as: Two classes are coupled when methods declared in one class use methods or instance variables of the other class. The more independent a class is, the easier it is to reuse it in another application. The larger the number of couples, the higher the sensitivity to changes in other parts of the design, and therefore maintenance is more difficult. The higher the inter-object class coupling, the more rigorous the testing needs to be.

e) Response Of A Class (Rfc)

It is defined as number of methods in the set of all methods that can be invoked in response to a message sent to an object of a class. Greater the number of methods to be invoked, greater is the complexity of the class.

f) Lack Of Cohesion In Methods (Lcom)

It is defined as the number of different methods within a class that reference a given instance variable. To promote encapsulation, cohesiveness of methods within a class is desirable. To decrease the possibility of errors during development process, high cohesion decreases complexity.

B. Mood (Metrics For Object Oriented Design)

Metrics for Object Oriented Design (MOOD) are used to measure object-oriented programs. These metrics are language independent and can be obtained in the early phases of software development life cycle.

a) Method Hiding Factor (Mhf)

MHF is defined as the ratio of the sum of the invisibilities of all methods defined in all classes to the total number of methods defined in the system under consideration. The invisibility of a method is the percentage of the total classes from which this method is not visible.

b) Attribute Hiding Factor (Ahf)

AHF is defined as the ratio of the sum of the invisibilities of all attributes defined in all classes to the total number of attributes defined in the system under consideration.

c) Method Inheritance Factor (Mif)

MIF is defined as the ratio of the sum of the inherited methods in all classes of the system under consideration to the total number of available methods (locally defined plus inherited) for all classes.

d) Attribute Inheritance Factor (Aif)

AIF is defined as the ratio of the sum of inherited attributes in all classes of the system under consideration to the total number of available attributes (locally defined plus inherited) for all classes.

e) Polymorphism Factor (Pf)

PF is defined as the ratio of the actual number of possible different polymorphic situation for class C_i to the maximum number of possible distinct polymorphic situations for class C_i .

f) Coupling Factor (Cf)

CF is defined as the ratio of the maximum possible number of couplings in the system to the actual number of couplings not imputable to inheritance.

V. SUMMARY OF METRICS

Software metrics are becoming the basis of the software management and crucial to the accomplishment of software development. Consequently their values help in determining the complexity and hence the maintainability of the code. The below tables summarizes the above discussed metrics for the complexity and maintainability of code.

Here, the impact of increased or higher value of the metrics on the Complexity and hence the Maintainability is analyzed (Table I - Table III). It illustrates, in general, whether a high or low value is desired for the metric for better code quality [8] [9] [13]. We have marked the high value as „1“and low value as „0“to represent in a graphical form (Fig. 1 - Fig. 3). It is shown(dark line) that the higher value of metrics increase the complexity of code, while the metrics with low value and hence lower the complexity are shown in light shade lines.

TABLE I. STATIC METRICS

Static Metrics (High Value)	Complexity	Maintainability	Desired Value of Metrics
SLOC	High	High	Low
CP	Less	Low	High
HM	High	High	Low
MCC	High	High	Low

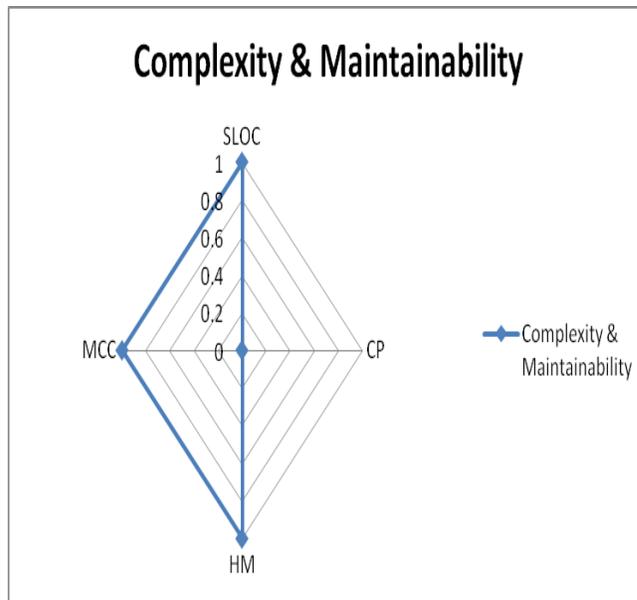


Fig. 1. Complexity and Maintainability of software code with high value of Static metrics

TABLE II. OO METRICS

OO Metrics (High Value)	Complexity	Maintainability	Desired Value of Metrics
WMC	High	High	Low
DIT	High	High	Low
NOC	High	High	Low
CBO	High	High	Low
RFC	High	High	Low
LCOM	Less	Less	High

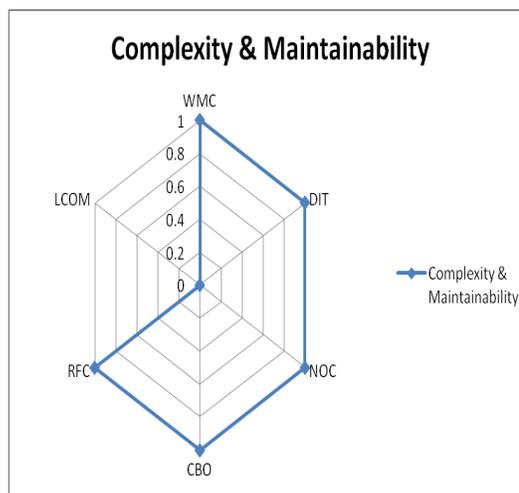


Fig. 2. Complexity and Maintainability of software code with high value of OO metrics

TABLE III. MOOD

MOOD Metrics (High Value)	Complexity	Maintainability	Desired Value of Metrics
MHF	Less	Less	High
AHF	Less	Less	High
MIF	High	High	Low
AIF	High	High	Low
PF	High	High	Low
CF	High	High	Low

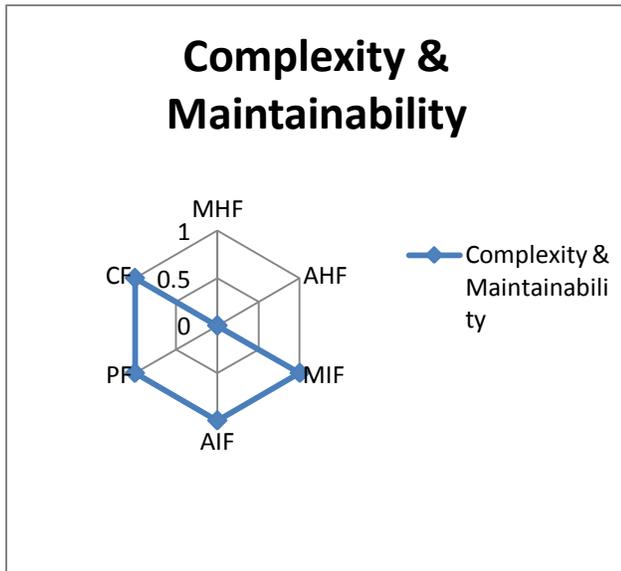


Fig. 3. Complexity and Maintainability of software code with high value of MOOD

VI. CONCLUSION

With the advancements in the software industry, measuring the software quality is complex for the development of the software product. Therefore the need for the development of better software metrics has increased over time. Since the

metrics plays a significant role in determining the complexity and thus the maintainability of the software code. Subsequently appropriate survey and study should be done to select the best metrics for the code. Each metric describes important features as, how to use it, interpretation guidelines, published thresholds whenever is possible, and assesses its appropriateness and usefulness. This would result in guiding and accessing the software to produce a robust, high-quality result, which enhances the potential reuse of the software and reduce the software maintenance cost.

REFERENCES

- [1] A. Deepak, K. Pooja, T. Alpika, S. Sharma, "Software quality estimation through object oriented design metrics", IJCSNS International journal of computer science and network security, april 2011, pp 100-104.
- [2] A. Henderson, seller, "object oriented metrics:measure of complexity", Prentice Hall, 1996
- [3] Lorenz, M. and Kidd, J. 1994. Object-oriented software metrics: a practical guide. Prentice-Hall, Inc.
- [4] Sharble, R. C. and Cohen, S. S. 1993. The Object Oriented Brewery: A Comparison of Two object oriented Development Methods, ACM SIGSOFT Software Engineering Notes, Vol. 18, No. 2., pp. 60 -73.
- [5] McCabe Software. 2012. Metrics & Thresholds in McCabe IQ. Available at: <http://www.mccabe.com/pdf/McCabe%20IQ%20Metrics.pdf>
- [6] Chidamber, S. R. and Kemerer, C. F. 1994. A Metrics Suite for Object Oriented Design, IEEE Transactions on Software Engineering, Vol. 20, No. 6, pp. 476-493.
- [7] Watson, A. H., McCabe, T. J., and Wallace, D. R. 1996. Structured testing: A testing methodology using the cyclomatic complexity metric. National Institute of Standards and Technology Special Publication 500-235.
- [8] Basili, V. R., Briand, L. C., and Melo, W. L., "A Validation of Object Orient Design Metrics as Quality Indicators," IEEE Transactions on Software Engineering, vol. 21, pp. 751-761, 1996.
- [9] <http://www.aivosto.com/project/help/pm-oo-ck.html>
- [10] <http://agile.csc.ncsu.edu/SEMaterials/OOMetrics.htm>
- [11] <http://www.cc.uah.es/drg/b/RodHarRama00.English.pdf>
- [12] <http://www.enggjournals.com/ijcse/doc/IJCSE11-03-09-003.pdf>
- [13] <http://www.ukessays.com/essays/information-technology/polymorphism-in-object-oriented-design-information-technology-essay.php>
- [14] <http://www.aivosto.com/project/help/pm-oo-mood.html>

Multithreading Image Processing in Single-core and Multi-core CPU using Java

Alda Kika

Department of Informatics
Faculty of Natural Sciences, University of Tirana
Tirana, Albania

Silvana Greca

Department of Informatics
Faculty of Natural Sciences, University of Tirana
Tirana, Albania

Abstract—Multithreading has been shown to be a powerful approach for boosting a system performance. One of the good examples of applications that benefits from multithreading is image processing. Image processing requires many resources and processing run time because the calculations are often done on a matrix of pixels. The programming language Java supports the multithreading programming as part of the language itself instead of treating threads through the operating system. In this paper we explore the performance of Java image processing applications designed with multithreading approach. In order to test how the multithreading influences on the performance of the program, we tested several image processing algorithms implemented with Java language using the sequential one thread and multithreading approach on single and multi-core CPU. The experiments were based not only on different platforms and algorithms that differ from each other from the level of complexity, but also on changing the sizes of the images and the number of threads when multithreading approach is applied. Performance is increased on single core and multiple core CPU in different ways in relation with image size, complexity of the algorithm and the platform.

Keywords—multithreading; image processing; multi-core; Java

I. INTRODUCTION

In recent years, Java language has become a popular choice for development of multithreaded applications due to the language multithreading support. Multithreaded programming allows simple identification of the sections of code that can be executed concurrently to exploit parallelism. The programs must be able to exploit this kind of parallelism in order to get performance gains in computing. There are two different ways to implement concurrent applications. The first approach is through the creation of processes, with all the communication made through messages, which are responsible for keeping all the necessary information for the programs, including register content and memory space [1]. The second approach uses thread, which are also known as light processes [2]. A thread is a point of execution within a process and they represent a key concurrency model supported by modern computers, programming languages, and operating systems. The threads exchange information only through shared memory and can be up to 20 times faster in their creation time when compared to processes [3].

A multi-threaded process has multiple points of concurrent execution within the process [4]. The use of multiple threads allows an application to distribute long running tasks so that they can be executed in parallel. This is also possible with

significant advances of multi-core systems. Today, multi-core processors are widely deployed in both server and desktop systems. The performance of multi-threaded applications could be improved on multi-core based systems because the workload of threads could be dispatched to cores, which work in parallel [5].

One of the good examples of applications that benefit from multithreading in a multi-core processor is image processing. Image processing “refers to the manipulation and analysis of pictorial information” [6]. The main idea of parallel image processing is to divide the problem into simple tasks and solve them concurrently, in such a way the total time can be divided between the total tasks (in the best case) [7]. The general idea behind image processing involves examining image pixels and manipulating them. Image processing can be a time consuming task based on the matrix structure of the image leading this process towards a multithreading algorithm.

Many authors have addressed the multithreading topic in java on multi-core systems. The java as a suitable programming language for parallel software and the power of multi-core processing is studied by Peter Bertels and Dirk Stroobandtin [8]. In [9], authors studied and implemented parallel multi-threaded implementations of two popular clustering algorithms: k-means and mean-shift. The experimental results show that good parallel implementations of those algorithms are able to achieve nearly linear speedups on multicore processors. In [10] Mahmood has implemented the imaging filter on multi-core processors for win32 platform. He shows that for large images, the parallel implementation approaches Amdahl’s ideal curve. In [4] authors observed that multithreading leads to tune the application performance considerably. The performance and scalability issues of multithreaded Java applications on multicore systems are studied in [11]. This interesting topic has solicited many scientific works so far.

In our work we studied how the algorithm performance changed when multithreading approach is applied on different single core and multi-core platforms. We also studied how complexity of the algorithm and the image size of the images influence the performance.

In order to exploit how multithreading can improve the performance of the algorithm, we have done some experiments with several image processing algorithms such as brightness, contrast and steganography executed on either single-core CPU or multi-core CPU using single-thread and multithreading

approaches in Java. The algorithms that we have taken in consideration differ from each other from the level of complexity. As the image size increases, the number of calculations increases too. For this reason we have experimented with different sizes of images. The difference in the algorithms and image size would influence on a further change of the level of complexity.

We measured the algorithms processing time on different platforms and image sizes that we took in consideration.

In this paper, we have provided insight into Java multithreading approach performance on single and multi-core CPU platforms and hardware suggesting the combination complexity and image size that gives the best performance.

Section I gives a brief introduction and related work. The rest of the paper is organized as follows. Section II presents an overview of multithreading programming and the algorithms implemented with this approach. Section III outlines the experimental setup including test programs and specifications of the experimental platforms. Section IV presents the results and section V gives a brief analysis of the results. The paper is concluded in Section VI and VII by presenting the conclusions and further work.

II. MULTITHREADING PROGRAMMING

First, in creating a multithreaded algorithm, there are three basic considerations. The first step is to identify the parallelism. This may mean simply decomposing the problem domain of a conventional algorithm into several sections. The second step is to control the access to shared data items. The third step is to optimize the algorithm [4].

The application running in the operating system is either single-threaded or multi-threaded. Single threaded applications require one thread to run on the CPU. Whereas a multithreaded program contains two or more parts that can run concurrently.

We have implemented three image processing algorithms using Java. Java has become a leading programming language soon after its release and it is an emerging option for High Performance Computing [12]. Java supports multithreading and we can construct single-thread as well as multi-thread application with it. A multi-threaded program in java has many entry and exit points, which are run concurrently with the main() method. The three constructed algorithms have been implemented using single thread approach and multithreading approach. In the multithreading approach the shared memory in which the threads operate is the matrix of the image pixels. We have used the Java packages to grab the pixel matrix of the image that has to be processed. Then different threads manipulate different parts of the matrix depending on the algorithm. The work task and the part of the matrix that each thread has to manipulate are determined by the main thread. The time that is necessary to manipulate all the matrix either by a single thread or by all the threads is registered.

III. EXPERIMENTAL SETUP

A. Testing programs

Initially, three image processing algorithms have been chosen to test the multithreading approach on different single-

core and multiple-core platforms. The first and the second algorithms change brightness and contrast of the chosen image and display the changing image. Brightness and contrast are very important features of an image. Brightness refers to the overall lightness or darkness, whereas contrast is the difference in brightness between objects or regions. For the algorithm of the brightness we have used the formula :

$$\text{Arithmetic mean model} = (r + g + b) / 3 \quad (1)$$

After finding the mean from the formula 1 we use a value input from the user to change the brightness of the image following these steps:

- Subtract the overall mean from every color value of every image pixel.
- Add the mean value multiplied by the brightness value to every color value.

In the Figure 1 a view from the brightness application is presented.



Fig. 1. A screenshot from the brightness application

The contrast algorithm will multiply every color value by a scale factor determined by the user. It has a little less calculation than the brightness algorithm.

The third algorithm is from the field of steganography. Steganography is the art of hiding the fact that communication is taking place by hiding information in other information [13]. In our implementation of steganography algorithm we hide a message in an image in such a way that the modification of the image is imperceptible. Since people perceive red and green color brighter than the blue color, the image would be darker in the red and green pixels and brighter in the blue pixels. In order to hide the original message in the image we have constructed an algorithm that change only the last bit of the blue color of the image pixels using the message and a private key as well [14]. The complexity of this algorithm is greater than the other two. A view from the application that implement the third algorithm designed with multithreading approach is given in the Fig. 2.



Fig. 2. A screenshot from the steganography algorithm application

B. Experimental platforms

Since the first multicore processor was released over a decade ago, today's processors implement up to a dozen cores and this number is expected to increase because of Moore's law [15]. Multi-core CPU-s support advanced capabilities, such as multithreading and parallel processing. The advantage of a multi-core processor is increased speed. They are widely used across many application domains including embedded, network, graphics etc. The experiments were conducted on the Windows Operating Systems (Windows XP and Windows 7), single-core and multiple-core CPU-s platforms.

In Table 1 is given a brief description of the four used experimental platforms.

TABLE I. THE DESCRIPTION OF EXPERIMENTAL PLATFORMS

Platform	Processor	Operating System	Number of Cores
Platform1	AMD Athlon™, 1.11GHz	Windows XP	1
Platform2	Intel Pentium 4, 3.2 Ghz	Windows XP	1
Platform3	Intel Core™2 Duo Processors, 2.2 GHz	Windows 7	2
Platform4	Intel Core™ i5-4570R Processor, 2.7 GHZ	Windows 7	4

We used the first and second platform to test the algorithms according to the single thread and multithreading approach in single core, and we used the third and fourth platform to test the algorithms in the multi-core CPU.

C. Input Data

Not only the platforms but also the size of the input images influence on the execution time of the algorithms. We perform the experiments using three image sizes: 371*281, 500*500 and 1024*768. We refer to the dimension of the images in the experiments as S(Small), M(Medium) and L(Large). The results of these experiments are presented in the next section.

IV. EXPERIMENTAL RESULTS

Every algorithm was repeated 10 times and the average of all the results was recorded.

From the implementation of the tests of the single thread approach algorithms on the single-core and multiple-core platforms we recorded the results shown in Table 2:

TABLE II. SINGLE-THREAD APPROACH RESULTS

Platform Image Size	Algorithms performance(ms)		
	Contrast	Brightness	Steganography
1(S)	30	40	190
1(M)	50	40	351
1(L)	110	100	872
2(S)	29	29	62
2(M)	29	29	172
2(L)	47	62	296
3(S)	15	20	50
3(M)	16	21	70
3(L)	16	31	172
4(S)	10	14	32
4(M)	11	16	47
4(L)	18	26	109

We redesigned the same algorithms with multithreading approach. Then we varied the number of threads from 1 to 10 and then the multiples of 5 thus 15, 20, 25, 30 ...100.

The results from the experiments showed that on the platform with single core (platform 1 and platform2) the processing time of the algorithms decreases in a considerably rates when multithreading approach is applied. However when the number of threads increases over 10 threads the processing time does not decrease, but remains almost constant with a tendency to increase more than in the case when the single thread approach is applied. Generally when the number of threads increased from 1 to 10 the average processing time has the tendency to remain constant and when the number of threads increases more than 10 the processing time increases and it becomes bigger than the executing time in single thread approach.

Table 3 shows the result for the first platform. The result refers to the average executing processing time for the 10 first threads for the small, medium and large images.

TABLE III. AVERAGE RESULTS OF THE FIRST 10 THREADS FOR THE FIRST PLATFORM

Image Size	Algorithms		
	Contrast	Brightness	Steganography
Small	18	20	151.3
Medium	41	27.1	295.5
Large	109	69	824.3

The tables 4,5,6 display the average executing processing time for the 10 first threads in case of small, medium and large images for the second, third and fourth platform.

TABLE IV. AVERAGE RESULTS OF THE FIRST 10 THREADS FOR THE SECOND PLATFORM

Image Size	Algorithms		
	Contrast	Brightness	Steganography
Small	15	15	48.4
Medium	20.6	15.5	100
Large	46.5	29.9	271.8

TABLE V. AVERAGE RESULTS OF THE FIRST 10 THREADS FOR THE THIRD PLATFORM

Image Size	Algorithms		
	Contrast	Brightness	Steganography
Small	10	11	31.3
Medium	18.8	15.4	59.6
Large	51.3	46.9	162.3

TABLE VI. AVERAGE RESULTS OF THE FIRST 10 THREADS FOR THE FOURTH PLATFORM

Image Size	Algorithms		
	Contrast	Brightness	Steganography
Small	13.5	12.3	19.7
Medium	17.8	15.7	43.7
Large	31.2	23.2	101

V. PERFORMANCE ANALYSIS

As we expected when single thread approach of the algorithms is applied, the processing time decrements if we increase the speed and the number of the cores from one platform to the other (table 2). When multithreading is applied,

the CPU without cores improves the processing time with a greater rate than when the number and cores of the processor increase in the case when the size of the image is small or medium and the complexity of algorithm is not big.

In the first and second platform where the number of cores is 1, the increase rate of processing time is more than 48.27 % when the size is small and the algorithms are brightness and contrast. In the other cases this rate varies from 15.9 to 44.82. The results do not fall in this interval when the size of the image is large and the algorithm is contrast (first platform 0.9 and the second 2) and when the algorithm is steganography (first platform 5.5 and the second 8.1).

The third and the fourth platform give different results. The result of performance increase rates can be even negative which means that the processing time of the single thread approaches are better than multithreading approach. The result for the third and fourth platform is displayed in the tables 7 and 8.

TABLE VII. PERFORMANCE INCREASE RATE FOR THE THIRD PLATFORM

Image Size	Performance increase rates(%)		
	Contrast	Brightness	Steganography
Small	33,3	45	44,7
Medium	-18,7	28,5	14,3
Large	-218,7	-51,6	5,8

TABLE VIII. PERFORMANCE INCREASE RATE FOR THE FOURTH PLATFORM

Image Size	Performance increase rates(%)		
	Contrast	Brightness	Steganography
Small	-40	14,2	37,5
Medium	-63,6	0	6,3
Large	-72,2	11,5	7,3

The third platform gives the best result when the size of the image is small for all the algorithms as shown in the table 7. When the size of the image is large it gives the worst result. When the complexity of the algorithm is bigger the processing time tends to be less.

In the fourth platform the image size seems to have no influence as much as in third platform in the processing time. But the complexity of the algorithm influence the processing time by reducing it.

From both platforms with multi core CPU the combination small image size and complex algorithm gives the best results, improving the processing time when multithreading occurs.

VI. CONCLUSIONS

Java language is very suitable to develop image processing applications due to its features and the free packages that it offers for this purpose. New open source java image processing packages is often added [16].

It is very important to understand how to improve the performance of this kind of applications using managed languages like Java. Through some experiments with the image processing algorithms the impact that multithreading approach has on performance is analyzed in single-core and multi-core platforms.

The results showed that the multithreading approach improves the performance processing time of algorithms either in single-core or multi-core CPU platforms but this improvement is different.

In single core the best results is given by the combination of small image size and less complex algorithm whereas in multi-core CPU the combination of small image size and more complex algorithm improves the performance. Multithreading programming can improve the performance on multi-core CPU when complex image processing algorithms is applied.

VII. FUTURE WORK

Others studies are needed to be done to understand the causes of these behaviors. Future work includes developing a C# environment with image processing algorithms in order to compare it with Java applications in different platforms. The influence that parallel implementation of these multithreading approach algorithms will be another interesting task for the future. We assume that we can improve the processing time as it was shown for Wu-Lee Steganography Algorithm [17].

REFERENCES

- [1] J. F. CORSINI, L. G. FREITAS, and F. D. ROSSI, "Comparando Processos entre Unix e Windows". Revista INFOCOMP. March, 2006.
- [2] SCHEFFER, R. "Uma visao Geralsobre Threads". Revista Campo Digit@l. Volume 2.N_umero 1.Paginas: 7-12, 2007.
- [3] B. BARNEY, "POSIX Threads Programing". Lawrence Livermore National Laboratory. 2011. <<https://computing.llnl.gov/tutorials/pthreads>>.
- [4] M. Shanthi, A. Anthony Irudhayaraj. Multithreading, "An Efficient Technique for Enhancing Application Performance", International Journal of Recent Trends in Engineering, Vol 2, No. 4, November 2009, pp. 165-167.
- [5] Kuo-Yi Chen, Fuh-Gwo Chen, "The Smart Energy Management of Multithreaded Java Applications on Multi-Core Processors", International Journal of Networked and Distributed Computing, Vol. 1, No. 1, January 2013, pp.53-60.
- [6] Gregory A. Baxes, Digital Image Processing: Principles and Applications. John Wiley and Sons, Inc, New York, NY, 1994.
- [7] S. Saxena, N. Sharma, Sh. Sharma, "Image Processing Tasks using Parallel Computing in Multi core Architecture and its Applications in Medical Imaging", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 4, April 2013
- [8] Peter Bertels, Dirk Stroobandt, "Java and the Power of Multi-Core Processing". The Second International Conference on Complex, Intelligent and Software Intensive Systems, CISIS 2008, pp. 627 – 631
- [9] Honggang Wang, Jide Zhao, Hongguang Li, Jianguo Wang. "Parallel Clustering Algorithms for Image Processing on MultiCore CPUs". International Conference on Computer Science and Software Engineering, 2008.
- [10] Faran Mahmood, "Parallel Implementation of Imaging Filters on Multi-Core Processors for Win32 Platform". Proceedings of the 4th International Conference on Open-Source Systems and Technologies, ICOSST, 2011.
- [11] Kuo-Yi Chen, J. Morris Chang, Ting-Wei Hou. :Multithreading in Java: Performance and Scalability on Multicore Systems:, IEEE Transactions on Computers 2011 (vol. 60 no. 11), pp. 1521-1534
- [12] G.L. Taboada, J. Touriño, R. Doallo, "Java for high performance computing: assessment of current research and practice:, in: Proc. 7th Intl. Conference on the Principles and Practice of Programming in Java, PPPJ'09, Calgary, Alberta, Canada, 2009, pp. 30–39.
- [13] Obaida Mohammad Awad Al-Hazaimeh, "Hiding Data in Images Using New Random Technique", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 2, July 2012, pp 49-53.

- [14] S. Greca, A. Kika, O. Qirici, "A new steganography method for hiding the message in the image", *Buletin of the Natural Sciences Faculty*, No. 14, 2012, pp. 211-220.
- [15] G. E. Moore. "Readings in computer architecture. Chapter Cramming more components onto integrated circuits", Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2000, pp 56-59.
- [16] Tobias Pietzsch, Stephan Preibisch, Pavel Tomanc and Stephan Saalfeld, "ImgLib2 generic image processing in Java", *BioInformatics*, Vol. 28 no. 22 2012, pp. 3009-3011.
- [17] S. Greca, E.Martiri, "Wu-Lee Steganographic Algorithm on Binary Images Processed in Parallel", *International Journal of Video & Image Processing and Network Security IJVIPNS-IJENS* Vol: 12 No: 03,2012, pp 1-4.

EICT Based Diagnostic Tool and Monitoring System for EMF Radiation to Sustain Environmental Safety

K Parandham
Aisect University,
Bhopal-Chiklod Road,
Raisen, Bhopal, India

RP Gupta
Aisect University
Bhopal-Chiklod Road
Raisen, Bhopal, India

Amitab Saxena
Aisect University
Bhopal-Chiklod Road
Raisen, Bhopal, India

Abstract—the adverse effects of electromagnetic radiation from mobile phones and communication towers on health issues are being well documented today. However, exact correlation between radiation of communication towers and their radiation levels, are not monitored.

Aim of this paper is to study, analyze, apply networking and data mining technologies to develop an EICT based Diagnostic tool and Monitoring system for electromagnetic radiation levels into environment. This system is to network all mobile towers of each service provider as a single entity and then connect all service providers to a central monitoring agency online for continuous monitoring. Since very large numbers of mobile towers exist in India, each state can have its own regional network which is further networked with national central network. This can be enlarged to entire world for monitoring the EMF radiation levels near every mobile tower. For these regional national and international networks the connectivity is to be instituted by the respective service provider.

In this paper an attempt is made to logically apply Data Mining and networking technologies to develop a central EICT based diagnostic tool and monitoring system for EMF radiation from each transmission tower. With this system regional, national and international agencies/authorities can monitor the EMF radiation at each and every transmission tower area continuously and verify them with exposure standards. It is proposed to display this information using Integrated Display System in front of monitoring authority at appropriate levels.

Keywords—EICT Based Diagnostic tool; Electromagnetic Fields(EMF) Radiation; Mobile Telephony; Data Mining; Data Warehousing; Electronics; Information and Communication Technologies(EICT); International Commission on Non-Ionizing Radiation Protection(ICNIRP); Compressed Natural Gas(CNG)

I. INTRODUCTION

Exponential growth and developments in various fields of science and technology in the last few decades have intensified the human interface into the natural environment and associated physical, biological and ecological systems resulting in various unintended and undesirable impacts on environment, human health and society.

Electronics, Information and Communication Technologies (EICT) is ushering in a revolution in every field of daily life and the technological advantages brought to society are unimaginable. EICT is helping mankind in many ways and at the same time giving away many negative effects

on environment, wildlife, human health and society at large. In addition, with the growth of mobile subscribers and their transmission towers, India is also witnessing a rapid population growth which is going to overtake China. For growing population the agricultural productivity and the problems influencing them should be of concern. The population of many species such as honey bees, which is one of the most important pollinator and useful factor for agricultural productivity, has seen a drastic population drop. In literature there is no much data about the effects of electromagnetic radiation available for most of our free living floral and faunal species in India.

Mobile communication industry is one of the fastest growing industries in the world. In recent years, there has been an exponential increase in the usage of mobile telecommunication devices, which has become an easy means for communication. The use of mobiles have become more conspicuous, during the last decade and this has led to construction of transmission towers in large numbers, built in urban, as well as in rural areas including other sparsely populated areas. Transmission towers are based on the electromagnetic waves, which over prolonged usage have adverse impacts on humans as well as on other fauna. The adverse effects of electromagnetic radiation from mobile phones and communication towers on health of human beings are being well documented today. Recently the electromagnetic fields from mobiles and other sources have been classified as “possibly carcinogenic to human “by the WHO,s International Agency for Research on Cancer (IARC). However, exact correlation between radiation of communication towers and wildlife, are not yet very well established.

II. HISTORICAL BACKGROUND AND NECESSITY

The existing literature survey shows that the Electro Magnetic Radiations (EMRs) are interfering with biological systems in many ways. There had already been many warning bells sounded in the case of bees and birds, which probably heralds the seriousness of this problem, indicating vulnerability of other living beings as well. The electromagnetic radiations are being associated with the observed decline in the sparrow population in London and several other European cities (Balmori, 2002, Balmori, 2009, Balmori & Hallberg, 2007) [1-3]. A vast majority of scientific literature published across the world serious effects of EMFs in various other species too.

The pollution from EMRs being a relatively new environmental issue, there is a lack of established standard procedures and protocols to study and monitor the EMF impacts on humans especially among wildlife, which often make the comparative evaluations between studies difficult. In addition to the gap areas in research, the necessary regulatory policies and their implementation mechanism also have not kept pace with the growth of mobile communications industry. The present guidelines on exposure limits to EMF need to be refined since the ICNIRP Standard [4] currently followed in India is coined based on thermal effects of Radio Frequency and are dismissive of current epidemiological evidence on impacts of non-thermal nature on chronic exposure from multiple transmission towers.

Generally EMF Radiation hazards attributed are sleep disorders, headache, depression, discontent, irritability, nausea, dizziness, appetite loss, muscle spasms, numbness, tingling, altered reflexes, effects on eyes, brain, cells and many biological systems. Also birds losing orientation and reduction in their population have been published after research in many countries. A new study from Brazil found the direct link to 4924 cancer deaths due to cellular antennas radiation between 1996 and 2006 which was published in 17 May 2011[5]. About 80% of the victims lived within 500m away from cell phone towers. A study in ISREAL found that the incidence of cancer cases increased by four fold among people living within 350m of transmission towers[6]. Another study in Germany found that a three fold increase in new malignancies among people living within 400m from a mast after 5 years of EMF exposures and 7 fold increase in cancer cases in Berlin[7-8]. Another study in Spain realized increased health problems among people living within 300m from transmission towers[9]. Study by LB Despande in Sher-e-Punjab colony in Andheri East, Mumbai, India found that many house wives have been diagnosed with many forms of cancer living within 91m from a mobile tower. There are very large number of research papers and case studies are available on the negative impacts of Electromagnetic radiations on environment, human health, birds & bees and wildlife. Most of them are quoted in expert group study report on the "Possible impacts of communication towers on wildlife including Birds and Bees", constituted by Ministry of Environment and Forests, Government of India[10].

Since the inception of mobile communication technologies, there have been concerns about the ill-effects of the mobile towers and mobile phones. Despite being a relatively newly acknowledged form of pollution, EMRs and their negative impacts on environment and biological systems have already been reported by several studies. However, most of the existing scientific literature on the negative environmental effects of electromagnetic fields, reports the results of experimental and epidemiological studies examining the serious impact on various aspects of human health. WHO and Government of India disapproves majority of these cases due to lack of proof. There is a conflict and difference of opinion exists among many agencies concerned about EMF radiation. One side Government agencies along with World Health Organization, another side cellular mobile operators along with manufacturers and third party is the society

concerned with health issues. To resolve this issue initially monitoring of EMF radiation levels is the basic necessity so that mitigation activities can be initiated.

Therefore there is a strong case for design and development of EICT based diagnostic tool and on line monitoring system for EMF radiation in and around every transmission tower across the world. EICT can be effectively applied for design and development of diagnostic tool and monitoring system for continuous monitoring of EMF radiation from every transmission tower for ensuring environmental safety.

The large amount of varied EMF radiation data available across the country/world is to be surveyed, captured, meaningfully processed, sequenced, classified, structured, identified and hierarchically stored in various tiers based on policies decided. Data mining technology will be a solution involving high end advanced systems storage and state of the art software tools need to be included in the EICT Based Diagnostic tool and Monitoring system for continuous monitoring and analysis of EMF radiation levels from all the transmission towers. This information can be displayed in front of designated authorities through DLP for real time decisions to mitigate the irregularities in the EMF radiation patterns of any of the mobile service providers to confirm to safety standards. This requirement is that of the hardware, software and solution technologies.

III. METHODOLOGY

The data inputs considered for application to develop EICT based Diagnostic tool and Monitoring System for continuous monitoring of EMF radiation levels from each transmission tower is the radiated power density at specific distances.

The proposed methodology for networking of various EMF radiation values as inputs for EICT based Diagnostic tool and Monitoring system for on line monitoring EMF radiation levels around each transmission tower is an shown in figure.1, 2 and 3.

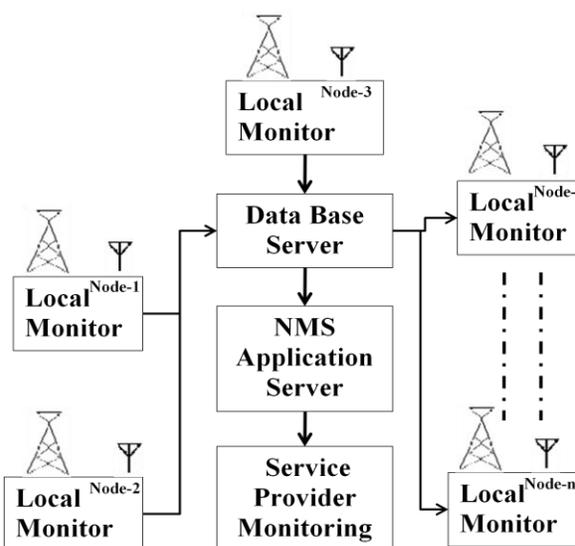


Fig. 1. Networking of EMF Radiation levels from each transmission tower service provider

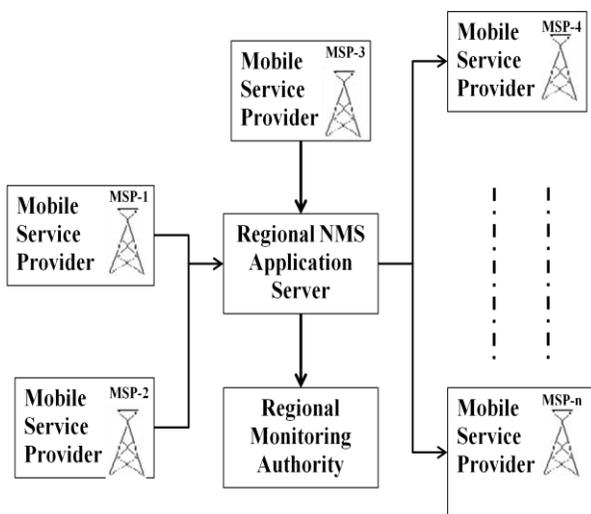


Fig. 2. Networking of EMF Radiation levels from each service provider in a region

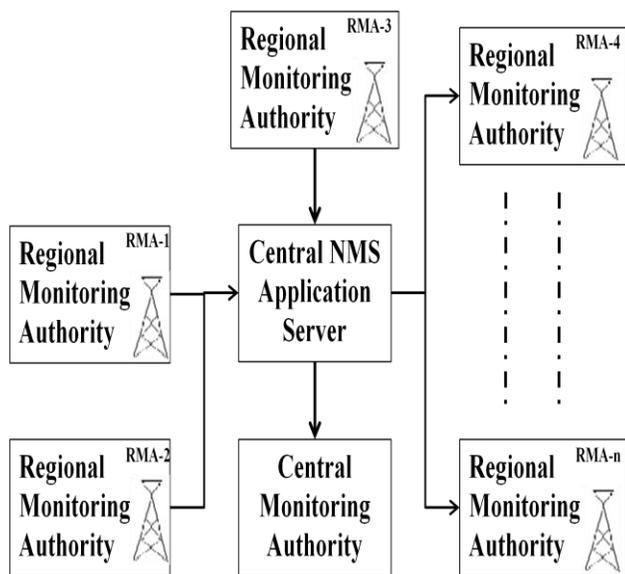


Fig. 3. Networking of EMF Radiation levels from each region to central monitoring authority

At every transmission tower wideband area monitor or smart monitor is envisaged at around 100m distance. Frequency switching to monitor EMF radiation to be inbuilt into this system which can be selected from either locally, regionally or nationally through respective network management software. The local monitoring set up at each transmission tower can be termed as local monitoring node which is as shown in figure 1.

From each monitoring node the measured EMF radiation data will be made to ride through respective service provider network. The EMF radiation data inputs from each node will pass through standard communication network such as multiplexer, sampling server, switch to database servers based at regional centre. The database servers are connected to NMS application server which is then connected through router to regional gateway server as shown in figure 1.

Similarly all the cellular mobile service providers will provide respective EMF radiation inputs to regional gateway server of the regional monitoring authority. The monitored EMF radiation levels as inputs from each regional gateway server are linked to the central NMS application server. These data inputs then stored sequenced and displayed through a high end data wall display system.

These data inputs are required to be mapped in respect of each transmission tower, networked, processed, stored, and fused which are required to be displayed in front of monitoring authority. To integrate all these inputs for maximum situational awareness a multi-input, multi-window display processor is proposed in this paper. A variety of incoming EMF radiation levels will need to be displayed in a compact arrangement in front of monitoring authority for viewing for taking right decisions for mitigation of imbalance when compared to standard radiation limits accepted by the Government.

For this a display system that could combine multiple incoming inputs and display them using Digital Signal Processing, sound processing system, integrated controller, switcher, and Digital Light Processor based projectors, projecting the comparative EMF radiation levels and accepted standard limits on a glass beaded screen requirements are proposed in this research paper which are structured in figure 4.

A. Salient Features of Customized Web NMS For This Application.

The customized web Network Management Software (NMS) of this proposed EICT base Diagnostic tool and monitoring system for EMF radiations is to have the following salient features.

- Comprehensive monitoring capabilities to monitor EMF radiation levels, health of infrastructure, components, network protocols, system metrics with a single tool.
- Centralized view of entire status of EMF radiation levels selectively.
- Fast detector of alarming increase in the EMF radiation levels and their alerts to designated authorities.
- Alert acknowledgements provide communication to particular cellular mobile service provider on alarming EMF radiation levels and problem response.
- In case of regional NMS management facility to reduce / switch of EMF radiation power levels.
- Generation of historical reports and EMF radiation levels compared with standard radiation levels.
- Multiuser access to customized Web NMS interface to health department authorities, environmental specialists and department of commutation to view the EMF radiation levels.
- Extendable architecture for integration with in house and global applications.

- Scales to monitor many regions and thousands of nodes.
- To have fail over capabilities to ensure nonstop monitoring.

B. Data Mining Concepts Application.

A high end advanced data capturing and analysis set up is proposed for this solution. The solution will involve high end advanced systems and storage and state of the art software tools for data analysis. The set of requirements proposed for this Diagnostic tool and monitoring system, EICT infrastructure needs are given below:-

- The EICT infrastructure should be capable of handling large data volumes of generally un-structured data available in various formats including but not limited to textual, video, audio and structured encrypted information.
- The total data volume will be in Gega Bites or in Tera Bites.
- The server infrastructure should be suitable for online analysis and processing and should be capable of handling large data volume providing near real time processing of large number of complex queries over the data volumes, generating required meaningful information. The input data flow will be from large numbers of transmission towers across the region and country from different mobile service providers.
- There should be provision for hierarchical storage, archiving, content management and data backup.
- The EICT infrastructure should be secure enough to the standards of security.
- The EICT infrastructure should provide support for data conversion from the available media to digital format and there should be appropriate mechanism for digital asset tracking.

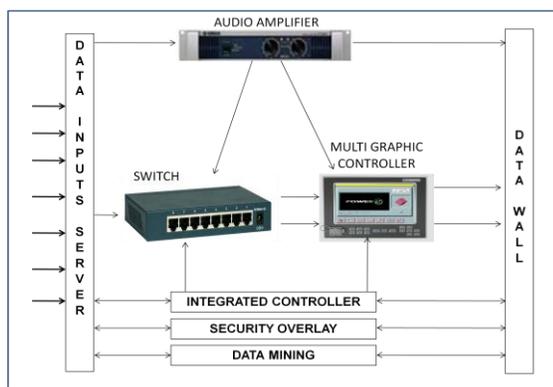


Fig. 4. Central Monitoring and Display System for EMF Radiation levels

- **Data Inputs Server.** This server receives and processes various EMF radiation inputs provided to the system and forward this information to various other sub-systems.

- **Switch.** A network device used to interconnect various subsystems for intra-system communications and passage of various EMF data.
- **Audio Amplifier.** The amplifier is used to amplify the audio signal received by data input server and forward this multi graphic controller as well as over the network through switch.
- **Multi Graphic Controller.** This controller processes the multiple EMF graphic/video signals and multiplexes them with the amplified audio signal for further display over data wall.
- **Data Mining.** Analytical tool for analyzing data from different perspective and summarizing it into useful information after comparison with standard accepted levels. In addition data mining is also the process of finding correlations or pattern among dozens of fields in large relational database.
- **Security Overlay.** Architecture to ensure secure communications between the sub-systems/nodes and regions. It includes encryption, media secrecy as well as confidentiality.
- **Integrated Controller.** It provides enterprise protection for servers and enhances the server performance. It provides support for all RAID levels to ensure safety of critical EMF levels which have exceeded standard levels.
- **Data Wall.** Data wall is a big multi-screen LCD display system for graphics output as well as for audio-visual display. The data wall provides easy projection of critical information (EMF levels) for further assimilation and analysis.

C. Data Fusion

The data fusion solution will be comprising of the following three technologies:-

- Data Warehousing (DW) for capturing and organizing the EMF radiation data for fast near real time processing and retrieval of meaningful information.
- Content Management (CM) for classifying and managing EMF radiation levels in a well defined logical layout.
- Hierarchical Storage Management (HSM) for identifying, structuring and management of EMF radiation data from various cellular mobile service providers from different regions and nodes into various storage tiers based on the policies decided.

IV. RESULTS & DISCUSSION

The IT infrastructure in figures 1, 2, 3 and 4 together forms the EICT based Diagnostic tool and Monitoring system for EMF radiation. The proposed system is to assist decision makers at regional and central (National) level to arrive at correct decision and have control over all cellular mobile providers at their disposal. This is an automated system for EMF radiation levels situation monitoring, to know status of

all data inputs in that particular node, region and to arrive at right decisions at right time to mitigate higher or dangerous levels of EMF radiation.

The proposed network of networks is geographically distributed system. Distribution of proposed network based EMF radiation monitoring is key to management of impacts of this radiation which will have geographically distributed networks. This also makes network management efficient and scalable, enabling the central monitoring authorities to focus on the analysis of exceeding levels of radiation data collected across the nodes and regions of different cellular mobile service providers.

Distributed Mediation Servers (DMS) are servers that can be deployed along with central NMS server in large scale or remotely distributed network such as the one proposed in this paper. Each DMS server performs network facing functions for a regional monitoring network of each cellular mobile server provider. The collected EMF data can then be correlated to transfer relevant EMF information to the central NMS server.

In distributed scenarios distribution is effective in providing scalability and remote management capabilities to NMS management applications since cellular mobile services are expanding in an exponential growth.

When deploying NMS management application to manage a very large EMF radiation data monitoring networks, the application needs to be highly scalable to manage the nodal, regional devices, ports and connectivity, both logical and physical of different service providers. This scalability need can be addressed by the customized web NMS distributing the network facing the functionality via the distributed mediation servers. The central EMF radiation monitoring NMS server can collate EMF radiation levels from different distributed mediation servers of different regions and provide a single console view of the entire network to the monitoring authority.

When NMS management applications are deployed by cellular mobile service providers, they do require management of their IT infrastructure health monitoring along with EMF radiation levels across multiple/remote locations. Here only EMF radiations are required to be transferred to regional or central monitoring location which can be addressed in NMS application server by deploying the distributed mediation servers at remote locations.

In a distributed set up like in the proposed EICT based diagnostic tool and monitoring system for EMF radiations the following components are very significant while deploying customized web NMS using DMS servers.

- Central Server
- Distributed Mediated Server and
- User interface

The central server here is a logical server and is made up of two servers, namely the back-end server and the front-end server. Front-end servers are servers to which the monitoring authorities and the DMS are connected. The numbers of front-end servers depend on the planned performance and scalability

metrics. The back-end servers store all the networks EMF radiation data in a centralized database and processes requests from monitoring authorities through the front-end servers. Both these servers are located at the central monitoring site. The collated EMF radiation data stored in a local database and correlated by the DMS. This ensures that only summary and critical data are sent to the central customized web NMS server. User interface facilitates the monitoring authority to connect to the DMS server at the specified node of particular region of the selected cellular mobile service provider to view the EMF radiation data available in the DMS database.

Network Management software for this proposed system is required in an operational environment where large amounts of information and sources need to be managed. The built in automation features of the software will be required to allow operators in the control room to focus fully on the important tasks. It will allow operators in the control room to focus fully on the important tasks. It has to allow right integration with custom applications, so that operators do not have to learn new tools or user interfaces after installation.

V. CONCLUSION

The world's entrance into the mobile telephony certainly had a profound impact on our society. It seems clear that the trend will continue to expand in many ways that will, no doubt, continue to surprise us. The expanding potential, however, is not an unmitigated blessing. It is very evident that there will be both positive and negative effects on society and environment.

The diagnostic tool and monitoring system proposed in this paper is an ideal solution for online monitoring of EMF radiations across all nodes, regions of different cellular mobile service providers to ensure environmental safety. This can give instant inputs of exceeding EMF radiation levels in any of the nodes of any region of any cellular mobile service provider. This can monitor in near real time and will be a great aid for knowing the EMF radiation levels across the nation.

However a diagnostic tool and monitoring system for respective controlling agencies such as Medical, Communication, Environment, Nongovernmental Organizations, Cellular mobile service providers and general public is felt appropriate for mitigating the higher EMF radiation levels.

A warehouse is more than an archive for data and more than a new way of accessing data. A warehouse is a subject oriented repository and provides tools to satisfy the information needs, not just for complex data queries, but as a general facility for getting quick, accurate, and often insightful information. It is to be designed in such a manner that users can recognize the information they want and access that information using this tool.

As per ICNIRP [11] guidelines EMF exposure limit in Canada is 3w/m^2 , in India it is 9.2w/m^2 (now 0.92w/m^2), 2w/m^2 in Australia, 0.09w/m^2 in Germany and 0.001w/m^2 in Austria and in New South Wales(Australia) it is 0.00001w/m^2 . Also successful communication with GSM mobile established in Germany [12] with 0.001microwatts per square meter. A mobile phone requires -80 to -100 dbm power for its

operation. Thus it is seen that at a distance of 50m the power level is 50 to 60 dbm higher in reality, meaning 1, 00,000 to 10, 00,000 times more power is radiated for mobiles operation. This is gravely hazardous to environment, health and to society.

Until now, society has been absorbing the harmful, invisible EM radiations without even being aware of it. With exponential growth in mobile communications this EM radiation pollution has started showing ill effects on environment, health of human beings, birds, bees and animals. Hence there is an urgent requirement to take precautionary steps to safe guard the environment and society. In this first step is to monitor for which the proposed EICT based diagnostic tool and monitoring system will be of great help. This system can give legal evidence and for regulatory authorities it will be a practical tool.

The growth of modern technologies in electronics, computers, and communications with availability of hardware for data mining and warehousing technologies, it is easy to realize the diagnostic tool and monitoring system proposed in this paper.

The EMF radiation from cellular mobiles and transmission towers is continuous and additive in nature. Stricter EMF radiation norms are required to be enforced. It is clear that we society has come up with alternate solutions for automobiles air pollution through unleaded petrol, CNG driven vehicles and hybrid vehicles. Similarly there is a need to come up with alternate solutions for EMF radiation problems and mandatory to monitor the EMF levels.

Cellular phone industries are multibillion dollar companies and products are linked to illness. Generally these Industries deny any health problem in spite of large number of health problems reported by many researchers from many countries. Mobile companies should not be in denial mode and have to accept that this is a real world problem.

Monitoring proposed in this research paper is the first step by regulatory authorities given the reason that in various countries the exposures limits vary from 12w/m^2 to 0.00001w/m^2 . Entire world has to come up and should have a single exposure limit which is sufficient for cellular mobiles operation for sustainable environmental safety.

REFERENCES

- [1] Balmori Martinez, A. (2003) The effects of microwaves on the trees and the other plants. Valadolid, Spain,
- [2] Balmori, A (March 2009) Electromagnetic pollution from phone masts. Effects on wildlife, Pathophysiology. 2009 March 3.
- [3] Balmori, A. & Hallberg, O. (2007) The urban decline of the house sparrow (*Passer Domesticus*): a possible link with electromagnetic radiation. Electromagnetic biology and medicine, 26,141-151.
- [4] Sainudeen Sahib. (2011) impact of mobile phones on the density of Honey Bees. Munis Entomology and Zoology, 6, 396-399.
- [5] <http://www.next.up.org/pdf/Brazil>, New study direct link to 4924 cancer deaths from cellular antennas radiation 28 07 2011 pdf.
- [6] Wolf R et al, International Journal of Cancer Prevention, 1(2) April 2004.
- [7] Eger H et al, Umwelt Medizin Gesellschaft, 17, 4, 2004.
- [8] Berlin, Head of Cancer Registry, 2006.
- [9] Santini R et al, Pathol Biol (Paris) 2002 July; 50(6) : 369-373.
- [10] Report on possible Impacts of communication Towers on wildlife including Birds and Bees, Ministry of Environment and Forest, Govt of India, 30 Aug 2011.
- [11] ICNIRP International Exposure Standards and Guidelines, <http://www.icnirp.de/documents/emfgdl.pdf>.
- [12] Haumann Thomas, et al, HF Radiation levels of GSM cellular phone towers in residential areas, Germany.
- [13] <http://thepeopleschronicle.in>

AUTHOR'S INFORMATION



K Parandham Gowd obtained B.Tech in Electronics and Communication Engineering with distinction from S.V. University, Tirupati. In 1994 he has conducted RCS Reduction experiments on coated (by pasting of absorber sheets) and uncoated scaled models of aircraft during his ME (Microwaves and Radar) as thesis work which is first time in India at IIT Roorkee. He has 48 research publications and 06 Technical reports to his credit most of them on RCS/RCS Reduction/Stealth Technology. He has one copyright to his credit on Dynamic RCS Range Validation Procedure from Govt of India. He is a Life Member of All India Management Association (AIMA), AeSI and Fellow of IETE. He had authored a book on Stealth Aircraft Technology in Hindi and English.



Dr R P Gupta is PhD, ME & MBA. He has 37 years of experience in Industrial Research & Development; Production; Quality Management; ISO Certification in Indian telephone Industries and Optel Communications Ltd; Bhopal. Since 2003 he is associated with Engineering colleges in the capacity of Principal/Director in BITS, Globus Engg College, RGPM, Surabhi college of Engineering and Scope College Bhopal. Presently he is life member of nine societies. His area of interests are Satellite communication, Nanotechnology and book writing. Presently he is Chairman IETE Centre (MP & CG), Bhopal and Director Development at Radharaman Group of Institutions, Bhopal.

Facing the challenges of the One-Tablet-Per-Child policy in Thai primary school education

Ratchada Viriyapong
Department of Mathematics
Naresuan University
Phitsanulok, Thailand

Antony Harfield
Dept. of Computer Science & Information Technology
Naresuan University
Phitsanulok, Thailand

Abstract—The Ministry of Education in Thailand is currently distributing tablets to all first year primary (Prathom 1) school children across the country as part of the government’s “One Tablet Per Child” (OTPC) project to improve education. Early indications suggest that there are many unexplored issues in designing and implementing tablet activities for such a large and varied group of students and so far there is a lack of evaluation on the effectiveness of the tablet activities. In this article, the authors propose four challenges for the improving Thailand’s OTPC project, consisting of: developing contextualised content, ensuring usability, providing teacher support, and assessing learning outcomes. A case study on developing science activities for first year primary school children on the OTPC devices is the basis for presenting possible solutions to the four challenges. In presenting a solution to the challenge of providing teacher support, an architecture is described for collecting data from student interactions with the tablet in order to analysis the current progress of students while in a live classroom setting. From tests in three local Thai schools, the authors evaluate the case study from both student and teacher perspectives. In concluding the paper, a framework for guiding mobile learning innovation is utilised to review the qualities and shortcomings of the case study.

Keywords—educational technology; m-learning; mobile computing; tablet-based education

I. INTRODUCTION

Mobile devices have become more relevant in all aspects of everyday life, with a significant impact on education. Approximately 10% of the 650,000 apps available on the Apple App Store (and about 300,000 on the Android Market) are under the Education category and 72% of these are for preschool and elementary age children [1]. Reasons for the popularity of tablets for education include: (i) portability makes it more convenient to use in the classroom, (ii) overall cost of tablets is lower compared to the cost of computers, (iii) children like using touch-driven and highly interactive technology [2,3] and (iv) preschool children learn to use tablets quickly, independently, and confidently, showing freedom of exploration [2,4].

As schools begin to accept the growing mobile culture, potential issues are being identified in the widespread adoption of mobile technology in the classroom. In the Bring Your Own Device (BYOD) schemes prominent in the U.S.A., devices can range from smartphones and tablets, to netbooks and laptops. According to a report by Tierney [5], one of the positive findings from BYOD projects is that instead of teachers

controlling the learning process and knowledge, students have become “empowered learners and active proponents of their understanding and ability to connect ideas in new ways”. However, the report also describes how BYOD can bring some difficulties to the school in practice, particularly in maintaining equality due to the different social/financial backgrounds and the wide range of devices available. Pilot projects found that the most common model, and the most successful, is for the school to define a single platform for BYOD [5]. Another well-known large-scale project is the One Laptop Per Child (OLPC) scheme which has been supported by non-profit organizations to provide affordable educational devices for use in developing countries (over 2.5 millions devices shipped worldwide) [6]. Several large studies have been undertaken and some have shown that laptop use in school have only a little or no correlation with students’ test scores [7] whereas in others the OLPC project was found to be effective in raising test scores [8]. An evaluation in Ethiopia reported that most teachers found trouble changing their teaching approach, which limited the use of the laptops in the class, whereas there was evidence in Haiti showing greater teacher engagement led to an enjoyable environment for students in the class [6]. Such studies indicate a wide range of high and low level issues for even the most successful widespread schemes like BYOD and OLPC.

Despite Thailand’s status as a developing country, the current government’s policy is radically in favour of adopting a tablet culture in schools. The government has given a tablet to every Prathom 1 (primary school, grade 1) child, loaded with approximately 336 learning objects (e-books, videos and interactive content) for five subjects including mathematics, science, Thai language, social studies and English language. To date, the project has delivered approximately 800,000 tablets to Prathom 1 students across Thailand – the biggest individual tablet procurement and the largest tablet experiment in the world. Their goal is for schoolchildren to “increase knowledge beyond textbooks”, especially for those in remote areas who lack access to learning resources, as described by the Prime Minister, Ms. Shinawatra, during the official launch of the One-Tablet-Per-Child (OTPC) scheme on 7th June 2012 [9]. Education minister, Mr. Suchart Tadathamrongvech, explained that some studies have shown the use of tablets is “a revolution in the education system”, which will elevate the learning quality of Prathom 1 students at a fast-growing rate and could be another step of Thailand for upcoming integration into the ASEAN Community [10]. Within this scheme, the Office of the Basic Education Commission (OBEC) at the Ministry of

Education is responsible for the installation of applications and content for all compulsory subjects. They also plan to train 549 supervisors to help instruct 54,900 Prathom 1 teachers in using tablets effectively (as yet, there is no information on the status of teacher training).

In comparison to the BYOD scheme, the OTPC project alleviates some of the problems of the BYOD project by using a single platform, single device model which eliminates the inconsistency of devices. Furthermore, the fact that they are given free to all students eliminates any social divisions. Compared to the OLPC project, tablets can provide advantages over laptop computers in terms of portability, cost, and efficiency (e.g. battery consumption and ease of recharging). Despite the practical advantages to tablets, there are still considerable challenges in realising an OTPC project as explored in this article.

II. CHALLENGES FOR OTPC

The goals of the Thai government's OTPC project are ambitious and the future pedagogical challenges are even greater than the initial challenge of delivering 800,000 tablets to primary school children. This article specifically identifies four new challenges for the success of the OTPC project, as identified during early observations of school children using tablets and from discussions with primary school teachers using OTPC devices. While these new challenges are derived from experiences in Thailand, they are also deemed to be relevant to other large-scale mobile learning projects in schools.

A. Contextualised content (challenge 1)

Early OTPC observations have reported that due to a variation of educational competency in Thai schools (c.f. the rural versus city social divide), the tablet activities need to take into account the learning abilities in different regional areas across Thailand [11]. This may also include differences in culture and language that vary from region to region.

Hence content accessibility is a key requirement and special attention is needed in producing suitable content for different contexts. Teachers at local schools often complained that many students did not have adequate reading skills for the activities provided on the OTPC devices. Moreover, there can be a significant difference in literacy levels among schools within the same region, so the content would ideally be adaptable to the context of the individual child and their current ability.

While the tablet activities provided on the OTPC devices take a „one size fits all“ approach, the future challenge will be to provide learning content that is suitable for different regions, different social and cultural backgrounds, and different levels of literacy.

B. Usability (challenge 2)

A number of studies have reported that the usability of mobile technology can be a factor in the success of educational activities (e.g. see Corlett *et al.* [12]). Within the OTPC project, battery life, screen brightness and button defects have had negative impact, but these hardware issues are not considered in this study as they are factors outside the control of schools

and teachers. Instead, consideration is given to the software issues, as they fall within the scope of improving the current OTPC project.

Teachers have reported that students can complete the entire body of activities in as little time as one month. This points to a flaw in that the tablet activities provide insufficient depth to the learner. In the same way that games such as Angry Birds (one of the bundled games included in OTPC) are providing a progression for the player, the learning activities should have a progression element to improve usability.

Another related issue is that the user interface in the OTPC affords a somewhat passive style of interaction whereby students „watch and then click next“. Maintaining young children's attention requires interactive content that actively engages students. The design and implementation of such content can be time consuming, requiring significant testing and iterative development to ensure that usability is appropriate for primary school children.

C. Teacher support (challenge 3)

There have been widespread calls for better teacher training [13] that is vital to the success of OTPC. In many schools, the teachers do not have tablets, which are causing severe problems for class preparation. From observations by the authors, many teachers use the tablets as an alternative to teaching (e.g. one period per day the children undertake tablet activities) rather than as a complement to traditional classroom activities. Clearly there is a need to support the teacher in the classroom environment so that more blended learning experiences are possible. Aside from the need for teacher training, Van de Bogart [14] suggests that the solution is to design the tablet software for OTPC such that „teachers would become engaged as much as the pupils“.

The challenge for providing better teacher support is intimately connected with how to provide better visibility to the teacher. In a traditional classroom, the teacher is directly aware of the students' interaction with the material. In a classroom with tablets, there needs to be similar support for the teacher by providing sensory information on the status of student's interaction with the tablet content.

D. Learning outcomes (challenge 4)

The final challenge, and perhaps the most difficult, is to address classroom management issues including evaluation of learning outcomes. One of the main concerns in a recent study of Thai primary school students using tablets in classrooms [15] was that „it is necessary for the teacher to monitor how students are using the tablet computer to achieve the learning objectives set out in the curricula“.

If the focus is on assessing the learning outcomes of children using OTPC tablets, there is first a need for lesson plans that link the curriculum to specific tablet activities, so that a teacher has a basis for incorporating the use of tablets into their existing classes. Then, building on challenge 3, well-designed tablet activities should have suitable data collection and analysis techniques, such that teachers and schools (as well as educators and technologists) can assess the learning outcomes of individuals, classes and schools.



Fig. 1. Two of 50 activities designed for the first year science curriculum (on the left, for learning “things that have life”, and on the right, for learning “parts of a tree/plant”)

III. MEETING THE CHALLENGES: A CASE STUDY

The “OTPC @ NU” research project began in 2012 with the aim of supporting local primary schools in Phitsanulok, Thailand, in obtaining greater benefit from the mobile devices provided by the OTPC project (see Mobile Computing Lab [16]). The project brief includes scope for developing and evaluating new tablet activities which are better suited to the needs of students, teachers and schools for primary school education in the local area of Phitsanulok (Prathom 1). The case study in this article focuses on how these new tablet activities provide some solutions to the challenges identified above, and where further work is still required.

A. Method

The current study follows the Lifecycle approach to evaluating educational technology, as described and applied by [17]. Unlike typical educational technology where evaluation is a particular phase of the project, the Lifecycle approach considers evaluation to play an important role in all stages of development, from the early stages of design through to implementation and testing.

Evaluation activities are undertaken at key stages in the lifecycle of the project, and inform the decisions in subsequent stages of the project. In this way, it shares common characteristics with the agile manifesto [18], which welcomes open evaluation and changes in requirements throughout the software development process. An agile approach was used during the software development process and hence, although the implementation is before and separate from the testing, the reality of the development process was that the preliminary testing contributed to further implementation cycles.

This case study is roughly described in three phases: (1) design of the learning activities, (2) implementation of the learning activities and related infrastructure, and (3) testing the learning activities and related infrastructure in live classroom environments. Teachers were involved in the evaluation of the design phase, and both students and teachers were involved in the evaluation of both implementation and testing phases. The

evaluation was mostly informal and qualitative, but some quantitative was collected in the testing phase.

B. Design

In order to provide some background to the design process, it is useful to consider a typical OTPC classroom environment. The government currently has provided every first year primary school child with a 7 inch tablet running version 4.0 of Google’s Android operating system. The tablets are pre-loaded with the OBEC’s LSystem learning environment [9] that contains content for the entire first year Thai primary school curriculum (standardised across the country). Additionally, individual schools may choose to load extra applications or games. Individual schools set their tablet usage policies for how much time is spent with the tablets and what types of activities are expected. Typical usage involves a teacher and 10-20 students in a classroom, sat at desks formally performing tablet activities for up to an hour at a time.

The purpose of the case study was to develop a new set of activities for the science curriculum of first year primary students. At the initial stage of the design, consultation was sought from Thai teachers familiar with using tablets in schools. The first year curriculum for science was examined and topics suitable for tablet activities were selected. The science curriculum is broken down into 5 main topics, and 10 different learning activities were selected for each topic. Storyboards were created for the 50 activities. Examples of the storyboards are shown in Fig. 1. Primary school teachers evaluated the designs and after several rounds of iteration the activities were approved for a prototype implementation.

With consideration to the four challenges, the designed activities were directly based on content from textbooks used in several schools in the local area of Phitsanulok, thus partially attempting to satisfy the needs of challenge 1 in providing contextualised content. Furthermore, the activities were designed with typical lesson plans in mind due to the close relation to content in the textbooks. Thus each of the activities is linked to a learning outcome defined in the curriculum, providing support for challenge 4.



Fig. 2. The three phases for one activity and the result (topic: objects for work vs objects for play)

The new activities are designed to complement, not substitute, the Ministry of Education's science activities bundled with the OTPC devices. However, the purpose of this study is not to simply „provide more content“. The four challenges identify issues with the current usage of the OTPC devices in schools, and therefore the main purpose of the study is to explore solutions to these somewhat broad issues by implementing new activities and testing them in a typical classroom environment.

C. Implementation

The activities were implemented using Adobe Flash CS5.5 with the programming in ActionScript 3.0, and deployed to Android tablets using AIR for Android [19]. The AIR platform was selected because it is a similar approach to the standard first year activities provided by the Ministry of Education [9]. Out of the 50 storyboards, 44 were fully implemented (the remaining 6 were not completed due to technical difficulties). An example of one activity is shown in Fig. 2. Each of the activities consisted of 3 phases: teaching, example and exercise. In the teaching phase, the topic is introduced and the key piece of knowledge is explained to the student. In the example phase, the student is given an example of what they must do in the exercise. Finally the student will undertake the exercise, which might consist of multiple screens where they must perform a similar task. If they perform incorrectly then the task will be repeated, otherwise they can continue to the next activity. The activity in Fig. 2 first shows the teaching phase (a) explaining to the student that for the objects shown, some objects for play and some are for work. Next there is the example phase (b) showing how the student must select the objects for play and place them in the basket. Finally, the student will perform the task himself or herself (c) by dragging

items into the appropriate basket and then the student will receive feedback (d) as to whether their answer was correct or incorrect.

Evaluation from testing early prototypes on school children played a significant role in ensuring that the usability (challenge 2) was suitable for the target group. Furthermore, the testing was essential for adding depth to the activities such that students could return to the activities to replay or progress further into the activities, as highlighted as part of challenge 2.

Challenge 3 was partially addressed during the implementation by developing a logging system for collecting data that would enable the teacher to view the progress made by each student. To accomplish this, it was necessary to send precise data on the interaction between student and tablet to a central server. A simplistic solution to this would be to send the data in real-time to the server. However, in a rural school environment (even more so in a home environment) there is no guarantee that a permanent wireless network is available, or that an Internet connection is present. Hence the need for an “offline” solution to data collection.

The architecture for collecting data from the tablet activities is shown in Fig. 3. The solution involves sending the logging data from the Flash activity to a background Android service that stores it in a local database. The service periodically polls the network to check if a connection to the cloud-based server can be made. When the connection is available, the data from the local database is forwarded to the server and removed from the local database. The data from multiple devices (from multiple sessions) is stored on the cloud-based server ready for analysis. A detailed description of the architecture is given by Nakrang *et al.* [20].

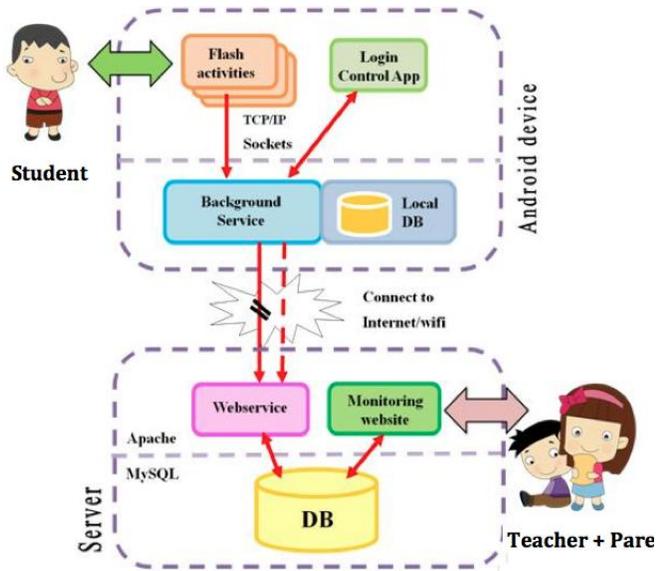


Fig. 3. The overall architecture for tablet activities and data collection

The data collected is valuable on two levels, corresponding directly to the two issues set out in challenges 3 and 4. Firstly, assuming the classroom has provision for the data to be collected in real-time (via a wireless Internet connection), then the teacher has a live source of information for understanding the current progress of every student in the classroom. The information can be made available to the teacher via a website viewable on a laptop or tablet (bottom-right of Fig. 3).

A prototype for possible visualisation of the data is shown in Fig. 4, consisting of test data from four student groups. The metrics displayed to the teacher include the last 12 interactions of the four groups, the total number of activities completed and the number of failures that have occurred. In the current progress of the class in Fig. 4, group 2 has yet to complete any activities which are one simple observation that would interest a teacher. Another observation would be that group 1 is producing incorrect answers to a large number of activities. Such insights could prompt the teacher to engage with certain students and goes some way to provide support for the teacher in terms of challenge 3. Teacher evaluation of this part of the system has been particularly positive, but further work is needed to understand the precise needs of the teacher.

The second valuable contribution of the collected data is in assessing the learning outcomes as per challenge 4. At the end of class, the teacher can view the individual progress of each student or the class as a whole. A teacher might be interested in which activities the students had difficulties with in order to plan extra offline work or future lessons. Aside from the teacher, the school as a whole might be interested in comparing learning outcomes between classes and parents might be interested in comparing their child with others in the class. The possible uses of the data for evaluation are broad and as yet unexplored by the authors. Further analysis on the potential for evaluating learning outcomes for challenge 4 is discussed in the following section.

D. Testing

Tests were performed on two occasions. The “preliminary tests” were undertaken in May 2012 at two schools in Amphur Bang Krathum, Phitsanulok. Following the feedback from the preliminary tests, the activities were revised and the software updated. The “final tests” were undertaken in September 2012 on a different group of students at one school in Amphur Muang, Phitsanulok.

The first of the preliminary tests was at a small local village school called Hnong Sra Phang School. We tested the activities on Prathom 1 and Prathom 2 students, approximately 12 students. We tested the students individually with a set of chosen activities for 20 minutes each. A researcher observed the student at all times, and provided guidance when difficulties occurred. The second was at a larger school called Wat Dong Hmee School where there were 20 Prathom 1 students. During this test, the students worked in groups of 2-3 per tablet, overseen by a researcher. They undertook the same activities for a 20 minute period. At least one teacher was on-hand during the tests and their feedback also contributed to the evaluation.

The final tests were undertaken at Wat Kung Waree School where there are 20 Prathom 1 students. They were split into groups of 2-3. The boys and the girls were tested independently, the first set of groups (1-4) consisted of all boys, and the second sets of groups (5-6) were all girls. Each group undertook a set of activities for 30 minutes. In the final tests, we collected detailed usage data from each tablet, including: (1) the number of activities played, (2) the amount of time spent on each activity, and (3) for each activity, the number of tasks that were passed and failed. The data was recorded via the architecture presented in Fig. 3. Some videos were also recorded for subsequent analysis. Due to the young age of the participants (Prathom 1 school children) it was not deemed appropriate to use questionnaires or interviews.

The evidence gathered from the preliminary tests was purely qualitative, and provided the evaluation that contributed to the iterative development of the activities. As discussed below, the preliminary tests provided essential feedback in overcoming challenges 1 and 2. The final tests provided some qualitative evidence verifying the improvements made over the preliminary tests, but they also focussed more on teacher support by gathering qualitative data that could be used as a basis for approaching challenges 3 and 4.

IV. RESULTS AND DISCUSSION

Of the four challenges for OTPC identified in this work, the first two can be considered as focusing more on supporting the student, whereas the last two focus more on supporting the teacher.

A. Contextualised content and usability (student-focussed)

Despite consultation from teachers during the design and implementation, the observation of students in the preliminary tests highlighted a number of problems as detailed in Table 1. The majority of these problems could be categorised as usability issues (challenge 2) or content issues (challenge 1). To overcome the issues it was necessary to redesign some areas

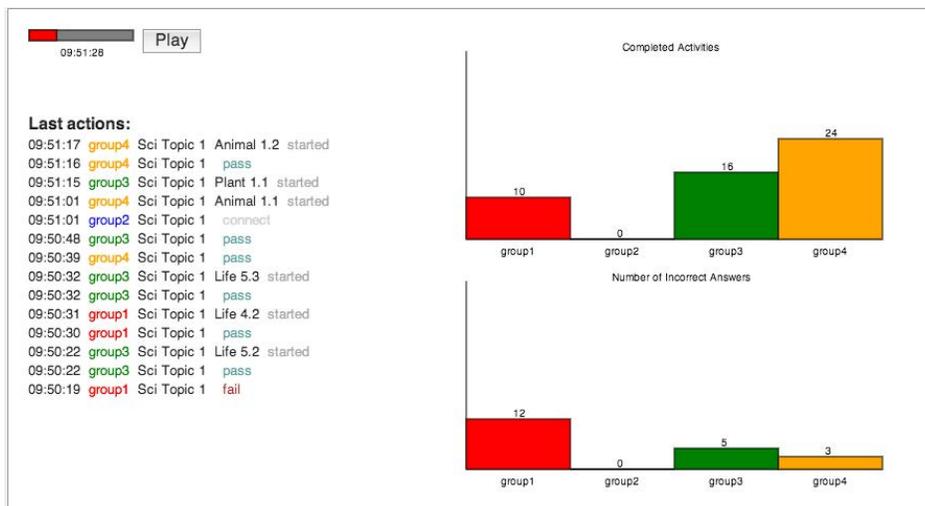


Fig. 4. Sample of the live monitoring website for teachers

of the user interface and to provide better support for students of different literacy levels (e.g. removing text and adding icons and voice-over).

TABLE I. OBSERVATIONS FROM THE PRELIMINARY TESTS

Observation	Problem	Solution
Students often do not know which parts of the screen to press.	The buttons are not clear enough and they are inconsistently used (challenge 2, badly designed interface).	Use one style of button throughout the activities. Use highlighting/animation to indicate that the button can be pressed.
Some students are not familiar with dragging objects on a touchscreen.	There is no training for the students on how to drag, or what can be dragged (challenge 2, unfamiliar user interface).	Animation can be added to show an example of how to drag. A tutorial page would also be useful.
Many Prathom 1 students cannot read	These students cannot read the text from the activities and therefore cannot do the activities without support (challenge 1, different levels of reading skills).	Provide voice-over on all activities so that reading is not required.
Some activities were too difficult and required teacher support/explanation.	Activities require explanation before the task begins (challenge 1 or 2, inappropriate content or badly designed interface)	Before the task, add a teaching page and an example page to show the what and how.

After improving the activities from preliminary results, in the final tests the students required much less support in using the activities. Several groups were able to use the activities independently for more than 10 minutes without support from the teacher. High levels of attention and engagement were observed. For example, in one video recorded there is evidence that the two boys were focussed on solving the task set in the activity. They each took it in turns to make an attempt, the other offering advice at the same time. After one of the boys got it incorrect the other showed clear signs of amusement and pleasure that it was his turn to show his friend. This was then

repeated the other way around until eventually one of the boys solved the activity. These results are purely speculative as there may be other explanations, e.g. the students in the final tests were more familiar with computers or mobile technology.

One obvious result is that the activities in final tests were more successful because the preliminary tests helped to resolve many interaction difficulties that the previous group of students had experienced. Whereas the difficulties in the preliminary tests were technical in nature, in the final tests the difficulties were more pedagogical or organisational. The following issues were observed:

- When in small groups, the students often tended to argue to be in control, or one student would take control. This would not be a problem if they each had their own tablet.
- Although the activities in the final tests contained voice-over to improve accessibility, a few of the activities still required some reading ability, and this was a problem for some students.
- Occasionally the students would not listen at the right time to get the voice-over instructions, and then they would not know how to solve the task. This could be solved by implementing an idle sensor and repeating the instructions if necessary, or by adding a repeat button.

These observations demonstrate that activities can be continually improved in terms of contextualised content and usability. The eventual outcome of the case study has been that a number of local schools have requested the set of science activities and, following final improvements to content and usability, we plan to roll-out the activities to ten schools in May 2013.

B. Teacher support and learning outcomes (teacher-focussed)

In contrast to the challenges of supporting the student with content that is contextualised and usable, the challenges of supporting the teacher with visibility of student progress and learning outcomes require a different style of evaluation.

	A	B	C	D	E	F
1	ActivityDate	ActivityType	AppId	UserId	SessionId	ActivityData
2	2012-09-25 09:41:51.391	connect	Sci Topic 1	group1		
3	2012-09-25 09:42:48.272	started	Sci Topic 1	group1		
4	2012-09-25 09:43:20.592	pass	Sci Topic 1	group1		
5	2012-09-25 09:43:20.904	started	Sci Topic 1	group1		
6	2012-09-25 09:43:42.486	pass	Sci Topic 1	group1		
7	2012-09-25 09:43:42.801	started	Sci Topic 1	group1		
8	2012-09-25 09:44:01.900	pass	Sci Topic 1	group1		
9	2012-09-25 09:44:30.064	started	Sci Topic 1	group1		
10	2012-09-25 09:44:43.838	fail	Sci Topic 1	group1		

Fig. 5. Raw data collected from tablets in the final tests

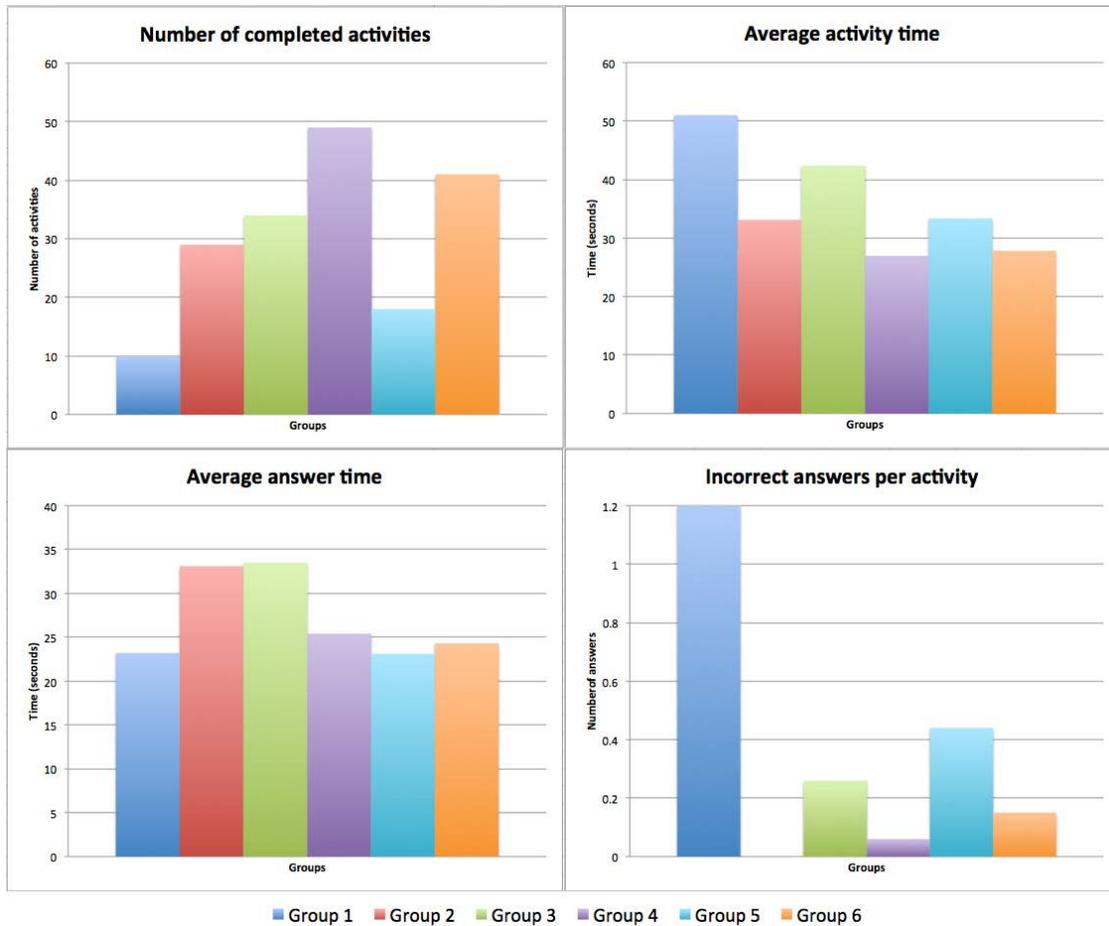


Fig. 6. Data of all six groups of students: (top-left) total number of completed activities per group, (top-right) average time spent to complete each activity (seconds), (bottom-left) average time spends to answer one task (seconds) and (bottom-right) failure rate in each activity (number of incorrect answers per activity).

During the final tests, data on every interaction between student and tablet was collected in real-time on the tablet and sent to the server. The claim to be explored in this section is that the collected data is valuable for supporting the teacher in understanding the progress of the students and the learning outcomes of the class. In order to explore the potential support for the teacher in more detail, the data set was exported for analysis in the form of a spreadsheet as shown in Fig. 5. The data consists of: (A) date and time of the event, (B) the type of event (e.g. “started” activity, “pass” / “fail” activity), (C) the activity topic, (D,E) the user identity, and (F) the activity name.

From simple data analysis, a number of features can be visually extracted through plots and graphs, for example: (i) the total number of activities completed by each student, (ii) the number of correct or incorrect answers given by each student, (iii) the total time spent in each activity; (iv) the average time between the start of an activity and the student giving a correct or incorrect answer (called the “answer time”). The average answer time is different from the total activity time because each activity has multiple tasks (as described in the implementation), and if a task is answered incorrectly then the student must try again. Therefore the average answer time is a

measure of how quickly the student enters an answer irrespective of whether it is correct or not.

As shown previously in Fig. 4, the aim of collecting data is to support the teacher in monitoring the live progress of their class. However, the purpose of extracting the data for further analysis as a one-off exercise is to explore potential solutions to challenges 3 and 4. The eventual goal is to make this extended analysis available to teachers in real-time through a website as in Fig. 4.

The four graphs plotted in Fig. 6 show the data collected from six groups of students at Wat Kung Waree School over a period lasting 30 minutes. The number of activities they completed within the time provided was different depending on their ability, concentration and behaviour. On average there were 30 activities completed by each group and the average time spent in each activity was 33 seconds. From the top-left graph, group 4 completed the most activities, and group 1 completed the fewest. As might be expected, group 1 spent a considerably longer period of time completing activities, with group 4 spending the least time per activity.

According to Fig. 6 top-right, groups 1 and 3 spent more time to complete each activity than the rest of students whereas group 4 spent the least time to complete each activity. An initial analysis of this might conclude that group 4 was the best among all six groups. However, further investigation into the students' average answering time (defined as the time from starting the task to pressing „submit“) and the number of failures made by each group was performed as shown in Fig. 6 (bottom-left and bottom-right). From the graph in the bottom-right, group 1 had the worst failure rate (1.2 fails/activity) among others, whilst group 2 had a zero failure rate (they passed all activities on their first attempt). Failure rate of students in group 5 was also high compared to the rest of the groups. The high rate of failure in group 1 and 5 is valuable information to the teacher who cannot observe all of the students at once in order to tell who is failing. Is the reason for failure because the students' knowledge and understanding of the subject is insufficient? Or were the students over-zealous when they did the activity without any intention to think through the task? Hence the considerations of the students' average answer time, bottom-left graph. Interestingly, the graph shows that both group 1 and 5 spent the least time thinking before they submitted their answer in each activity at 23.2 and 23.1 seconds, respectively. On the other hand, students in group 2 spent 33.1 seconds before they submitted each activity.

Hence, according to the data obtained, we could divide the groups of students into three categories; (i) students who quickly submitted their work/activity and had a high rate of failure – they were group 1 and 5, (ii) students who gently and carefully did their work before they submitted, leading to no or little failure rate – they are group 2 and (iii) students who quickly submitted their work but had so little failure rate still – they are group 4. Among all the groups, this analysis shows that group 4 was the most efficient in terms of time spent and failures. The group that needs most support from the teacher would be group 1.

The failure rate and thinking time might be suitable metrics for analysing overall student behaviour, but the teacher is likely

to also be interested in which activities or topics that the class as a whole found difficult, as it relates to assessing the learning outcomes (challenge 4). Some further analysis was carried out on the data and found that most students had a problem doing exercise named “Life 2.1”. Within this activity, students in group 1 failed three times and group 4 failed one time, whereas students in group 2 and 3 spent 42 and 34 seconds, respectively in completing the activity correctly which was rather long time spent comparing to other's activity time spent. This problem could be because the content relating to Life 2.1 activity was more difficult than others, which would prompt the teacher to consider giving further explanation or examples on the topic. Such insights would enable the teacher to assess the learning outcomes in ways that are not currently possible with the regular OTPC activities. At an even higher level, it could be claimed that the data is valuable for analysing the performance across an entire school, in terms of comparing class achievement and assessing whether achievement matches expectations based on the curriculum. Each of these claims require individual investigation, the observation here is that the current study offers a solution to providing relevant data to the teacher to inform their own decision on the progress of individual students or the entire classroom.

V. CONCLUSION

Within the relatively recent field of mobile learning, prominent scholars Vavoula & Sharples [21] have proposed a framework for evaluating mobile learning „M3“ which includes five precepts for guiding the development of mobile learning innovation. In concluding this article, it is relevant to consider the case study in light of these five precepts in order to reflect on the contribution to a wider mobile learning community. The qualities and shortcomings of the case study in terms of the five precepts are reviewed in Table 2. It shows the reality of the work that still needs to be done to better understand the needs of the OTPC project.

Although BYOD and OLPC schemes are leading large-scale mobile technology facilitation in schools with support nationally and internationally, Thailand's OTPC project is an equally impressive experiment that merits interest from the mobile learning community. While the Thai Ministry of Education focuses on the logistics of distributing and maintaining 800,000 tablets, there is much work needed on addressing the pedagogical challenges. This article highlights, and provides some solutions to, only four challenges that are deemed relevant to current OTPC issues in Thailand. The challenges address issues on both the side of the learner, in providing deeper more engaging learning experiences to the primary school child, and the side of the teacher, in providing relevant support in a classroom environment to assess progress and learning outcomes. Considered in light of the five precepts for mobile learning innovation (as shown in Table 2), considerable work is needed on a much wider range of issues if the impact of OTPC on primary school education is to be better understood. The need is urgent, as the Ministry of Education is preparing a larger roll-out to second year primary school children beginning October 2013. Time will tell whether the OTPC project can have a lasting effect on improving educational in Thailand – certainly it promises to be an exciting opportunity for mobile learning research.

TABLE II. THE FIVE PRECEPTS FOR GUIDING MOBILE LEARNING INNOVATION, TOGETHER WITH SUGGESTED QUALITIES AND SHORTCOMINGS OF THE CURRENT CASE STUDY

Precept	Case study qualities/shortcomings
P1. Capture and analyse learning in context, with consideration of learner privacy	New activities successfully tested in the classroom context, and data captured/analysed from the perspective of the teacher No consideration given to learner privacy
P2. Assess the usability of the technology and how it affects the learning experience	Multiple tests in schools showed improvements in usability Later tests demonstrated increasingly independent learning, but difficult to qualify the improvement
P3. Look beyond measurable cognitive gains into changes in the learning process and practice	Not yet well addressed, but clear that the tablets can change the learning process with both positive and negative (c.f. increased enthusiasm for technology in the classroom vs. decreased student-to-student interaction)
P4. Consider organisational issues in the adoption of mobile learning practice and its integration with existing practices and understand how this integration affects attributes of in/formality	Case study proposes solutions to the challenge of teacher support within a formal classroom environment Hardware and infrastructure issues not yet considered (c.f. effects of usability on classroom management addressed by Lim [22])
P5. Span the lifecycle of the mobile learning innovation that is evaluated, from conception to full deployment and beyond	Case study involved teacher/student input from conception through to final deployment No opportunity yet to compare the new activities with the old activities over a longer period of time – possible future work

ACKNOWLEDGMENT

The authors would first like to thank the three schools that allowed us to undertake the testing and evaluation as well as the many teachers who gave us feedback on the design and content of the activities. Two talented final year undergraduates in computer science, Jenjira Nakrang and Donnapa Yata, developed the storyboards, activities and software. Further thanks go to Me&Friends Studio in Chiang Mai for their guidance throughout the design and implementation.

REFERENCES

[1] C. Shuler, "iLearnII: An Analysis of the Education Category of the iTunes App Store", New York, NY: The Joan Ganz Cooney Center at Sesame Workshop, 2012.
 [2] M. Cohen, "Young Children, Apps & iPad", Research report prepared for U.S. Department of Education Ready to Learn program, 2011. URL: http://mcgrc.com/wp-content/uploads/2012/06/ipad-study-cover-page-report-mcg-info_new-online.pdf
 [3] C. Shuler, "Pockets of potential: Using mobile technologies to promote children's learning", New York, NY: The Joan Ganz Cooney Center at Sesame Workshop, 2009.

[4] L.J. Couse & D.W. Chen, "A tablet computer for young children? Exploring its viability for early childhood education" *Journal of Research on Technology in Education*, 43(1), 75-98.
 [5] S. Tierney, "Bring Your Own Device to school", report by Microsoft Corporation, 2012.
 [6] D. Nugroho & M. Lonsdale, "Evaluation of OLPC programs globally: a literature review (Version 4)", Australian Council for Educational Research, 2010.
 [7] M. Warschauer, "Laptops and Literacy: A Multi-Site Case Study", *Pedagogies: An International Journal*, 3, 52-67, 2008.
 [8] L.L. Linden, "Complement or Substitute? The effect of Technology on Student Achievement in India", Poverty Action Lab, 2008. URL: <http://www.povertyactionlab.org/publication/complement-substitute-effect-technology-student-achievement-india>
 [9] L. Intathep, "One Tablet PC Per Child scheme launch at Government house", 7th June 2012. Bangkok Post. URL: <http://www.bangkokpost.com/news/local/297070/yingluck-wants-tablet-use-tracked>
 [10] National News Bureau of Thailand, "OBEC demonstrates use of tablet PCs for Prathom 1 students", April 4th 2012. URL: <http://202.47.224.92/en/news.php?id=255504040013>
 [11] L. Intathep, "Tablet study finds city, rural schools need different content", 21st February 2012, Bangkok Post. URL: <http://www.bangkokpost.com/news/local/280760/tablet-study-finds-city-rural-schools-need-different-content>
 [12] D. Corlett, M. Sharples, S. Bull & T. Chan, "Evaluation of a mobile learning organiser for university students", *Journal of Computer Assisted Learning*, 21: 162-170, 2005.
 [13] C. Saengpassa & W. Khaopa, "The Fun They Had: Government urged to train teachers on use of computer tablets to enhance pupils' learning", 13th February 2012, The Nation. URL: <http://www.nationmultimedia.com/national/The-fun-they-had-30175709.html>
 [14] W. Van De Bogart, "Child Development Issues Related to Thailand's Tablet Computer Policy within the ASEAN Community", *Proceedings of ASEAN Scenario Analysis 2015-2020 Conference*, Bangkok University, May 25, 2012.
 [15] W. Van de Bogart, "Developing Multi-literacy Skills in e-Learning Environment using Tablet Computers at the Primary Grade Level in Thailand". *Proceedings of The Fourth TCU International e-Learning Conference Smart Innovations in Education & Lifelong Learning*, Muang Thong Thani, Thailand. June 14-15, 2012.
 [16] Mobile Computing Lab. The One-Tablet-Per-Child @ NU project kicks off. April 2003. URL: <http://www.mobcomlab.com/?p=129>
 [17] G. Vavoula, J. Meek, M. Sharples, P. Lonsdale & P. Rudman, "A Lifecycle approach to evaluating MyArtSpace". *Proceedings of 4th IEEE International Workshop on Wireless, Mobile and Ubiquitous Technology in Education, WMUTE '06*. November 16-17, 2006. pp.18-22.
 [18] K. Beck, et al. "Manifesto for Agile Software Development". Agile Alliance, 2002. URL: <http://agilemanifesto.org/>
 [19] V. Brossier, "Developing Android Applications with Adobe AIR", O'Reilly Media Inc, 2012.
 [20] J. Nakrang, D. Yata, & A. Harfield, "Interactive tablet activities for science education in Thai primary schools", *Proceedings of 1st ASEAN Undergraduate Conference in Computing (AUCC)*, Chantaburi, Thailand, February 14-15, 2013.
 [21] G. Vavoula & M. Sharples, "Meeting the challenges in evaluating mobile learning: a 3-level evaluation framework", *International Journal of Mobile and Blended Learning*, 1(2), pp. 54-75, 2009.
 [22] C.P. Lim, Y.H. Teo, P. Wong, M.S. Khine, C.S. Chai & S. Divaharan, "Creating a Conducive Learning Environment for the Effective Integration of ICT: Classroom Management Issues", *Journal of Interactive Learning Research*, 14(4), 405-423. Norfolk, VA: AACE, 2003.

A Remote Health Care System Combining a fall Down Alarm and Biomedical Signal Monitor System in an Android Smart-Phone

Ching-Sung Wang

Department of Electronic
Engineering Oriental Institute of
Technology, Taipei, Taiwan

Chien-Wei Liu

Department of Information
Management St. Mary's College,
Yilan, Taiwan

Teng-Hui Wang

Department of Computer Science
and Information Engineering,
National Taiwan University of
Science and Technology, Taipei
Taiwan

Abstract—First aid and immediate help are very important following an accident. The earlier the detection and treatment is carried out, the better the prognosis and chance of recovery of the patients. It is even more important when considering the elderly. Once the elderly have an accident, they not only physically injure their body, but also impair their mental and social ability, and may develop severe sequela. In the last few years, the continuously developed Android cell phone has been applied to many fields. Despite the nature of the GPS positioning system that the mobile phone currently uses, most applications used are SMS and file transfers. However, these biomedical measurement signals, passing through a transferring interface and uploading to the mobile, result the little really successful cases with the remote health care feasibility. This research will develop an Android cell phone which combines the functionality of an ECG, pulsimeter, SpO₂, and BAD (Body Activity Detector) for real-time monitoring of the activity of a body. When an accident occurs, the signals go through Android smart phone, immediately notifying the remote ends and providing first time help.

Keywords—Health care; Biomedical signal; fall down alarm; Real-time; Android smart phone

I. INTRODUCTION

In 1991, the American Heart Association suggested, the "Chain of Survival" of early access, early CPR, early defibrillation, and early advanced care. It emphasized that immediately after cardiac arrest, early CPR in the first 4 min. and early advance care within the first 8 min would have about a 43% chance of successful rescue rate, the alternative was lower than 20%. Cumnius R.O.'s research points out that there is about a 7-10% decrease in rescue rate success per minute delayed after the patients' heart arrest [1].

In traumatic care, there is the so-called "golden hour" concept that refers to the first sixty minutes after the occurrence of multi-system trauma. It is widely believed that the victim's chances of survival are greatest if they receive definitive care in the operating room within the first hour after a severe injury (Committee on Trauma, 1993; Division of Trauma & EMS, 1992).

For a cerebrovascular accident, it also highlights the "golden 3 hours", in that providing proper care in the first 3 hours after the accident can increase the chance of survival. Despite the kind of emergency, they focus on immediate response and properly provided early detection and early treatment.

Accidents are more significant for the elderly, morbid populations, and for those who have unstable life signs. According to studies, once an elderly person falls down, they would reduce their living activity [2][3], indirectly reducing their quality of life. Due to the habit of using an Android mobile phone, we suggest a tool set that is able to monitor the user's state of health [4][5][6][7]. It operates through Bluetooth SPP (Serial Port Profile) [8] to communicate with the mobile phone. When the BAD (Body Activity Device) detects an abnormal response, it will send an emergency signal to the Android mobile phone. After that, the Android mobile starts real-time biomedical signal monitoring, and efficiently transfers both the GPS and biomedical signals back to the server.

When an accident occurs, it can immediately and effectively provide relative rescue information by the mobile phone and Bluetooth, expecting that is able to improve first aid effect. In addition, it provides the senior population with an effective accident prevention guarantee, enabling them to avoid decreasing their life span because of an accident.

II. SYSTEM ARCHITECTURE

This system is categorized into three main structures, client, server, and First Aid unit. The client includes BAD, a biomedical signal monitoring device (including ECG, SpO₂ monitoring device) [9][10], and Android mobile phone with Bluetooth communication ability. When BAD detects an abnormal health response from the user, it will give the mobile a report by means of the Bluetooth. In the meantime, the operating biomedical device, in sending signals back, allows GPS coordination to take place. These messages would be sent back to server through the mobile phone's internet system. The server end is predominantly used to save the data, to analysis it, and to communicate with others. (Figure 1)

A. Client Side System

This system includes BAD, ECG, SpO₂, and Android mobile phones. They communicate with each other by means of the Bluetooth:

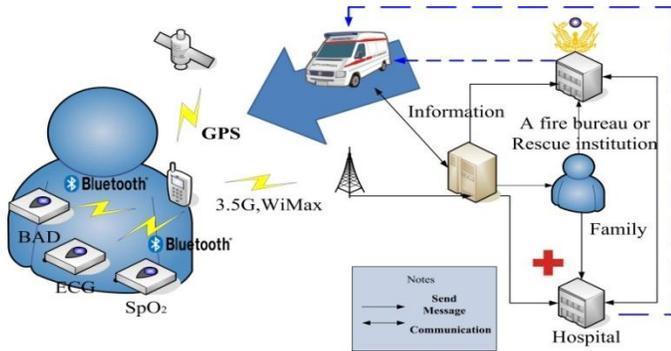


Fig. 1. System Architecture

1) *BAD*: The main function is to monitor the unusual, and the phone immediately sends such messages to the monitoring server through the public mobile network.

2) *Biomedical signal detector*: Uses the already set-up electric circuit and device to detect biomedical signals. After obtaining the signals, it digitalizes them and communicates with the mobile phone to do transferring the device includes ECG and SpO₂ monitoring.

3) *Android smart phone*: Uses Android 2x version [11] smart new mobile technology as a communicating host. It carries a GPS coordinating device and a 3.5 wireless communication system. When an abnormal signal is detected, it immediately sends the user's address and signal back to the server for processing.

B. Server Side System

The server end is composed of a high level server, back end processing program, emergency report system and SQL database. After receiving the emergency signal and biomedical information, the processing plot is shown as in Figure 2.

C. First Aid Clients

The First Aid circuits are divided into four parts, first aid unit, medical unit, family unit, and mobile care unit. After the server receives the signal, it will inform the associated units:

1) *First Aid unit*: The server will provide the position and health information to these units. Such as the fire control unit emergency care unit.

2) *Medical unit*: The local hospital can contact the emergency care unit to access the biomedical information of the user.

3) *Family*: It will contact the user's requested family member. If the system has not contacted other units for any reason, the family member can contact assistance by another medium.

4) *Mobile care unit*: Ambulance, for an example, can access the user's position and biomedical information from the server.

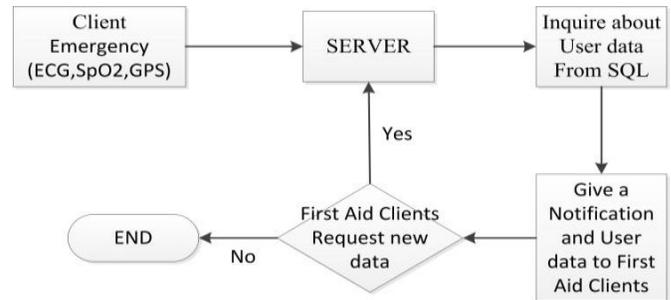


Fig. 2. Server Side System

III. HARDWARE

The whole hardware system, not only the server predominantly the user devices, includes BAD and other communicating devices:

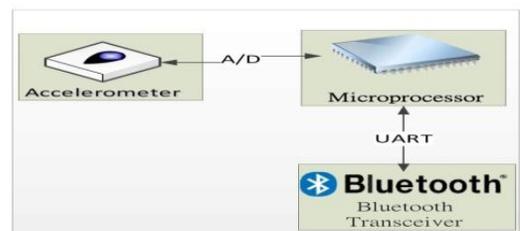


Fig. 3. BAD Hardware Architecture

A. BAD Device

Figure 3 details the block diagram of the BAD hardware architecture. The BAD is mainly built from a microprocessor, an accelerometer, Bluetooth transceiver and other elements. Its function is to detect the body's reaction to signals, abnormal analysis, and the control of signal transfer and receive.

Explanations for each device are listed below:

1) *Microprocessor*: SCI4431 [12] is used as the controller of BAD in this research. SCI4431 contains a RISC (Reduced Instruction Set Computer) micro controller; DSP (Digital Signal Processor), and PCM (Pulse Code Modulation) codec. The micro controller is used to operate the Bluetooth protocol stack, and analyze the accelerometer signals. DSP and PCM codec are used for sound signal operation and decoding.

2) *Accelerometer*: A MXA2500 Dual Axis accelerometer [13] is used in this project. MXA2500 is an electro-mechanical integrated device which can be used to detect the changes in acceleration. The change of acceleration is a control point of the body condition in this research. The MXA2500 communicates with the microprocessor through the A/D (Analog/Digital) interface.

3) *Bluetooth Transceiver*: iWRAP [14] is used as the baseband IC of the Bluetooth transceiver, which communicates with the microprocessor through an UART interface.

B. Communication Devices

Figure 4, represents the Bluetooth composed Pico-Net [15]. It forms communication devices by operating SPP.

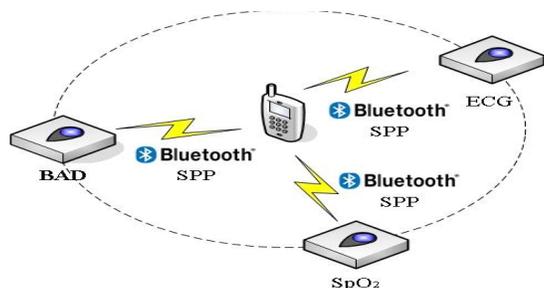


Fig. 4. Bluetooth Pico-Net

Figure 5 presents the system communication plot. BAD, ECG and SpO₂ will shut down the system into a power save mode after Bluetooth and mobile phone pairing. When the BAD detects an emergency, it will resume normal power, and contact the mobile phone to call out ECG and SpO₂ processing.

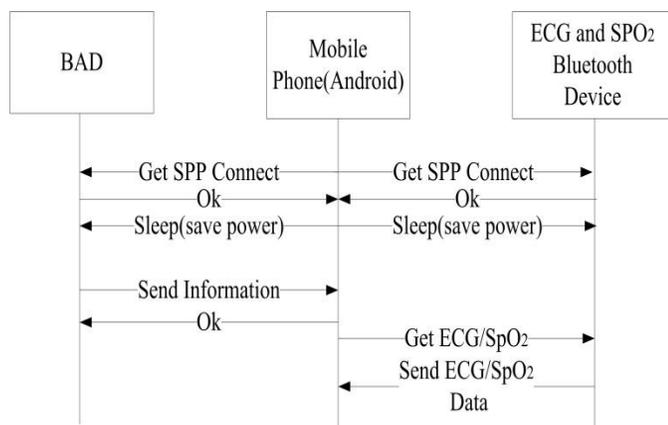


Fig. 5. System Communication

IV. SOFTWARE

This part of the software design includes BAD motion editing and its analysis process. When the system is told to run into the part of emergency procedure, the processing details like before:

A. BAD

BAD in this research contains two different modes according to different body conditions, namely the body stimulation mode and body activity detection mode [15]. The details of which are as follows:

- 1) Body Stress-Reaction Mode Signal Processing
- 2) Body Activity Detection Mode Signal Processing

This mode is opposite to the body stress-reaction mode, which means it detects the body activity changing from the normal level to a level of low activity. It represents that something wrong might have happened, and BAD immediately starts the emergency procedures. A microprocessor reads the signal from the accelerometer every 10 milliseconds, and passes the signal through one order passive HPF, and then calculates the RMS. If the RMS is lower than the preset static RMS threshold for a period of time, it means that an emergency has occurred. The system would beep out the warning sound in the first instance, and if the user does not turn off the alarm, the system would start emergency procedures.

B. Emergency procedure

After receiving the BAD emergency signal, the system will enter the emergency procedure mode, as shown in Figure 6. Firstly, it will recheck the signal. If it is still an emergency signal, it would start the GPS to determine where the user is and identify whether the target moves. If the target stays still, it will start the ECG and SpO₂ Bluetooth devices. After receiving the data, it processes mobile phone and monitoring and displaying and alerts nearby people for further assistance. As the same time, the data is sent back to the server which also informs the First Aid Units to send out further help.

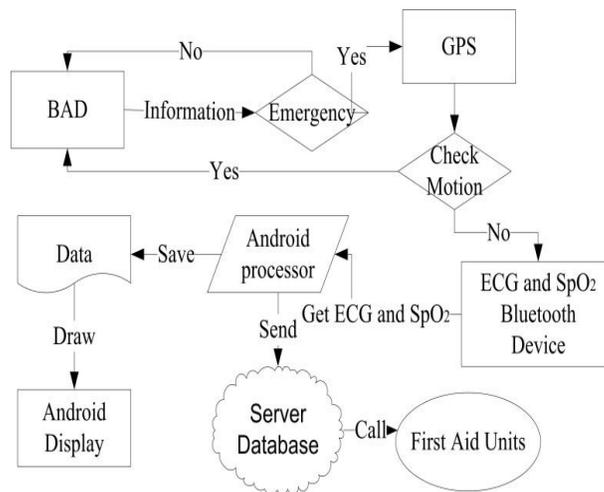


Fig. 6. Emergency Procedure

V. MEASUREMENT RESULTS

After the Android phone receives the signal from BAD and enters the emergency procedure mode, the Android phone will start receiving an ECG signal. For receiving the ECG signal we used a Bluetooth class of Android mobile phone to receive the ECG data with a 16 bit microchip.

This chip edits two groups of 8 bits data and uses a SPP profile to transfer data. Finally it converts the ECG data using formula (1) to determine the value, which is the ECG detection value.

$$ECG = |(Data [0] \times 256 + Data [1]) \times 32.23| \div 10000 \quad (1)$$

When we receive the data, we convert these data to the signal shown in Figure 7. Upon standing up, both psychology and physiology conditions of the user become normal as shown in Figure 8.



Fig. 7. Fall down status



Fig. 8. Normal status



Fig. 9. Server Web-End

The remote and other members through the WEB interface confirm the biomedical information such as shown in Figure 9, the server and receives signal when simulating people falling down. At the same time, it uses GPS to confirm the client's location, for an efficient response time, and sends an ambulance to help. Strives see the Figure 10.



Fig. 10. GPS Location

VI. CONCLUSION

To develop an Android mobile phone transfer function, the proposed design with the fall down alarm and GPS functions provides a simple, practical and portable system of real-time

monitoring of human body activity by BAD, ECG, pulsimeter, SpO2, or other future developed biomedical devices. These devices detects the user being in both a static and dynamic emergency by separate body stress-reaction mode and body activity detection mode, and suggests a solution with ECG, pulsimeter and SpO2, or others, in the emergency situation. In particular this is designed for the elderly and patients who are at risk; we expect that people who carried this device would have more effective care and assistance from others.

REFERENCES

- [1] Cummins RO, Eisenberg MS, Hallstrom AP, Litwin PE. "Survival of out-of-hospital cardiac arrest with early initiation of cardiopulmonary resuscitation", *Am J Emerg Med.* pp.114-pp.119, 1985.
- [2] Albareda JL, Lemieux A, Vellas B, Groulx B. "Psychological factors in falls in elderly patients", *Can J Psychiat* ; 34(2):94-6,1989.
- [3] S. Dagtas, G. Pekhteryev, Z. Sahinoglu, H. Cam, and N. Challa."Real-Time and Secure Wireless Health Monitoring", *International Journal of Telemedicine and Applications*, 2008.
- [4] Stanford, V. "Pervasive health care applications face tough security challenges", *IEEE Pervasive Computing*, April-June, Vol 1, Issue: 2, pp 8-pp12, 2002.
- [5] Sultan S., Mohan P., Sultan N, "Experiences from a Newe-Health Initiative for Patients with Diabetes and Cardiovascular Disease", In 1st IEEE International WoWMoM Workshop on Interdisciplinary Research on E-Health Services and Systems. Kos, Greece, June, 2009.
- [6] Ren-Guey Lee, Kuei-Chien Chen, Chun-Chieh Hsiao, and Chwan-LuTseng,"A Mobile Care System With Alarm Mechanism", *IEEE Transactions on Information Technology in Biomedicine*, vol.11, no.5, September, 2007
- [7] Maki BE, Holliday PJ, Topper AK. Fear of falling and postural performance in the elderly *J Gerontol*; M123-31, 1991.
- [8] Bluetooth Specification, "Serial Port Profile", Version 1.1, February,2001.
- [9] U.Anliker, et al."Amon:A wearable multi-parameter medical monitoring and alarm system",*IEEE Transactions on Information Technology in Biomedicine*, pp415-pp427, 2004.
- [10] Foad Dabiri, Tammara Massey, Hyduke Noshadi, Hadop Hagopian, C.K.Lin, Robert Tan, Jacob Schmidt, and Majid Sarrafzadeh,"A Telehealth Architecture for Networked Embedded Systems: A Case Study in Vivo Health Monitoring", *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 3, pp351-pp359, May, 2009.
- [11] Gotop Information Inc., The experiment of Google Android 2.X Application program development, , October ,2010.
- [12] National Semiconductor Inc., "SC14440/431/432/435/436 /437/438. Baseband Processor for PP/FP DECT and WDCT", VI.1, June 16, 2005.
- [13] Memsic Inc.,"MXA2500G/M Improved Ultra Low Noise ± 1.7 g Dual Axis Accelerometer with Absolute Outputs", Rev. G, 3/6/2007
- [14] Bluegiga Technologies Inc. "iWRAP³ User Guide", 2008
- [15] Ching-Sung Wang, Chien-Wei Liu, Teng-Wei Wang, "Tele-care for emergency announcements "J. Biomedical Science and Engineering, vol.3, no.8, pp.822-pp.827, 2010.

Detecting LinkedIn Spammers and its Spam Nets

Víctor M. Prieto*, Manuel Álvarez* and Fidel Cacheda*

*Department of Information and Communication Technologies, University of A Coruña, Coruña, Spain 15071

Email: {victor.prieto, manuel.alvarez, fidel.cacheda}@udc.es

Abstract—Spam is one of the main problems of the WWW. Many studies exist about characterising and detecting several types of Spam (mainly Web Spam, Email Spam, Forum/Blog Spam and Social Networking Spam). Nevertheless, to the best of our knowledge, there are no studies about the detection of Spam in LinkedIn. In this article, we propose a method for detecting Spammers and Spam nets in the LinkedIn social network. As there are no public or private LinkedIn datasets in the state of the art, we have manually built a dataset of real LinkedIn users, classifying them as Spammers or legitimate users.

The proposed method for detecting LinkedIn Spammers consists of a set of new heuristics and their combinations using a kNN classifier. Moreover, we proposed a method for detecting Spam nets (fake companies) in LinkedIn, based on the idea that the profiles of these companies share content similarities. We have found that the proposed methods were very effective. We achieved an F-Measure of 0.971 and an AUC close to 1 in the detection of Spammer profiles, and in the detection of Spam nets, we have obtained an F-Measure of 1.

I. INTRODUCTION

Currently, the WWW is the biggest information repository ever built, and it is continuously growing. According to the study presented by Gulli and Signorini [1] in 2005, the Web consists of thousands of millions of pages. In 2008, according to Official Blog of Google¹, the Web contained 1 trillion unique URLs.

Due to the huge size of the Web, search engines are essential tools in order to allow users to access relevant information for their needs. Search engines are complex systems that allow collecting, storing, managing, locating and accessing web resources ranked according to user preferences. A study by Jansen and Spink [2] established that approximately 80% of search engine users do not take into consideration those entries that are placed beyond the third result page.

This fact, together with the great amount of money that the traffic of a web site can generate, has led to the appearance of persons and organizations that use unethical techniques to try to improve the ranking of their pages and web sites. Persons and organizations that use these methods are called spammers, and the set of techniques used by them, are called Spam techniques.

There are different types of Spam based on the target client: Web Spam [3] [4] or Email Spam [5]. Web Spam contains several Spam types such as: Blog/Forum Spam, Review/Opinion Spam and Social Networking Spam. Blog/Forum Spam is the Spam created by posting automatically random comments or promoting commercial services to blogs, wikis, guestbooks. Review/Opinion Spam tries to mislead readers or

automated opinion mining and sentiment analysis systems by giving undeserving positive opinions to some target entities in order to promote them and/or by giving false negative opinions to some other entities in order to damage their reputations. Finally, Spam is also becoming a problem in social networks. There are existing studies in the literature about Spam in Video Social Networks [6] or Twitter [7] [8]. There are several features of Social Networks that could make Spam even more attractive:

- The target client is directly the final user. Web Spam is focused on content, so the Web Spammers try to improve the relevance of a web site by, for example, keyword stuffing. When the user conducts a search, it is likely that a Web Spam page will appear. However, it depends on the user clicking on this Web Spam page. Social Networks allow direct Spam, therefore the user will receive the Spam no matter what.
- It is focused on specific user profiles. In the case of Email Spam the content and the products of the email are generic because Spammers do not have information about target users. However, Social Networks (Facebook [9], Twitter [10] or LinkedIn²) allow us to know a great amount of user data, so spammers use this data to aim each type of content or product at a specific audience.
- Social networks contain social network search tools to target a certain demographical segment of users.

This article focuses on Spam in the LinkedIn social network. LinkedIn is a social networking web site for people in professional occupations. It was founded in 2002 and in 2013, LinkedIn had more than 200 million registered users in more than 200 countries.

Although some existing works have been performed to detect Spam in some well-known social networks, to the best of our knowledge this is the first one focused on the LinkedIn social network and it presents a different approach to detect Spam in Social Networks. First, due to the lack of public or private LinkedIn Spam datasets in the state of the art, we have generated one by means of a honeypot profile and searches of the Spam phrases. The process used to create the dataset is explained in Section V.

Second, we have created a method for detecting LinkedIn Spammers. For that, we have analysed the Spammers profiles on LinkedIn, and we have proposed a set of new heuristics to characterise them. Finally, we have studied the combination of these heuristics using different types of classifiers (Naïve Bayes [11], SVM [12], Decision Trees [13] and kNN [14]).

¹<http://googleblog.blogspot.com.es/2008/07/we-knew-web-was-big.html>

²<http://press.linkedin.com/about>

Third, we present a method for detecting Spam LinkedIn nets, that is, to detect sets of fake users created to send Spam messages to the real users connecting with them. This allows filtering those legitimate companies among fake companies, which creates a large amount of profiles for the unique purpose of generating Spam. The method is based on the similarity of their profiles and contacts. For that, the method uses distance functions (Levenshtein [15], Jaro-Winkler [16], Jaccard [17], etc.) which calculate the similarity value of each company.

After performing the experiments, we have determined that for the Spammers detection method the best classifier is kNN, and for the Spam nets detection method the best distance function is Levenshtein.

In short, these are the main contributions of this article: a) a detection method for LinkedIn Spammers, b) a detection method for Spam nets and c) the first LinkedIn Spam dataset.

The structure of this article is as follows. In Section II we comment on the works presented in the literature regarding the different types of Spam, and Spam techniques, as well as the ones that deal with the distinct detection methods. Section III shows the presence of Spam in social networks, specifically in LinkedIn. Section IV explains the two proposed detection methods. In Section V the LinkedIn Spam dataset we have created is explained. Section VI analyses the results obtained detecting Spammers profiles and Spam nets by applying the proposed methods. Finally, in sections VII and VIII we comment on our conclusions and the future works respectively.

II. RELATED WORK

Spam has existed since the Web appeared and it has been growing in importance with the expansion of the Web. Currently, Spam is present in various applications, such as email servers, blogs, search engines, videos, opinions, social networks, etc. Different approaches to Spam detection have appeared [18] [19], however, the best results have been obtained by the methods based on the machine learning approach. Below, we analyse some of the more important articles for the different types of Spam.

There are many articles about Web Spam. Henzinger *et al.* [4] discuss the importance of this phenomenon and the quality of the results that search engines offer. Gyöngyi and Garcia-Molina [3] propose a taxonomy of this type of Spam. Ntoulas *et al.* [20] highlight the importance of analysing the content to detect this type of Spam.

On the other hand, there are studies focused on the detection of Email Spam. Among them, we highlight the work performed by Ching-Tung *et al.* [21]. This article presents a new approach for detecting Email Spam based on visual analysis, due to Spam emails embedding text messages in images to get around text-based anti-spam filters. They use three sets of features: a) embedded-text features (text embedded in the images), b) banner and graphic features (ratio of the number of banner images and ratio of the number of graphic images) and c) image location features. One of the first studies which focused on the detection of Email Spam based on machine learning, was that proposed by Sahami *et al.* [5].

Currently, Spam in social networks is booming, due to the wide use and the easy access to user data. There are several

articles focused on this type of Spam.

Gao *et al.* [22] present an initial study to quantify and characterize Spam campaigns launched using accounts on online social networks. They analyze 3.5 million Facebook users, and propose a set of automated techniques to detect and characterize coordinated Spam campaigns. Grier *et al.* [23] present a characterization of Spam on Twitter. The authors indicate that 8% of 25 million URLs studied point to phishing, malware, and scams listed on popular blacklists. However their results indicate that blacklists are too slow at identifying new threats. In 2010, Wang presented an article [24], where he proposed a Spam detection prototype based on content and graph features. Another interesting articles focused on Twitter Spam, are the ones carried out by Yardi *et al.* [25] and Stringhini *et al.* [26], which study the behavior of Twitter Spammers finding that they exhibit different behavior (tweets, replying tweets, followers, and followees) from normal users (non-spammers).

With respect to the forum/blog Spam, it is necessary to highlight the study performed by Youngsang *et al.* [27]. In this work, the authors study the importance of forum Spam, and the detection of these web pages by using several new heuristics and an SVM classifier. Another study presented by Mishne [28], describes an approach for detecting blog Spam by comparing the language models used in different posts.

In the literature there are other articles related to other types of Spam. An example is the article by Jindal and Liu [29], where they present a detailed analysis about the Spam in the context of product reviews. Mukherjee *et al.* [30] presented an article focusing on detecting fake reviews. The authors propose an effective technique to detect such groups, using the following features: ratio of group size, group size, support count, time window between fake reviews, etc. Lim *et al.* [31] presented another interesting study about this type of Spam. The authors propose a supervised method to discover review Spammers. To achieve that, they identify several characteristic behaviors of review spammers and model these behaviors to perform a ranking of the different reviewers.

Finally, we want to highlight an interesting article performed by Benevenuto *et al.* [6], where the authors propose a method for detecting Spam in video social networks. They use three sets of features: a) quality of the set of videos uploaded by the user, b) individual characteristics of user behavior and c) social relationships established between users via video response interactions.

In this Section, we have presented a wide set of articles related with the different types of Spam and detection methods. Among them are several focused on Social Networking Spam, however, to the best of our knowledge, this is the first study that analyses and detects Spam in LinkedIn. Due to significant differences with other social networks, it is necessary to analyse in detail its characteristics and propose new heuristics to detect Spammers and their Spam nets. Some of its major differences are: it is a professional network, premium accounts allow access to detailed user data or users can be filtered (using search tools) to conduct Spam campaigns. Moreover, its interesting and different characteristics, from the Spammers point of view, make this study both useful and necessary.

	Google+	Facebook	Linkedin	Twitter
Users	+500	+1100	+200	+500

TABLE I: Number of users, in millions, for the main social networks

III. MOTIVATION

Spam is one of the most important challenges on the Web. Currently, due to the boom in social networks, the Spam generated in them is growing constantly. Probably, one of the main reasons for this growth is the large amount of users that they contain. In Table I we show the number of users for: Google+³, Facebook⁴, LinkedIn⁵ and Twitter⁶.

There are several reasons why Spammers are using the social networks. These reasons can be divided into 3 topics:

- Audience:
 - Huge audience (see Table I), this means big profits for spammers, even if only a small percent visit the page or buy the product.
 - It is very easy to create Spam, because social networks allow direct Spam, that is, the Spammer knows the name and data (job, contacts, skills, etc.) of each of its victims. That is the difference with Web Spam or Email Spam where the Spammer does not know these data about its victims, only the email and perhaps the name.
- Different options and tools to create Spam:
 - Fast distribution of Spam, due to user trust and their curiosity. The users trust anything that they see posted by one of their contacts. An example of this, is the use of popular hashtags in Twitter to lure users to their Spam sites.
 - This type of Spam allows the creation of fake relationship contacts, to make the user's profile appear more real on the social network.
 - Spammers can send messages to the users and include embedded links to pornographic or other product sites designed to sell something.
 - Internet social networks contain common fan pages or groups that allow people to send messages to a lot of users even if the Spam user does not have these users as contacts.
- Little investment. Unlike other types of Web Spam, which requires investing in domains, hosting, developers, etc., social network Spam only needs accounts in social networks such as LinkedIn, Facebook or Twitter.

A. How much Spam is there in LinkedIn?

Currently, LinkedIn Spam presents a significant problem for its users. A large amount of forums and blogs show their grievances regarding this type of Spam and offer some

advice to users. Public reports of Panda Security⁷ about the intense Spam campaigns in LinkedIn [32], or the fake emails of LinkedIn to exploit Java and Adobe vulnerabilities, show the importance of the problem [33].

One approach to determine the incidence of one problem or topic in society is to measure its impact in Google searches (or other important search engines). This method has already been used to study the presence of certain diseases in society [34].

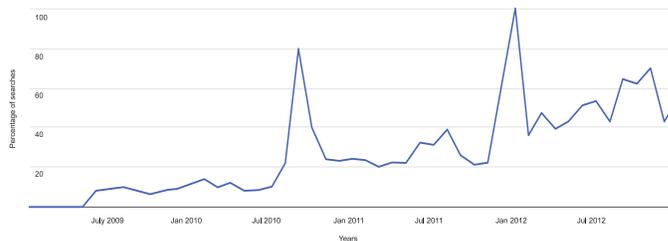


Fig. 1: Trend of the LinkedIn Spam

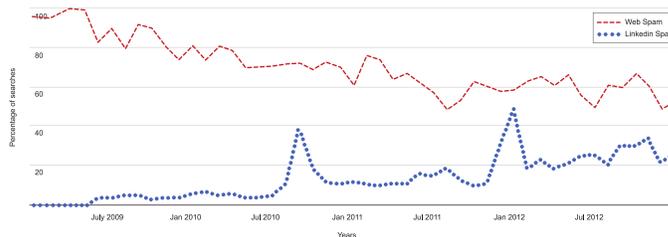


Fig. 2: Web Spam versus LinkedIn Spam

In our case, we have used two methods. First, we have measured the trends in the Google searches using Google Trends. Figure 1 shows the obtained results. It depicts the relative amount of searches by year according to the maximum value obtained in 2012 (100%). With the same tool, we have also analyzed the importance of Web Spam versus LinkedIn Spam. The results shown in Figure 2, show that the importance of Web Spam (the most important type of Spam) is decreasing compared to the increase in LinkedIn Spam.

On the other hand, we have used another approach: searching for the query in Google *LinkedIn Spam*. The number of results is higher than 53 million pages, which indicates the high presence and concern about social networking Spam.

³<http://googleblog.blogspot.com.es/2012/12/google-communities-and-photos.html>

⁴<http://investor.fb.com/releasedetail.cfm?ReleaseID=761090>

⁵<http://blog.linkedin.com/2013/01/09/linkedin-200-million/>

⁶<http://techcrunch.com/2012/07/31/twitter-may-have-500m-users-but-only-170m-are-active-75-on-twitters-own-clients/>

⁷<http://www.pandasecurity.com>

B. Can existing Spam detection techniques be used on LinkedIn?

Before analysing the existing Spam detection techniques we will explain the different types of Spam. There are different classifications for them, but, the classifications most relevant in the state of the art, performed by Gyongyi and Garcia-Molina [3], and Najork [35], suggest that the main types of Web Spam are: content spam, cloaking and redirection spam, click Spam and link spam.

- Content Spam, that is, a technique based on the modification of the content or the keywords of a web page with the purpose of simulating more relevance to search engines and attract more traffic. To detect this type of Spam the search engines use several algorithms and heuristics based on content analysis. However, there are a lot of Spam pages that avoid these detection algorithms. An example of these heuristics was presented by Ntoulas *et al.* [20].
- Cloaking and Redirection Spam, which consists in dynamically generating different content for certain clients (e.g.: browsers) but not for others (e.g.: crawling systems). There are several techniques to detect this type of Spam, among them we can highlight [36], [37] and [38]. The latter approach, proposed by Wu and Davison [38], where the authors detect this Spam by analysing common words across three copies of a web page.
- Click Spam: this technique is based on running queries against search engines and clicking on certain pages in order to simulate a real interest from the user. Search engines use algorithms to analyze certain logs clicks and detect suspicious behaviours.
- Link Spam, which is the creation of Web Spam by means of the addition of links between pages with the purpose of raising their popularity. It is also possible to create "link farms", which are pages and sites interconnected among themselves with the same purpose. In order to detect Link Spam, the search engines analyse relationships and the graphs between web domains. In this way, they can detect domains with a number of inlinks and outlinks or web graphs suspected of being Spam.

Finally, there is another detection technique which is not focused on a unique type of Spam. It was proposed by Webb *et al.* [39] [40], and the authors analyse the HTTP headers and their common values in Spam pages, to detect Spam.

As we can see, the existing detection techniques cannot be applied to detect LinkedIn Spam. Only the algorithms for detecting Content Spam could be used, however, the problem is that the existing heuristics to detect Spam in a typical web page, cannot be applied to web page profiles of LinkedIn because their features are completely different. Due to existing detection techniques being impossible to apply, LinkedIn has proposed its own particular detection techniques (see Section III-D).

C. How do Spammers create LinkedIn Spam and what are the differences between it and other social networks?

The first step for a Spammer is to decide whether the Spamming attack will be a focused or general attack. In the case of it being a focused attack, the Spammer will search for specific users by means of the LinkedIn tools. After that, the Spammer can create Spam by the following methods:

- Messages: these are sent by any one of our contacts. However, we often accept contacts because of having contacts in common, or because we think that the job, groups or skills of this user are appropriate for them to be our contact, and perhaps, it could be a job opportunity.
- Groups: that is, those notifications sent to the groups of each of the victims. These messages will be sent by email to the users of the group, and moreover, this post can be seen in the forum of the group.
- Updates: the Spammer makes updates to their profile to invite their contacts to visit their profile.

After we have analysed the operation of LinkedIn Spam, we explain the differences in the creation Spam between LinkedIn and other social networks.

- LinkedIn allows the use of social networks search tools to target a certain demographical segment of the users. This allows the Spam to be made more specific, and therefore, it is likely that the victim will click on the Spam link.
- LinkedIn is a social network which focuses on business, companies and professionals, which is very interesting from the point of view of the Spammers. In other words, the possible profit will be higher if the person that visits the Spam site is a businessman instead of a teenager from Facebook. Twitter, Facebook or email can be used for professional ends, however they do not contain the other advantages of LinkedIn.
- LinkedIn allows direct Spam, such as Twitter, Facebook or email, but in this case, the Spammer knows the name and data (job, contacts, skills, location, etc.) of each of their victims. So, the probability of success is higher in LinkedIn than in Web or Email Spam.
- Due to LinkedIn being defined as a professional social network, usually the data of its users are real, and the usage of it by them is usually a way to find a job or to find new professional contacts; this is not a game. For this reason, when a LinkedIn user receives a message, email or update from LinkedIn, they pay more attention to it than to notifications of other social networks. Again, the probability of Spam success is higher.

In summary, the LinkedIn Spam is made by means of emails or messages to the victims, in their interest groups or in the updates of the Spammer. So, as we have said, the Spam detection techniques based on content analysis could be used, however new heuristics would have to be created specifically for this environment. As a new approach, we propose the detection of Spammers and their Spam nets, instead of Spam

in a particular email, message or comment, which is more difficult. So, if we can detect a Spammer, we can also detect all their Spam messages (emails, comments and updates).

D. How does LinkedIn detect its Spam?

Currently, LinkedIn uses two techniques to detect Spam. On the one hand, when a user receives an invitation to become a contact of another user, he can indicate that this person is a Spammer. A LinkedIn user has numerous methods of contacting a specific user (as a friend, as a coworker, as a classmate, etc.). LinkedIn blocks a contact method to a user profile (Spammer), when it has received 5 requests rejecting said account by the same contact method indicating that it is Spam. Alternatively, the user can report the profile of the Spammer to the following e-mail address: `abuse@linkedin.com`.

However, from our opinion, these methods are not sufficient, due to two reasons. First, people are lazy, and because of that they will usually not accept this person but also will not usually notify that the user is a Spammer. For the same reason, only in a few cases, the user sends an email to report a Spammer. The second reason is because of the speed and ease with which criminal organizations and Spammers can create a lot of accounts, compared to the slow detection methods used by LinkedIn.

In summary, as we have explained, the presence and concern of Spam in social networks is high. Due to this, together with the lack of articles about LinkedIn Spam, existing methods for detecting Spam cannot be applied and the need for other methods to complement tools used by LinkedIn, we propose a method for detecting LinkedIn Spammers, and a method for identifying fake companies (Spam nets).

IV. DETECTION METHODS

We present two detection methods, one for detecting Spammers and another for detecting Spam nets, both in the LinkedIn social network. In order to know and understand the behaviour of LinkedIn Spammers and Spam nets, and also to test the proposed methods, we have manually built a dataset of LinkedIn profiles, classifying them as spammers and legitimate users (see Section V).

The method for detecting Spammers (section IV-A) is based on a set of new heuristics together with the use of machine learning. The heuristics have been obtained by means of the manual and statistical analysis of the legitimate and Spam LinkedIn profiles. So, we characterize a LinkedIn profile, and then decide whether or not it is Spam. For the appropriate combination of these heuristics, we have tried different classification techniques (decision trees, techniques based on rules, neuronal networks and kNN).

To detect Spams nets in LinkedIn (section IV-B), we have focused on the idea that we have observed during the manual analysis. The fake profiles of a fake company, usually share similarities that allow differentiation between legitimate companies and fake companies.

A. Method to detect LinkedIn Spammers

We will discuss a set of heuristics that aim to characterize and detect Spam profiles in LinkedIn. Some of the features we present below, have appeared because LinkedIn is a professional social network, and their users are very careful with the details of their profiles. LinkedIn users want to have an updated and complete profile.

The results obtained for each heuristic were tested on the dataset described in Section V. For each non binary feature, we include a figure showing a box and whisker diagram with the feature values, corresponding to Spam and Non-Spam pages. For binary features, we only present the percentage of use for each type of page.

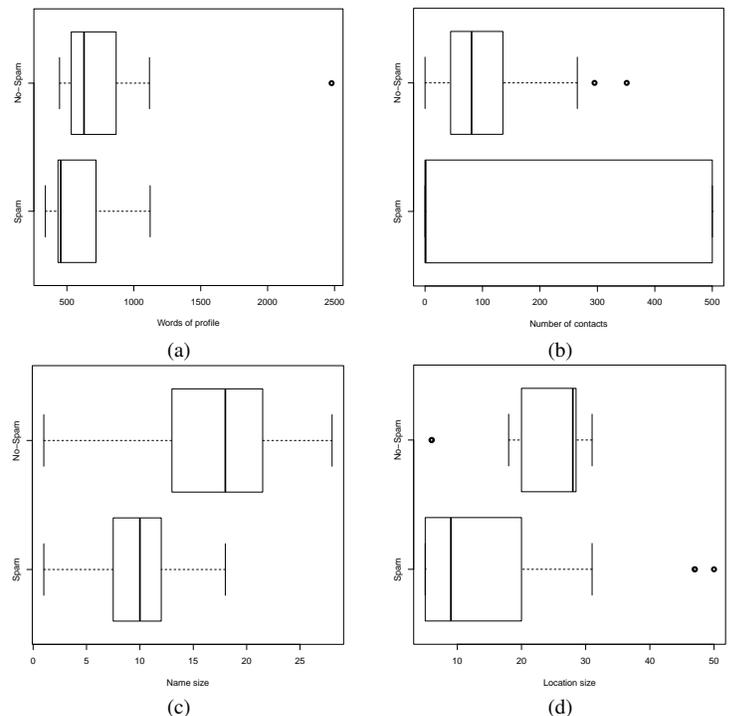


Fig. 3: Number of words (a), contacts (b), name size (c) and location size (d) in the Spam and No-spam profiles

The features analysed are the following:

- Number of words in profile: we have analysed this feature because during the manual labeling of the pages we have observed, that Spam pages usually contain less words than the legitimate LinkedIn profiles. Figure 3a shows that the median number of words in Spam profiles is 454 words. In other words, the Spam profiles contain on average 559.3 words and the No-Spam profiles 711.8 words, 24.6% lower.
- Number of contacts: due to the automatic generation of Spam profiles, they present two clear behaviors: profiles with very few contacts or profiles with many contacts. On the other hand, the legitimate profiles follow a uniform distribution of contacts, without these extreme differences. We observe in Figure 3b that in Spam profiles the median is 1 and in the No-Spam profiles it is 81. Moreover, the difference

between averages is very significant. Specifically, Spam profiles contain on average 204.8 contacts, while No-Spam profiles contain only 98.1.

- Name size: in this case, the deficiency of Spam profiles appears in the name of the person. Fake profiles usually contain shorter names and surnames than in legitimate profiles. This is because LinkedIn users are very careful and want their data profile to be correct and updated. To achieve this, they use their complete name and do not tend to use short names or nicknames.

As we thought, the results indicate that legitimate profiles have longer names than fake profiles. Figure 3c shows that the median and average in Spam profiles is 10 and 9.09 letters and in No-Spam profiles is 18 and 17.41, respectively.

- Location size: we have observed that Spam profiles usually contain a simple and smaller location than in the legitimate profiles. Moreover, the location among the fake profiles of the fake companies are very similar.

In Figure 3d shows the median and average location size in Spam profiles to be 9 letters and 14.72, respectively. In the case of No-Spam profiles, the median is three times higher, than the Spam profiles, 28 letters, and the average is 24.64 letters.

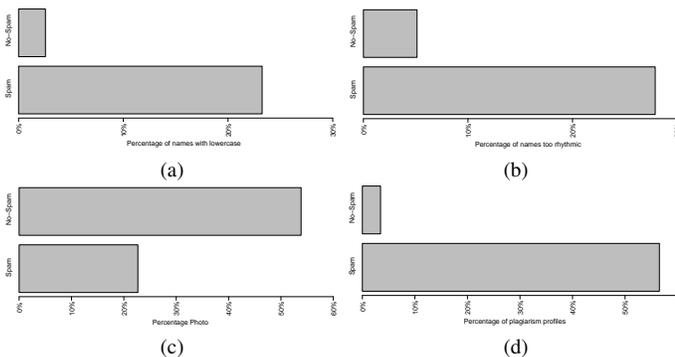


Fig. 4: Percentages of the names written in lowercase (a), percentage of rhythmic names (b), percentage of profiles with photo (c) and percentages of the plagiarism profiles (d) in the Spam and No-spam profiles

- Name written in lowercase: another big weakness of fake profiles, is that their name or surname, are often written in lowercase. Figure 4a indicates that more than 20% of the Spam profiles contain user names in lowercase, and in the legitimate profiles this value is almost 10 times smaller.
- Rhythmic name: a technique used to draw the users attention and build trust in the profile. Specifically, it was observed that often the Spam profiles contained people whose first and last names start with the same two or three letters. Figure 4b shows that the 5.12% of the No-Spam profiles contain a rhythmic name and lastname, but this figure raises up to more than five times, 27.90, in the Spam profiles.

- Profile with photo: we have noted two weaknesses in the fake LinkedIn profiles regarding this issue. First, in a social network the users usually have photo, in the case of the Spam profile, they usually do not. And second, if a Spam profile contains a photo, this photo can usually be found by a search engine. In Figure 4c we observe that only 22.67% of the Spam profiles contain photo, and in the case of legitimate profiles this value is more than double, specifically 53.84%.
- Plagiarism in profiles: another weakness of automatically generated Spam profiles, is that their content is small, or, due to the difficulty in generating logical content, the Spammers take texts from the Internet. We have used the Grammarly plagiarism checker⁸. This system finds unoriginal text by checking for plagiarism against a database of over 8 billion documents. The Figure 4d shows that multiple Spam profiles have copied or automatically generated content. The differences among the results are very significant, specifically 3.45% in No-Spam profiles and 56.53% in Spam profiles.

As we have seen, Spam profiles tend to be simpler and contain less detail than legitimate profiles. Moreover, the results obtained for each type of profile (Spam and No-Spam) show significant differences between them. In multiples cases the results are 2, 3, 5 or even 10 times higher or smaller in Spam profiles than in legitimate profiles. These important differences allow the proposed heuristics to be used to characterize and detect LinkedIn Spam.

The detection method proposed uses these heuristics together with machine learning techniques to identify Spammer profiles. The method is to not focus on a specific heuristic but to use all of them. In the case of failure of a particular heuristic, since the method uses all heuristics, the other heuristics will correct this error. For the appropriate combination of heuristics we have tried different machine learning techniques (decision trees, kNN, SVM and Naïve-Bayes). Based on the obtained results (see Section VI), the method for combining the heuristics is kNN.

B. How to Detect Spam Nets?

We have studied a method for detecting fake companies, companies whose unique purpose is to create fake profiles to generate Spam. The proposed method identifies this type of LinkedIn companies based on the similarity among the profiles that each company contains.

To measure the similarity between profiles, we have generated a text string that contains the different data of the profile separated by commas ", ". A generic example of the text string generated with the LinkedIn data profile and the results obtained with the proposed heuristics, is the following:

```
UserName, ProfileTitle, Location, NumberofContacts, Skills, Education, NumberOfWords, NameSize, LocationSize, RhythmicName, Photo, LowercaseName, Profile-Plagiarism
```

⁸www.grammarly.com/Plagiarism_Check

The value of the user name, title of the profile, user location, number of contacts, skills and education are extracted directly from the profile of the user. However, the variables: NumberOfWords, NameSize, LocationSize, RhythmicName, Photo, LowercaseName and ProfilePlagiarism are calculated previously, based on the analysis of the profile. Finally, the value of RhythmicName, Photo, LowercaseName and ProfilePlagiarism are boolean.

As a preliminary step, we specify the following concepts to help us formally define our method:

- Let be $pi = Profile\ of\ the\ user\ i$.
- Let be $N_p = Number\ of\ profiles\ of\ a\ company$.
- Let be $distance_{ij} = Similarity\ between\ the\ profiles\ i\ and\ j$.
To obtain this value, we have studied different functions:
 - Levenshtein [15]: is the minimum number of edits needed to transform one string into the other (using insert, delete or replace operations). This distance is usually denoted as edit distance.
 - Jaro: is a similarity function which defines the transposition of two characters as the only permitted operation to edit. The characters can be a distance apart depending on the length of both text strings.
 - Jaro-Winkler [16]: is a variant of the Jaro metric, which assigns similarity scores higher to those words that share some prefix.
 - Jaccard [17]: it defines the similarity between two text strings A and B as the size of the intersection divided by the size of the union of the corresponding text strings.
 - Cos TF-IDF [41]: given two strings A and B, and, $\alpha_1, \alpha_2 \dots \alpha_K$ and $\beta_1, \beta_2 \dots \beta_L$ their tokens respectively, they can be seen as two vectors, V_A and V_B , with K and L components. So, the similarity between A and B, can be calculated as the cosine of the angle of these two vectors.
 - Monge Elkan [42]: given two strings A and B, and, $\alpha_1, \alpha_2 \dots \alpha_K$ and $\beta_1, \beta_2 \dots \beta_L$ their tokens respectively. For each token α_i there is a β_j with maximum similarity. Then the Monge Elkan similarity between A and B, is the average maximum similarity between a couple (α_i, β_j)
- Let be S_{pi} the similarity of a profile, pi , with the other profiles of his company. This value is calculated as the sum of the distances between the text string of the corresponding profile with the text strings of the rest of profiles, divided by the number of profiles, N_p , minus 1.

$$S_{pi} = \frac{\sum_{j=0}^{N_p-1} distance_{ij}}{N_p - 1}$$

We now can define the method we propose to obtain the value, S_c , that summarises the similarity of a specific company.

S_c is calculated as the average of the similarities of the profiles, S_{pi} , of the staff of the company.

$$S_c = \frac{\sum_{i=0}^{N_p} S_{pi}}{N_p}$$

After we have obtained the value of similarity of a company, we have to decide if said company is fake or legitimate. In order to do that, we have calculated a threshold for each similarity function. These thresholds allow us to decide when a company contains very similar profiles, and this company will likely be fake, or conversely, the profiles are different enough to be a legitimate company. For that, we have created a training set that contains 4 fake and 4 legitimate companies with the highest number of profiles. In Table II we show the similarity results obtained in this training set. Among the results obtained, we have selected as thresholds those that have obtained the best results (precision, recall and F-Measure) in the training set. The thresholds selected were used to obtain the results (precision, recall and F-Measure) of the method in the full dataset.

Analysing the results we can see that there are differences about the similarity values obtained by each technique. Levenshtein and Cos TF-IDF have obtained the lower values of similarity, around 0.6 and 0.5 respectively. In the other hand, Jaccard obtains the highest results, close to 1. Jaro and Monge Elkan have obtained intermediate results, with values between 0.7 and 0.8. The results obtained by Jaro and Jaro-Winkler, as we expected, are different. This is because Jaro-Winkler scores words which share some prefix higher and we had observed that the string created contains prefixes that increase the similarity result.

If we study the results of the legitimate companies and the fake companies separately, we can observe that, as we thought, fake companies display more similarity between them than the normal companies. This fact can easily be seen in the results obtained by Levenshtein and Cos TF-IDF measures.

In Section VI-C we present the precision, recall and F-Measure applying the proposed method on the created dataset.

V. LINKEDIN SPAM DATASET

To the best of our knowledge, there is no public dataset of LinkedIn profiles. The dataset we have built contains legitimate and Spam profiles. The creation of the dataset was carried out during 30 days, from October 30th to November 30th, 2012.

For the legitimate profiles we have used profiles of users in well known companies (Google, Microsoft, Oracle, Twitter, IBM, etc.). The gathering process of the legitimate profiles was made automatically by means of the LinkedIn API⁹. We have used our LinkedIn profiles to obtain user profiles of the legitimate companies. The process starts in the public profile of an employee of a legitimate company and continues by the contacts of this user who works in the same company.

On the other hand, we have identified a set of Spam users, based on the Spam messages that they send to other users, and the profiles obtained by searching for words that

⁹<https://developer.linkedin.com/>

		Levenshtein	Jaro	Jaro-Winkler	Jaccard	Cos TF-IDF	Monge Elkan
Dinowill	Fake	0.745	0.838	0.903	0.993	0.640	0.840
Innovabiz	Fake	0.693	0.768	0.861	0.995	0.551	0.791
Online Pharm	Fake	0.739	0.706	0.824	0.994	0.630	0.819
Pharmacy	Fake	0.723	0.743	0.845	0.991	0.566	0.792
Cisco	Legitimate	0.571	0.730	0.838	0.996	0.403	0.778
Google	Legitimate	0.640	0.736	0.841	0.994	0.477	0.791
Hp	Legitimate	0.639	0.745	0.825	0.995	0.489	0.756
Motorola	Legitimate	0.662	0.756	0.850	0.995	0.539	0.759

TABLE II: Similarity of the analysed fake and legitimate companies

commonly appear in Spam comments in Google, such as "viagra", "growth hormone", "cialis", etc. The list of these words has been obtained by searching Spam words in the WordPress Codex¹⁰, the online manual for WordPress. After this, we created a fake profile on LinkedIn, as a honeypot, and sent contact requests to the Spam profiles. All the contact requests were accepted, and usually the Spammers responded 1 or 2 days after the request. Once accepted, as we thought, we could detect new fake profiles among their contacts. The labeling and gathering process of each fake profile was made manually due to the need to check whether each profile was really a fake.

We want to clarify that to obtain legitimate profiles we have not associated the created fake profile with the reputed companies. Both legitimate profiles and fake profiles have not been published anywhere and they have been stored encrypted.

To know more details about the generated dataset, we explain its structure:

- Its size is 1,4 GB.
- It contains 750 profiles, 250 Spammers and 500 legitimate users.
To decide the size of the dataset, and the subsets (Spammers and legitimate users), we have had a problem because there are no studies about how much Spam there is in LinkedIn. So, we have decided that the number of Spam profiles is similar to Spam contained in .com domain [20]. Moreover, we have used a percentage of Spam higher than .com domain, because we want to increase the variability of the profiles in order to make the detection process more difficult.
- The profiles are divided among 150 companies, of which 50 are fake companies and 100 are legitimates.
- Among the fake companies are companies focused on different drugs like Viagra or Cialis, and on Chinese products.
- 80% of the legitimate companies are focused on technology and computer science area.
- Companies are mainly located in USA.
- The size of the legitimate companies is variable, ranging from companies with 500 employees to those with more than 20,000.

The provided data about the used dataset allows other researchers to test and compare our methods. It is likely that

their dataset does not contain the same profiles as our dataset, however, in our opinion, they can create a dataset with the same structure and features, and therefore, the results should be very similar.

VI. EXPERIMENTAL RESULTS

In this Section, we discuss all the issues we found for both the execution and the assessment stages, and we show and analyse the results obtained. First, we show the results obtained when detecting Spammers in LinkedIn using the proposed heuristics (see Section IV-A), and then, the results obtained to detect Spam groups or companies, analysing the similarity between the company user profiles.

A. Experimental Setup

To execute the different classifiers, we used WEKA [43], a tool for automatic learning and data mining, which includes different types of classifiers and different algorithms for each classifier. The techniques tested were: SVM, Naïve Bayes, Decision Trees and Nearest Neighbour. To obtain the results, we have used the default parameters in WEKA for each of the machine learning algorithms, specially the value k used in kNN has been 1.

To evaluate the classifier we used the "cross validation" technique [44], that consists in building k data subsets. In each iteration a new model is built and assessed, using one of the sets as a "test set" and the rest as "training set". We used 10 as the value for k ("ten-fold cross validation").

The dataset used to obtain the results was created by us, because there is no public dataset of LinkedIn profiles. The method used to generate it and its characteristics were explained in Section V.

B. Results for LinkedIn Spammers

This section discusses the results obtained applying the proposed method to discover LinkedIn Spammers. Table III shows the precision, recall and F-Measure for each of the types of classifiers studied.

	Precision	Recall	F-Measure
Naïve Bayes	0.909	0.908	0.909
SVM	0.837	0.831	0.794
Decision Trees	0.967	0.967	0.967
kNN	0.969	0.979	0.971

TABLE III: Results of the proposed heuristics using different types of classifiers

¹⁰http://codex.wordpress.org/Spam_Words

The analysis of the results shows that the best results are obtained using kNN and the worst with SVM. We also observe that the results of kNN and decision trees are very similar. Specifically, kNN obtains a precision and a recall of 0.979, and an F-Measure of 0.971. These results are similar to those obtained by applying decision trees, where the recall, precision, and therefore, F-Measure is 0.967. This fact, that the three measures are equal occurs because the number of false positives and false negatives are the same. In this case, this occurs because, based on the results, the number of false positives and false negatives are very small, and it is likely that these values match.

On the other hand, SVM achieves 0.837, 0.831 and 0.794 of precision, recall and F-Measure, respectively. Despite these small differences we observe that the results are very satisfactory.

Another issue analysed to determine the right performance of the proposed heuristics is to study the ROC curve of the classifiers. In Figure 5 we show the ROC curve obtained by each of the studied classifier. Again, the best results are obtained using kNN and decision trees. Analysing the area under de ROC (AUC), kNN and decision trees achieves 0.984 and 0.976, respectively, that is, almost a perfect result. On the other hand, Naïve Bayes obtains an AUC of 0.934. Finally, SVM achieves the worst result, with an AUC of 0.629, that means that although, this classifier has obtained a good precision and recall, it is not reliable.

In short, the precision, recall and the ROC curve indicate that the proposed heuristics are very adequate to detect Spammer profiles in LinkedIn. Moreover, we have detected that the best classifier for this type of Spam is kNN or decision trees.

In general, the obtained results are very hopeful. In our opinion, this fact is in part due to the fact that Social Networking Spam is still at an early stage.

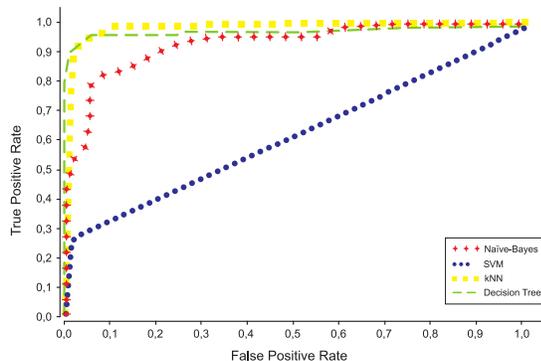


Fig. 5: ROC curve of the tested classifiers

C. Results for Spam LinkedIn Nets

In this section we discuss the results obtained to detect fake companies in LinkedIn, whose only purpose is to generate Spam. For that, we have applied the method explained in Section IV-B.

The results shown in Table II were used as training set to select the thresholds with best results. The proposed

method used the selected thresholds in all labeled companies, legitimate and fake, and we have calculated the precision, recall and F-Measure for each of them. The thresholds used and their corresponding results are shown in Table IV.

	Threshold	Precision	Recall	F-Measure
Levenshtein	0.693	1	1	1
Jaro	0.743	0.750	0.750	0.750
Jaro-Winkler	0.845	1	0.750	0.857
Jaccard	0.993	0.571	1	0.727
Cos TF-IDF	0.551	1	1	1
Monge-Elkan	0.791	0.8	1	0.8

TABLE IV: Results to detect Spam LinkedIn nets

The best results are obtained using Levenshtein and Cos TF-IDF. In these two cases the detection is perfect, with an F-Measure of 1. The results indicate that the method used by these two techniques to measure the similarity between profiles, is the most adequate for the text string generated by our method.

In the analysis of the results, we have detected two types of Spam profiles. On the one hand we have very simple Spam profiles and with little content, and, on the other hand, complex Spam profiles with too much content (in most of cases, automatically generated). To detect this second type all the proposed heuristics must be used because, otherwise, it could be skipped.

However, the fake companies only contain one of these two types of Spam profiles and always with very similar content. In short, the method proposed to detect Spam nets has achieved hopeful results, mainly due to two facts: a) the idea proposed is right, and the fake companies contain similar profiles and b) the Social Networking Spam is relatively new and its techniques are unsophisticated. In the future, it is likely that these techniques will improve. However, we have demonstrated that the proposed idea works perfectly, and can be used, in the future, as a base to be complemented with other new techniques.

VII. CONCLUSIONS

The presence of different types of Spam (Web Spam, Email Spam, Forum/Blog Spam and Social Networking Spam) on the Web is important and is constantly growing. There are many studies that analyse and present techniques for the detection of different types of Spam. However, to the best of our knowledge, there are no studies about Spam in LinkedIn.

In this article, we present a method to detect Spammers and Spam nets in LinkedIn social network. We have proposed a set of heuristics that characterize LinkedIn Spam profiles and help to identify LinkedIn Spammers. These heuristics were used as input to several classification algorithms (Naïve Bayes, SVM, Decision Trees, kNN). The best results are obtained by kNN and decision trees, with an F-Measure of 0.969 and 0.967, respectively, and an AUC close to 1.

Moreover, we have proposed a method for detecting Spam nets in LinkedIn. It is based on the idea that the profiles of fake companies share multiple similarities. The method calculates the similarity between different profiles of the companies,

using several distance functions (Levenshtein, Jaro, Smith-Waterman, etc.). The values of similarity obtained are used as thresholds to detect fake companies (Spam nets) among the legitimate companies. Again, the results are also very hopeful. We have achieved an F-Measure of 1 using Levenshtein and Cos TF-IDF.

In short, the results obtained in the study show that, on the one hand, the heuristics proposed are adequate to detect Spammer profiles, and, on the other hand, the new method proposed to detect Spam nets (fake companies) in LinkedIn performs very well.

VIII. FUTURE WORKS

Spam in LinkedIn social network is a relatively new Spam type, so our intention is to follow its evolution over time. Due to the continuous changing of Spam techniques, we want to find new and better heuristics to detect this type of Spam. Furthermore, we plan to increase and improve our labeled dataset. Finally, we will test the proposed heuristics and the method for detecting Spam LinkedIn nets, in other social networks.

ACKNOWLEDGMENT

This research was supported by Xunta de Galicia CN2012/211, the Ministry of Education and Science of Spain and FEDER funds of the European Union (Project TIN2009-14203).

The list of domains “es” of 2009, from which we conducted the crawling process, has been provided by the Spanish Business Public Entity Red.es.

REFERENCES

- [1] A. Gulli and A. Signorini, “The indexable web is more than 11.5 billion pages,” in *Special interest tracks and posters of the 14th international conference on World Wide Web*, ser. WWW '05. New York, NY, USA: ACM, 2005, pp. 902–903.
- [2] B. J. Jansen and A. Spink, “An analysis of web documents retrieved and viewed,” 2003.
- [3] Z. Gyongyi and H. Garcia-Molina, “Web spam taxonomy,” Stanford InfoLab, Technical Report 2004-25, March 2004.
- [4] M. R. Henzinger, R. Motwani, and C. Silverstein, “Challenges in web search engines,” *SIGIR Forum*, vol. 36, pp. 11–22, September 2002. [Online]. Available: <http://doi.acm.org/10.1145/792550.792553>
- [5] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz, “A bayesian approach to filtering junk e-mail,” 1998.
- [6] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, and K. Ross, “Video interactions in online video social networks,” *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 5, no. 4, pp. 30:1–30:25, Nov. 2009.
- [7] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, “Detecting spammers on twitter,” 2010.
- [8] K. Lee, J. Caverlee, and S. Webb, “The social honeypot project: protecting online communities from spammers,” in *Proceedings of the 19th international conference on World wide web*, ser. WWW '10, 2010, pp. 1139–1140.
- [9] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow, “The anatomy of the facebook social graph,” *CoRR*, vol. abs/1111.4503, 2011.
- [10] H. Kwak, C. Lee, H. Park, and S. Moon, “What is twitter, a social network or a news media?” in *Proceedings of the 19th international conference on World wide web*, ser. WWW '10. New York, NY, USA: ACM, 2010, pp. 591–600. [Online]. Available: <http://doi.acm.org/10.1145/1772690.1772751>
- [11] H. Zhang, “The optimality of naive bayes,” in *FLAIRS Conference*, V. Barr and Z. Markov, Eds. AAAI Press, 2004.
- [12] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [13] J. R. Quinlan, “Induction of decision trees,” *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, Mar. 1986.
- [14] G. Shakhnarovich, T. Darrell, and P. Indyk, *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing)*. The MIT Press, 2006.
- [15] V. Levenshtein, “Binary Codes Capable of Correcting Deletions, Insertions and Reversals,” *Soviet Physics Doklady*, vol. 10, p. 707, 1966.
- [16] W. E. Winkler, “String comparator metrics and enhanced decision rules in the fellegi-sunter model of record linkage,” in *Proceedings of the Section on Survey Research*, 1990, pp. 354–359.
- [17] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining, (First Edition)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2005.
- [18] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten, and I. Osipkov, “Spamming botnets: signatures and characteristics,” *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 171–182, Aug. 2008.
- [19] Z. Gyöngyi, H. Garcia-Molina, and J. Pedersen, “Combating web spam with trustrank,” in *Proceedings of the Thirtieth international conference on Very large data bases - Volume 30*, ser. VLDB '04, pp. 576–587.
- [20] A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly, “Detecting spam web pages through content analysis,” in *Proceedings of the 15th international conference on World Wide Web*, ser. WWW '06, pp. 83–92.
- [21] C.-T. Wu, K.-T. Cheng, Q. Zhu, and Y.-L. Wu, “Using visual features for anti-spam filtering,” in *Image Processing (ICIP 2005) IEEE International Conference on*, vol. 3, pp. 509–12.
- [22] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao, “Detecting and characterizing social spam campaigns,” in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, ser. IMC '10. New York, NY, USA: ACM, 2010, pp. 35–47.
- [23] C. Grier, K. Thomas, V. Paxson, and M. Zhang, “@spam: the underground on 140 characters or less,” in *Proceedings of the 17th ACM conference on Computer and communications security*, ser. CCS '10. New York, NY, USA: ACM, 2010, pp. 27–37.
- [24] A. H. Wang, “Don't follow me - spam detection in twitter,” in *SECURITY*, S. K. Katsikas and P. Samarati, Eds., 2010, pp. 142–151.
- [25] S. Yardi, D. M. Romero, G. Schoenebeck, and D. Boyd, “Detecting spam in a twitter network,” *First Monday*, 2010.
- [26] G. Stringhini, C. Kruegel, and G. Vigna, “Detecting spammers on social networks,” in *Proceedings of the 26th Annual Computer Security Applications Conference*, ser. ACSAC '10. New York, NY, USA: ACM, 2010, pp. 1–9.
- [27] Y. Shin, M. Gupta, and S. Myers, “Prevalence and mitigation of forum spamming,” in *INFOCOM, 2011 Proceedings IEEE*, pp. 2309–2317.
- [28] G. Mishne, “Blocking blog spam with language model disagreement,” in *In Proceedings of the First International Workshop on Adversarial Information Retrieval on the Web (AIRWeb)*, 2005.
- [29] N. Jindal and B. Liu, “Opinion spam and analysis,” in *Proceedings of the international conference on Web search and web data mining*, ser. WSDM '08. New York, NY, USA: ACM, 2008, pp. 219–230.
- [30] A. Mukherjee, B. Liu, J. Wang, N. Glance, and N. Jindal, “Detecting group review spam,” in *Proceedings of the 20th international conference companion on World wide web*, ser. WWW '11. New York, NY, USA: ACM, 2011, pp. 93–94. [Online]. Available: <http://doi.acm.org/10.1145/1963192.1963240>
- [31] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, “Detecting product review spammers using rating behaviors,” in *Proceedings of the 19th ACM international conference on Information and knowledge management*, ser. CIKM '10. New York, NY, USA: ACM, 2010, pp. 939–948. [Online]. Available: <http://doi.acm.org/10.1145/1871437.1871557>
- [32] “LinkedIn spam campaign,” <http://pandalabs.pandasecurity.com/linkedin-spam-campaign/>.
- [33] “LinkedIn spam serving adobe and java exploits,”

<http://pandalabs.pandasecurity.com/linkedin-spam-serving-adobe-and-java-exploits/>.

- [34] J. Ginsberg, M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data." *Nature*, vol. 457, no. 5, pp. 1012–4, 2009.
- [35] M. Najork, "Web spam detection." in *Encyclopedia of Database Systems*, L. Liu and M. T. Özsu, Eds. Springer US, 2009, pp. 3520–3523.
- [36] —, "System and method for identifying cloaked web servers," Patent 20030 131 048.
- [37] K. Chellapilla and A. Maykov, "A taxonomy of javascript redirection spam," in *Proceedings of the 3rd international workshop on Adversarial information retrieval on the web*, ser. AIRWeb '07. New York, NY, USA: ACM, 2007, pp. 81–88. [Online]. Available: <http://doi.acm.org/10.1145/1244408.1244423>
- [38] B. Wu and B. D. Davison, "Cloaking and redirection: A preliminary study," 2005.
- [39] S. Webb, J. Caverlee, and C. Pu, "Characterizing web spam using content and http session analysis."
- [40] —, "Predicting web spam with http session information," in *Proceedings of the 17th ACM conference on Information and knowledge management*, ser. CIKM '08, 2008, pp. 339–348.
- [41] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Commun. ACM*, vol. 18, no. 11, pp. 613–620, Nov. 1975.
- [42] A. Monge and C. Elkan, "The field matching problem: Algorithms and applications," in *In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 267–270.
- [43] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *SIGKDD Explor. Newsl.*, vol. 11, pp. 10–18, November 2009. [Online]. Available: <http://doi.acm.org/10.1145/1656274.1656278>
- [44] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," 1995, pp. 1137–1143.

Limit Cycle Generation for Multi-Modal and 2-Dimensional Piecewise Affine Control Systems

Tatsuya Kai

Department of Applied Electronics
Faculty of Industrial Science and Technology
Tokyo University of Science
6-3-1 Niijuku, Katsushika-ku, Tokyo 125-8585 JAPAN
Email: kai@rs.tus.ac.jp

Abstract—This paper considers a limit cycle control problem of a multi-modal and 2-dimensional piecewise affine control system. Limit cycle control means a controller design method to generate a limit cycle for given piecewise affine control systems. First, we deal with a limit cycle synthesis problem and derive a new solution of the problem. In addition, theoretical analysis on the rotational direction and the period of a limit cycle is shown. Next, the limit cycle control problem for piecewise affine control system is formulated. Then, we obtain matching conditions such that the piecewise affine control system with the state feedback law corresponds to the reference system which generates a desired limit cycle. Finally, in order to indicate the effectiveness of the new method, a numerical simulation is illustrated.

I. INTRODUCTION

Limit cycles are known to be quite important concept in various research fields [1]. We can find limit cycles in real world, for example, stable walking or gaits of humanoid robots in robotic engineering, oscillator circuits in electronic engineering, catalytic hypercycles in chemistry, circadian rhythms in biology, boom-bust cycles in economics and so on. Researches on limit cycles have been eagerly done from both mathematical and engineering perspectives so far [2], [3], [4], [5], [6], [7], [8], [9], [11], [12], [13], [14]. Especially, some conditions for nonlinear systems that generate periodic solutions and some applications were shown in [2], and in [7], a synthesis method of hybrid systems whose solution trajectories converge to desired trajectories was proposed. In these studies, it is guaranteed that solution trajectories of the systems converges to a desired closed curve, and the existence of limit cycles was confirmed by numerical simulations. However, the mathematical guarantee of the existence of limit cycles was not shown. On the other hand, the authors proposed a synthesis method of multi-modal and 2-dimensional piecewise affine systems that generate desired limit cycles in [10], [15] and showed a mathematical proof of the existence and the uniqueness of a limit cycle for the proposed system. In addition, some theoretical analysis on the rotational direction and the period of a limit cycle is derived. In this study, we assume that the whole of a system can be designed. A method to generate a desired limit cycle for a given piecewise affine control system with tuning some parameters of the system is more useful for a wide variety of situations. However, such a control method have not been proposed so far.

Hence, we consider a limit cycle control problem of multi-modal and 2-dimensional piecewise affine systems in this

paper. The outline of this paper is as follows. We first consider a limit cycle synthesis problem and derive its new solution in Section II. Some theoretical properties are also shown. Then, in Section III, a formulation of limit cycle control problem is presented, and necessary and sufficient conditions for the problem, which are called *matching conditions*, are derived. Finally, a numerical simulation is shown in order to confirm the effectiveness of the new method in Section IV.

II. LIMIT CYCLE SYNTHESIS OF PIECEWISE AFFINE SYSTEMS

A. Formulation of Limit Cycle Synthesis

In this section, we consider a synthesis problem of piecewise affine systems which generate desired limit cycles. First, this subsection give the formulation of the problem. Consider the 2-dimensional Euclidian space: \mathbf{R}^2 , its coordinate: $x = [x_1 \ x_2]^T \in \mathbf{R}^2$, and the origin of \mathbf{R}^2 : O . Let us set N ($N \geq 3$) points $P_i \neq O$ ($i = 1, \dots, N$) in \mathbf{R}^2 and denote the vector from O to P_i by $p_i = [p_i^1 \ p_i^2]^T$. We also denote the angle between the half line OP_i and the x_1 -axis by θ_i . Now, without loss of generality, we assume that the N points P_i ($i = 1, \dots, N$) are located in the counterclockwise rotation from the x_1 -axis, that is, $0 \leq \theta_1 < \theta_2 < \dots < \theta_N$ holds.

Next, we define the semi-infinite region D_i which is sandwiched by the half lines OP_i and OP_{i+1} and the line segment C_i joining P_i and P_{i+1} , where $P_{N+1} = P_1$. Set a polygon that is a union of C_i ($i = 1, \dots, N$) as

$$C := \bigcup_{i=1}^N C_i. \quad (1)$$

Fig. 1 shows an example of a Polygonal Closed Curve for $N = 5$.

We then consider the affine system defined in D_i :

$$\dot{x} = a_i + A_i x, \quad x \in D_i \quad (2)$$

where $x = [x_1 \ x_2]^T \in \mathbf{R}^2$ is the state variable, and $a_i \in \mathbf{R}^2$, $A_i \in \mathbf{R}^{2 \times 2}$ are the affine term and the coefficient matrix, respectively. Consequently, we treat the N -modal and 2-dimensional piecewise affine system that consists of N regions D_i ($i = 1, \dots, N$) and N affine systems (2). In this paper, we consider the following synthesis problem on limit cycles called “the limit cycle synthesis problem.”

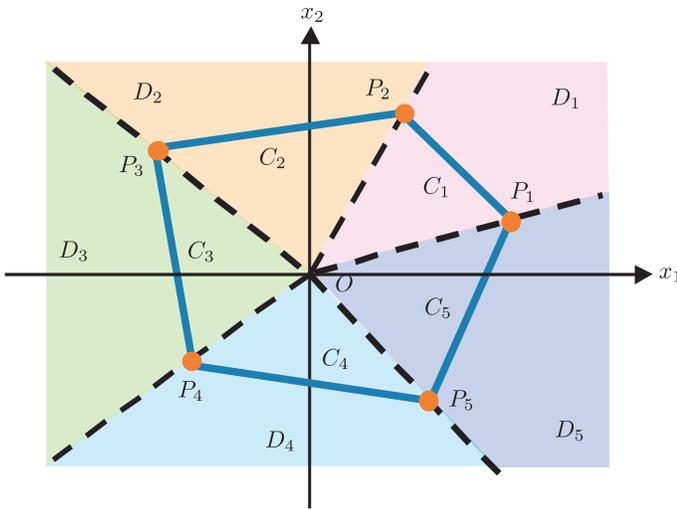


Fig. 1 : Example of Polygonal Closed Curve ($N = 5$)

Problem 1 : For the N -modal and 2-dimensional piecewise affine system (2), design a_i, A_i ($i = 1, \dots, N$) such that a given polygonal closed curve C (1) is a unique and stable limit cycle pf the system. \square

A solution of Problem 1 has been derived in the author's previous studies [10], [15]. However, in this paper we will derive another solution which can be utilized to consider the limit cycle control problem shown in Section III.

B. Proposed System and Existence/Uniqueness of Limit Cycle

Next, we shall derive a solution for Problem 1 in this subsection. We also consider the existence and the uniqueness of a limit cycle for the obtained system. We focus on a behavior of a solution trajectory of (2) in D_i . It is easily confirmed that the equation of C_i is represented by

$$(p_i^2 - p_{i+1}^2)x_1 - (p_i^1 - p_{i+1}^1)x_2 + p_i^1 p_{i+1}^2 - p_i^2 p_{i+1}^1 = 0. \quad (3)$$

Using (3), we now define a limit cycle function V_i as

$$V_i(x) = (p_i^2 - p_{i+1}^2)x_1 - (p_i^1 - p_{i+1}^1)x_2 + p_i^1 p_{i+1}^2 - p_i^2 p_{i+1}^1. \quad (4)$$

If V_i converges to 0 along a solution trajectory of (2), then the solution trajectory of (2) also converges to C_i . Hence, a_i and A_i should be determined so that V_i converges to 0 along a solution trajectory of (2). Now, an important result on design of nonlinear systems is shown as follows [2].

Theorem 1 [2] : Consider the 2-dimensional nonlinear system:

$$\dot{x} = f(x) + g(x), \quad (5)$$

where $x \in \mathbf{R}^2$ and $f, g \in \mathbf{R}^2 \rightarrow \mathbf{R}^2$ are vector fields defined in \mathbf{R}^2 . In addition, consider a radial and unbounded function define on \mathbf{R}^2 : $V : \mathbf{R}^2 \rightarrow \mathbf{R}$ such that $V(0) = 0$ and $V(x) \neq 0, \forall x \neq 0$ hold. We now define f and g as

$$f := U_f(x) \frac{\partial V}{\partial x}^\top, \quad g := -u_g(V) U_g(x) \frac{\partial V}{\partial x}^\top, \quad (6)$$

where a skew-symmetric matrix U_f , a positive definite matrix U_g , and u_g such that $u_g(V) > 0, V \neq 0$ holds. Then, for the system (5) with (6),

$$\lim_{t \rightarrow \infty} V(x(t)) = 0 \quad (7)$$

holds \square

Applying Theorem 1, we can derive an affine term a_i and a coefficient matrix A_i of the affine system (2) such that V_i converges to 0 along a solution trajectory of (2). As a specific form of (5) and (6), we use

$$\begin{aligned} \dot{x} &= f_i + g_i, \\ f_i &:= \begin{bmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{bmatrix} \frac{\partial V_i}{\partial x}^\top, \\ g_i &:= -V_i(x) \begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_i \end{bmatrix} \frac{\partial V_i}{\partial x}^\top, \end{aligned} \quad (8)$$

where ω_i and $\lambda_i > 0$ are design parameters. Substituting (4) into (8) and comparing it with (2), we can obtain a_i and A_i of (2) as

$$\begin{aligned} a_i &= \begin{bmatrix} -\lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 p_{i+1}^2 - p_i^2 p_{i+1}^1) - \omega_i(p_i^1 - p_{i+1}^1) \\ \lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 p_{i+1}^2 - p_i^2 p_{i+1}^1) - \omega_i(p_i^2 - p_{i+1}^2) \end{bmatrix}, \\ A_i &= \begin{bmatrix} -\lambda_i(p_i^2 - p_{i+1}^2)^2 & \lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 - p_{i+1}^1) \\ \lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 - p_{i+1}^1) & -\lambda_i(p_i^1 - p_{i+1}^1)^2 \end{bmatrix}. \end{aligned} \quad (9)$$

Compared to the piecewise affine system shown in [10], [15], (28) contains a new parameter λ_i and this additional parameter plays an important role in the limit cycle control problem in Section III. It is noted that that the system (2) with (28) satisfies only the convergence property (7), that is, its solution trajectory converges to C_i in D_i . Hence, we will discuss the existence of a unique and stable limit cycle of the system (2) with (28). To prove this, we first indicate three lemmas, and then we show the main theorem by using them. Now, we give the definition on the clockwise and counterclockwise rotations of limit cycle solution trajectories of the system (2) with (28) [10], [15].

Definition 1 [10], [15] : For limit cycle solution trajectories of the N -modal and 2-dimensional piecewise affine system (2) with (28), one that rotates in the clockwise direction in \mathbf{R}^2 is called a limit cycle solution trajectory in the clockwise rotation. On the contrary, one that rotates in the counterclockwise direction in \mathbf{R}^2 is called a limit cycle solution trajectory in the counterclockwise rotation (see Fig. 2). \square

Let us define a subset in D_i as

$$M_i(\varepsilon_i) := \{ x \in D_i \mid \varepsilon_i^- \leq V_i(x) \leq \varepsilon_i^+ \}, \quad (10)$$

where $\varepsilon_i^-, \varepsilon_i^+ \in \mathbf{R}$ satisfies $\varepsilon_i^- < \varepsilon_i^+$ and we set $\varepsilon_i = (\varepsilon_i^-, \varepsilon_i^+)$. We also define a sum of these subsets as

$$M(\varepsilon) := \bigcup_{i=1}^N M_i(\varepsilon_i), \quad (11)$$

where we use the notations: $\varepsilon^- = (\varepsilon_1^-, \dots, \varepsilon_N^-)$, $\varepsilon^+ = (\varepsilon_1^+, \dots, \varepsilon_N^+)$, $\varepsilon = (\varepsilon^-, \varepsilon^+)$, and the parameters ε^- and ε^+

are determined such that

$$M^-(\varepsilon^-) = \bigcup_{i=1}^N \{x \in D_i \mid V_i(x) = \varepsilon_i^-\} \quad (12)$$

$$M^+(\varepsilon^+) = \bigcup_{i=1}^N \{x \in D_i \mid V_i(x) = \varepsilon_i^+\}$$

form closed polygons. If $M^-(\varepsilon^-)$ and $M^+(\varepsilon^+)$ are closed polygons, that is, $M(\varepsilon)$ is a bounded and closed set, then ε is said to be *admissible* (see Fig. 3). We can derive the following proposition on $M(\varepsilon)$.

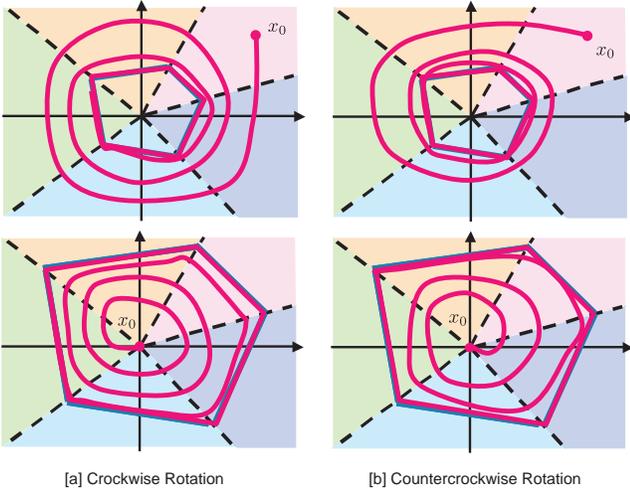


Fig. 2 : Clockwise and Counterclockwise Rotations of Limit Cycle Solution Trajectories

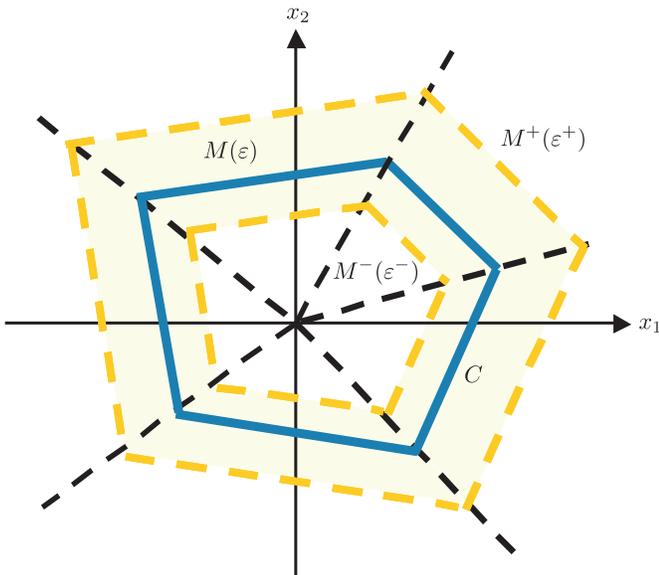


Fig. 3 : Example of $M^-(\varepsilon^-)$, $M^+(\varepsilon^+)$ and $M(\varepsilon)$

Lemma 1 : For the N -modal and 2-dimensional piecewise affine system (2) with (28), $M(\varepsilon)$ is a positively invariant, bounded and closed set for any admissible ε .

(Proof) Calculating the time differential of $1/2V_i^2$ along a

solution trajectory of the system (2) with (28), we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt}(V_i^2) &= V_i \dot{V}_i = V_i \frac{\partial V_i}{\partial x} \dot{x} = V_i \frac{\partial V_i}{\partial x} (f_i + g_i) \\ &= V_i \frac{\partial V_i}{\partial x} \begin{bmatrix} 0 & \omega_i \\ -\omega_i & 0 \end{bmatrix} \frac{\partial V_i}{\partial x} \\ &\quad - V_i^2 \frac{\partial V_i}{\partial x} \begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_i \end{bmatrix} \frac{\partial V_i}{\partial x} \\ &= -\lambda_i^2 V_i^2 \frac{\partial V_i}{\partial x} \frac{\partial V_i}{\partial x} < 0. \end{aligned}$$

Hence, it turns out that the velocity vector field of the system (2) with (28) points to the direction of the inner side of the bounded and closed set M at any points on the boundary $M^-(\varepsilon^-) \cup M^+(\varepsilon^+)$ of M . Consequently, $M(\varepsilon)$ is a positively invariant, bounded and closed set. \square

Next, we consider equilibrium points of the system (2) with (28). The following lemma on equilibrium points can be obtained.

Lemma 2 : Assume $\omega_i \neq 0$ ($i = 1, \dots, N$). Then, the N -modal and 2-dimensional piecewise affine system (2) with (28) does not have any equilibrium points in $M(\varepsilon)$ for any admissible ε .

(Proof) The unit vector which is on a parallel with C_i in D_i and points to the counterclockwise rotation is given by $(p_i - p_{i+1}) / \|p_i - p_{i+1}\|$. By considering the inner product of this unit vector and the velocity vector field of the system (2) with (28), we have the magnitude of the velocity component to the direction of $p_i - p_{i+1}$ for a solution trajectory v_i of the system (2) with (28) in D_i as

$$v_i = (a_i + A_i x) \cdot \frac{p_i - p_{i+1}}{\|p_i - p_{i+1}\|}. \quad (13)$$

Now, we denote a point in D_i by $x = \alpha_i p_i + \beta_i p_{i+1}$, $\alpha_i, \beta_i \geq 0$. Hence, we can calculate (13) as

$$\begin{aligned} v_i &= \{a_i + A_i(\alpha_i p_i + \beta_i p_{i+1})\} \cdot \frac{p_i - p_{i+1}}{\|p_i - p_{i+1}\|} \\ &= -\omega_i \sqrt{(p_i^1 - p_{i+1}^1)^2 + (p_i^2 - p_{i+1}^2)^2}. \end{aligned} \quad (14)$$

Note that λ_i does not appear in (14). From (14), we can see that the parameters α_i and β_i vanish, and hence v_i is constant at any point $x \in D_i$. Since v_i does not vanish at any point $x \in D_i$, the system (2) with (28) does not have any equilibrium points in $M(\varepsilon)$. \square

Now, a definition on the concept “traversal” for the system (2) with (28) is given as follows [11].

Definition 2 : Let Σ be a line segment in the positively invariant, bounded and closed set $M(\varepsilon)$. If the value of an inner product of the unit normal vector to Σ : e_Σ and the velocity vector of the N -modal and 2-dimensional piecewise affine system (2) with (28) is not equal to 0 and its sign does not change at any point in Σ , then Σ is said to be *traversal* with respect to the system (2) with (28). \square

In addition to Lemma 2, under the condition of $\omega_i > 0$ ($i = 1, \dots, N$), a solution trajectory vector of the system (2) with (28) always has a velocity component in the counterclockwise

rotation. On the other hand, under the condition of $\omega_i < 0 (i = 1, \dots, N)$, a solution trajectory vector of (2) with (28) always has a velocity component in the clockwise rotation. From this fact, we can derive the following lemma.

Lemma 3 : For the N -modal and 2-dimensional piecewise affine system (2) with (28), assume that $\omega_i > 0 (i = 1, \dots, N)$ or $\omega_i < 0 (i = 1, \dots, N)$ holds. Then, there exists a traversal line segment Σ at any point in $x \in M(\varepsilon)$, and it is satisfied that $x \in \Sigma$ and Σ infinitely intersects with solution trajectories of the system (2) with (28).

(Proof) We assume that $\omega_i > 0 (i = 1, \dots, N)$ or $\omega_i < 0 (i = 1, \dots, N)$ holds. Then, a solution trajectory of the system (2) with (28) always circles to the counterclockwise rotation or to the clockwise rotation. Now, for a point $x \in M$, we consider a half line whose origin is O and that passes through x , and define a subset $\Sigma \subset M$ as the intersection of the half line and M . Since the velocity vector field of the system (2) with (28) always has the velocity component of the counterclockwise rotation or the clockwise rotation, the inner product of a normal vector of Σ and the vector field of the system (2) with (28) at any point in Σ is not equal to 0 and its sign does not change, that is, Σ is traversal. Moreover, in each $M_i (i = 1, \dots, N)$, since the velocity vector field of the system (2) with (28) always has the velocity component of the clockwise rotation or the counterclockwise rotation, a solution trajectory of the system (2) with (28) $x(t)$ that intersected Σ intersects Σ in a finite time again. Consequently, it turns out the solution trajectory intersects Σ infinitely. \square

We have to note that the result in Lemma 3 does not depend on $\lambda_i (i = 1, \dots, N)$. Using Lemmas 1-3, we can derive the main theorem on the existence of the limit cycle of the system (2) with (28).

Theorem 2 : For the N -modal and 2-dimensional piecewise affine system (2) with (28), assume that $\omega_i > 0 (i = 1, \dots, N)$ or $\omega_i < 0 (i = 1, \dots, N)$ holds. Then, the unique and stable limit cycle of the system (2) with (28) is equivalent to C .

(Proof) By the result on the hybrid Poincare-Bendixson theorem derived in [3], [11], it turns out that sufficient conditions for the existence of stable limit cycles of the system (2) with (28) in $M(\varepsilon)$ are the following three: (i) $M(\varepsilon)$ is a positively invariant, bounded and closed set, (ii) there do not exist any equilibrium points at the boundary and in the interior of $M(\varepsilon)$ (iii) there exists a traversal line segment $\Sigma \subset M(\varepsilon)$ such that $x \in \Sigma$ and Σ infinitely intersects with solution trajectories of the system (2) with (28). Since we have confirmed these three conditions in Lemma 1, 2, and 3, we can see that there exists a stable limit cycle in $M(\varepsilon)$ for the system (2) with (28) for any admissible ε . Moreover, since $M(\varepsilon)$ converges to C as the values of ε goes to 0, it can be confirmed that C is a unique and stable limit cycle. Hence, the proof is completed. \square

In this paper, we consider an additional parameter λ_i in the system (2), (28). However, from the results obtained in this subsection, we can see that the existence and the uniqueness of a limit cycle of the system (2), (28) are independent of λ_i . This fact is quite important in the next section.

C. Theoretical Analysis

Finally, this subsection gives theoretical analysis on rotational directions and periods of limit cycle solution trajectories of the system (2) with (28). First, we consider the relationship between rotational directions of limit cycles and the parameters in (28). The following proposition can be derived.

Proposition 1 : For the N -modal and 2-dimensional piecewise affine system (2) with (28), its limit cycle solution trajectory moves in the counterclockwise rotation for $\omega_i > 0 (i = 1, \dots, N)$, and conversely it moves in the clockwise rotation for $\omega_i < 0 (i = 1, \dots, N)$.

(Proof) The proof of this proposition is trivial from the discussion in the previous section. \square

From Proposition 1, it is confirmed that the rotational directions of limit cycles do not depend on λ_i . Next, we analyze periods of limit cycles of the system (2) with (28). It can be expected that after a solution trajectory of the system (2) with (28) converges to C , it behaves as a periodic trajectory. By calculating the velocity component of the vector of the system along C , we can derive the next proposition.

Proposition 2 : When a limit cycle solution trajectory of the N -modal and 2-dimensional piecewise affine system (2) with (28) is sufficiently close to C , the period with which it rotates around C is given by

$$T \approx \sum_{i=1}^N \frac{1}{|\omega_i|}. \quad (15)$$

(Proof) The velocity component of a solution trajectory v_i of (2) with (28) in D_i to the direction of $p_i - p_{i+1}$ is given by (13). The length of C_i : L_i can be calculated as

$$L_i = \sqrt{(p_i^1 - p_{i+1}^1)^2 + (p_i^2 - p_{i+1}^2)^2}. \quad (16)$$

Therefore, we can obtain the period T as

$$T \approx \sum_{i=1}^N \frac{L_i}{|v_i|} = \sum_{i=1}^N \frac{1}{|\omega_i|}. \quad (17)$$

This completes the proof of this proposition. \square

From Proposition 2, we can also see that the period of a limit cycle solution trajectory of the system (2), (28) is not independent of λ_i . So, we can freely choose the value of λ_i .

III. LIMIT CYCLE CONTROL FOR PIECEWISE AFFINE SYSTEMS

A. Formulation of Limit Cycle Control

In this section, we consider a controller design problem on generation of limit cycles for given piecewise affine control systems. First, this sub-section gives the problem formulation. Consider the next piecewise affine control system defined in D_i :

$$\dot{x} = a_i + A_i x + b_i u, \quad x \in D_i, \quad (18)$$

where $u \in \mathbf{R}$ is the control input and $b_i \in \mathbf{R}^2$ is the coefficient vector for the control input. We next consider the state feedback law:

$$u = k_i x + l_i, \quad x \in D_i, \quad (19)$$

where $k_i \in \mathbf{R}^2$ and $l_i \in \mathbf{R}$. We assume that p_i, p_{i+1}, a_i, A_i are given parameters. Now, we formulate a problem on generating a desired limit cycle for the piecewise affine control system (2) and the state feedback law (19) as follows.

Problem 2 : For the N -modal and 2-dimensional piecewise affine control system (18) with the state feedback law (19), design $b_i, k_i, l_i, \omega_i, \lambda_i$ ($i = 1, \dots, N$) such that a given polygonal closed curve C (1) is a unique and stable limit cycle of the closed-loop system. \square

Throughout this paper, we call Problem 2 a *limit cycle control problem for piecewise affine control system*.

B. Matching Conditions for Limit Cycle Control Problem

The purpose of this subsection is to derive a solution method of Problem 2 for the piecewise affine control system (18) with the state feedback law (19). To fulfill this, we shall utilize the limit cycle synthesis method obtained in Section 2. The results in Section II show that the unique and stable limit cycle of the system (2), (28) coincides with C . Hence, by tuning design parameters $b_i, k_i, l_i, \omega_i, \lambda_i$ ($i = 1, \dots, N$), we conform the closed-loop system (18), (19) to the system (2), (28). We here call the system (2), (28) *the reference system*. Use the following notations for the system (2), (28):

$$\begin{aligned} a_i &= \begin{bmatrix} a_i^1 \\ a_i^2 \end{bmatrix}, \quad A_i = \begin{bmatrix} A_i^{11} & A_i^{12} \\ A_i^{21} & A_i^{22} \end{bmatrix}, \\ b_i &= \begin{bmatrix} b_i^1 \\ b_i^2 \end{bmatrix}, \quad k_i = \begin{bmatrix} k_i^1 & k_i^2 \end{bmatrix}. \end{aligned} \quad (20)$$

Conditions such that the closed-loop system (18), (19) is consistent with the reference system (2), (28) can be obtained by the following theorem.

Theorem 3 : The N -modal and 2-dimensional piecewise affine control system (18) with the state feedback control law (19) is equivalent to the reference system (2), (28) if and only if *the matching conditions*:

$$a_i^1 + b_i^1 l_i = -\lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 p_{i+1}^2 - p_i^2 p_{i+1}^1) - \omega_i(p_i^1 - p_{i+1}^1) \quad (21)$$

$$a_i^2 + b_i^2 l_i = \lambda_i(p_i^1 - p_{i+1}^1)(p_i^1 p_{i+1}^2 - p_i^2 p_{i+1}^1) - \omega_i(p_i^2 - p_{i+1}^2) \quad (22)$$

$$A_i^{11} + b_i^1 k_i^1 = -\lambda_i(p_i^2 - p_{i+1}^2)^2 \quad (23)$$

$$A_i^{12} + b_i^1 k_i^2 = \lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 - p_{i+1}^1) \quad (24)$$

$$A_i^{21} + b_i^2 k_i^1 = \lambda_i(p_i^2 - p_{i+1}^2)(p_i^1 - p_{i+1}^1) \quad (25)$$

$$A_i^{22} + b_i^2 k_i^2 = -\lambda_i(p_i^1 - p_{i+1}^1)^2 \quad (26)$$

hold.

(Proof) Substituting (19) into (18), we get the closed-loop

system:

$$\begin{aligned} \dot{x} &= a_i + A_i x + b_i(k_i x + l_i) \\ &= a_i + b_i l_i + (A_i + b_i k_i) x \\ &= \begin{bmatrix} a_i^1 + b_i^1 l_i \\ a_i^2 + b_i^2 l_i \end{bmatrix} + \begin{bmatrix} A_i^{11} + b_i^1 k_i^1 & A_i^{12} + b_i^1 k_i^2 \\ A_i^{21} + b_i^2 k_i^1 & A_i^{22} + b_i^2 k_i^2 \end{bmatrix} x. \end{aligned} \quad (27)$$

Comparing the components of the reference system (2), (28) to the closed-loop system (27), we can obtain the matching conditions (21)–(26). \square

The matching conditions (21)–(26) consists of 6 algebraic equations, and 7 unknown variables: $b_i^1, b_i^2, k_i^1, k_i^2, l_i, \omega_i, \lambda_i$. Hence, by solving them under the condition $\lambda_i > 0$, we can obtain these unknown variables, that is, a solution of Problem 2.

IV. SIMULATIONS

This section presents a numerical example in order to confirm the effectiveness of the results derived in the previous sections. We now give data of a polygon with $N = 4$ as $P_1 = (1, 0)$, $P_2 = (0, 1)$, $P_3 = (-1, 0)$, $P_4 = (0, -1)$. The polygon is shown in Fig. 4.

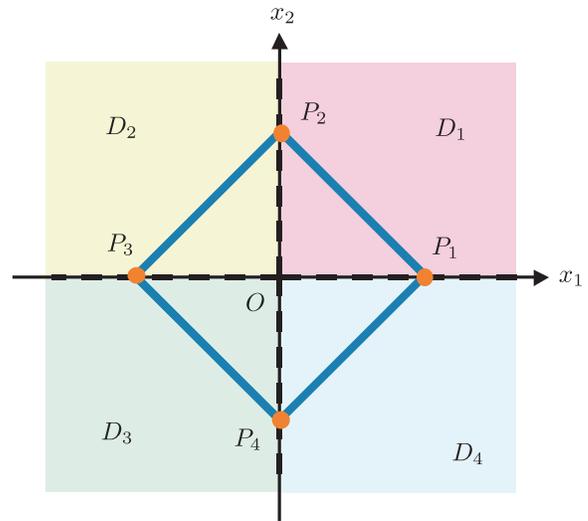


Fig. 4 : Polygonal Closed Curve of Example

The coefficients of the piecewise affine system are given by

$$\begin{aligned} a_1 &= \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix}, \\ a_2 &= \begin{bmatrix} -1 \\ -3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \\ a_3 &= \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}, \\ a_4 &= \begin{bmatrix} 9 \\ 1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} -8 & 0 \\ 0 & -8 \end{bmatrix}. \end{aligned} \quad (28)$$

By solving the matching conditions for Problem 2 (21)–(26) in terms of the problem formulation, we can calculate

design parameters as follows:

$$\begin{aligned}
 b_1 &= \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad k_1 = \begin{bmatrix} -1 & -2 \end{bmatrix}, \\
 l_1 &= 1, \quad \omega_1 = 3, \quad \lambda_1 = 1, \\
 b_2 &= \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad k_2 = \begin{bmatrix} 1 & -2 \end{bmatrix}, \\
 l_2 &= 2, \quad \omega_2 = 1, \quad \lambda_2 = 2, \\
 b_3 &= \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad k_3 = \begin{bmatrix} 1 & -2 \end{bmatrix}, \\
 l_3 &= 2, \quad \omega_3 = 2, \quad \lambda_3 = 3, \\
 b_4 &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad k_4 = \begin{bmatrix} 4 & 4 \end{bmatrix}, \\
 l_4 &= -4, \quad \omega_4 = 1, \quad \lambda_4 = 4.
 \end{aligned} \tag{29}$$

Note that $\lambda_i > 0$ ($i = 1, 2, 3, 4$) holds in (29). It can be confirmed that from Proposition 1, a limit cycle solution trajectory moves in the counterclockwise rotation since $\omega_i > 0$ ($i = 1, 2, 3, 4$) holds. In addition, from Proposition 2 we can estimate the period of a limit cycle solution trajectory as

$$T \approx \sum_{i=1}^4 \frac{1}{|\omega_i|} = \frac{17}{6}. \tag{30}$$

We set the initial state as $x_0 = [1, 1]^T$ for the numerical simulation. The simulation results are illustrated in Figs. 5–7. Fig. 5 shows the solution trajectory on the x_1x_2 -plane. In Figs. 6 and 7, the time series of x_1 and x_2 are shown, respectively. From these simulation results, we can see that the solution trajectory that starts from x_0 behaves as a limit cycle for the desired polygonal closed curve C , and hence Theorem 1 holds. As we expected above, the solution trajectory moves in the counterclockwise rotation, and this result is coincident with Proposition 1. Moreover, the estimated period $T \approx 17/6$ is mostly agree about the simulation result from Figs. 6 and 7.

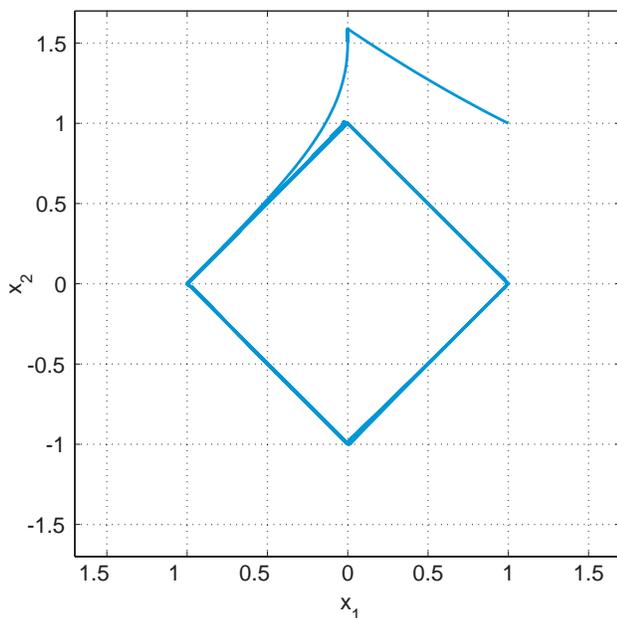


Fig. 5 : Solution Trajectory on x_1x_2 -Plane

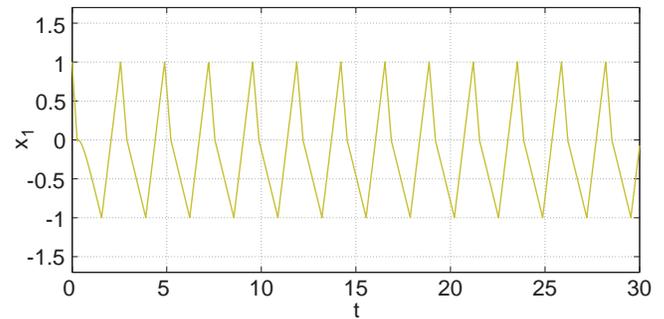


Fig. 6 : Time Series of x_1

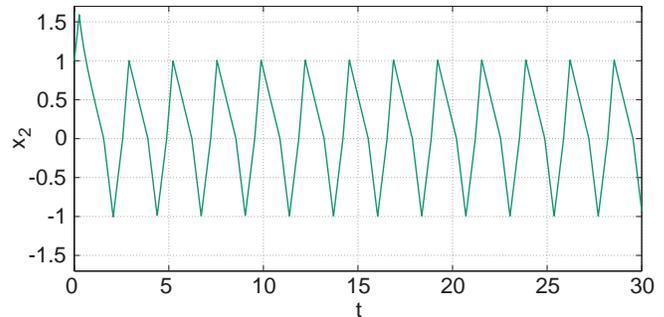


Fig. 7 : Time Series of x_2

V. CONCLUSION

In this paper, we have considered a limit cycle control problem for a multi-modal and 2-dimensional piecewise control affine system. We have derive the matching conditions such that the piecewise control affine system with the state feedback law corresponds with the reference system which generates a unique and stable limit cycle. It has been confirmed by solving the matching conditions we can obtain the values of design parameters. A numerical simulations show the availability and the application potentiality of the proposed method.

Our future work includes applications of the proposed control method to real systems and extensions to multi-dimensional piecewise affine systems.

REFERENCES

- [1] J. Buchli, L. Righetti and A. Ijspeert, "Engineering Entrainment and Adaptation in Limit Cycle Systems," *Biological Cybernetics*, vol.95, pp.645–664, 2006
- [2] D. N. Green, "Synthesis of Systems with Periodic Solutions Satisfying $\mathcal{V}(x) = 0$," *IEEE Trans. Circuits and Systems*, vol.31, no.4, pp.317–326, 1984
- [3] S. N. Simic, K. H. Johansson, J. Lygeros and S. Sastry, "Hybrid Limit Cycles and Hybrid Poincare-Bendixson," in *Proc. of IFAC World Congress*, Barcelona, Spain, pp.86–89, 2002
- [4] A. Girard, "Computation and Stability Analysis of Limit Cycles in Piecewise Linear Hybrid Systems," in *Proc. of 1st IFAC Conference on Analysis and Design of Hybrid Systems*, Saint-Malo, France, pp. 181-186, 2002
- [5] M. Adachi and T. Ushio, "Synthesis of Hybrid Systems with Limit Cycles Satisfying Piecewise Smooth Constant Equations," *IEICE Trans. Fundamentals*, vol.E87-A, No.4, pp.837–842, 2004
- [6] F. Gómez-Estern, J. Aracil, F. Gordillo and A. Barreiro, "Generation of Autonomous Oscillations via Output Feedback," in *Proc. of IEEE CDC 2005*, Seville, Spain, pp.7708–7713, 2005

- [7] A. Ohno, T. Ushio and M. Adachi, "Synthesis of Nonautonomous Systems with Specified Limit Cycles," *IEICE Trans. Fundamentals*, vol.E89-A, No.10, pp.2833–2836, 2006
- [8] F. Gómez-Estern, A. Barreiro, J. Aracil and F. Gordillo, "Robust Generation of Almost-periodic Oscillations in a Class of Nonlinear systems," *Int. J. Robust Nonlinear Control*, vol.16, no.18, pp.863–890, 2006
- [9] D. Flieller, P. Riedinger, and J. Louis, "Computation and Stability of Limit Cycles in Hybrid Systems," *Nonlinear Analysis*, Vol. 64, pp. 352–367, 2006
- [10] T. Kai and R. Masuda, "Limit Cycle Synthesis of Multi-Modal and 2-Dimensional Piecewise Affine Systems," *Mathematical and Computer Modelling*, Vol. 55, pp. 505–516, 2012
- [11] M. Suenaga and T. Hayakawa, "Existence Condition of Periodic Orbits for Piecewise Affine Planar Systems," in *Proc. of SICE 8th Annual Conference on Control Systems*, Kyoto, Japan, 2008
- [12] T. Kai and R. Masuda, "Controller Design for 2-Dimensional Nonlinear Control Systems Generating Limit Cycles and Its Application to Spacerobots," in *Proc. of NOLTA 2008*, Budapest, Hungary, pp.496–499, 2008
- [13] T. Kai and M. Katsuta, "Limit Cycle Control for 2-Dimensional Discrete-time Nonlinear Control Systems and Its Application to Chaos Systems," in *Proc. of NOLTA 2009*, Sapporo, Japan, pp.86–89, 2009
- [14] A. Schild, Xu Chu Ding, M. Egerstedt, J. Lunze, "Design of optimal switching surfaces for switched autonomous systems," in *Proc. of IEEE Conference on Decision and Control & Chinese Control Conference 2009*, Shanghai, China, pp. 5293–5298, 2009
- [15] T. Kai and R. Masuda, "A Limit Cycle Synthesis Method of Multi-Modal and 2-Dimensional Piecewise Affine Systems," in *Proc. of 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, USA, pp. 4759-4764, 2011

Privacy-Preserving Clustering Using Representatives over Arbitrarily Partitioned Data*

Yu Li, Sheng Zhong

Computer Science and Engineering Department

State University of New York at Buffalo

Amherst, NY 14260, U.S.A

Email: {yli32,szhong}@buffalo.edu

Abstract—The challenge in privacy-preserving data mining is avoiding the invasion of personal data privacy. Secure computation provides a solution to this problem. With the development of this technique, fully homomorphic encryption has been realized after decades of research; this encryption enables the computing and obtaining results via encrypted data without accessing any plaintext or private key information. In this paper, we propose a privacy-preserving clustering using representatives (CURE) algorithm over arbitrarily partitioned data using fully homomorphic encryption. Our privacy-preserving CURE algorithm allows cooperative computation without revealing users' individual data. The method used in our algorithm enables the data to be arbitrarily distributed among different parties and to receive accurate clustering result simultaneously.

I. INTRODUCTION

With the advent of the Big Data Era, people are beginning to care more about the security of their private data. For example, when people store their data on the cloud server, there are always concerns that the service provider will use their data illegally. Cryptographers provide many elegant encryption schemes (e.g., RSA) to ensure the security of data transmission, but they are not adequate for protecting the privacy of customers' sensitive information. After the invention of RSA, Rivest, Adleman, and Dertouzos introduced privacy homomorphism in [1]. For example, RSA is a multiplicatively homomorphic encryption scheme that can efficiently compute a ciphertext that encrypts the product of the original plaintext. Based on this privacy homomorphism method, people are realizing that they can store encrypted data instead of plaintext on the cloud. Therefore, the cloud service provider can use these data to conduct research without compromising the customers' privacy.

Clustering is the most commonly used data mining method to determine the relationship between objects. With the development of network technology, collecting large amounts of data about clients and customers has become much easier. In addition, many clustering algorithms have been proposed to solve the efficiency problem of clustering for large databases. Clustering using representatives (CURE) is such an algorithm that can efficiently cluster a large database in such a way that objects in the same group are more similar to each other than to those in other groups.

However, determining how to protect the privacy of customers is a big challenge in this age of information explosion. For example, in the field of customer behavior analysis and

targeted marketing, people are more willing to communicate with others who share the same interests, they are cautious about revealing their private information to the public. Privacy-preserving data mining is the only way to solve this problem and to not invade the privacy of customers. Another example is in the bioinformatics field when two medical researchers want to study a certain disease caused by a gene. They cannot share these data with each other due to the privacy rules established by HIPAA. But privacy-preserving data mining using homomorphic encryption can determine the relationship between these data without revealing any information to the other concerned parties.

Privacy-preserving data mining is usually used for horizontally or vertically partitioned database. In [2], Jagannathan introduced a stronger assumption called the "arbitrarily partitioned" database, meaning that any of the different attributes for different features can be owned by any party. In this paper, we propose a privacy-preserving solution to a simple clustering algorithm ("CURE") over arbitrarily partitioned data. To be concrete, recall the example of collaborative customer behavior analysis and targeted marketing that we mentioned above. We assume that the attributes involved in this research are interests in food, job, religion, and entertainment, but all these data are arbitrarily partitioned by two network companies (like Facebook and LinkedIn). The two companies want to jointly cluster the people into several different groups so that people are can more conveniently to communicate with others who share their interests. In addition, what happens if the two companies are not willing to share the real data with each other. To solve this problem, we propose a privacy-preserving method for clustering using representatives. We use a novel cryptographic techniques called fully homomorphic encryption, that can enable strong privacy protection for distributed clustering. In this paper, we present a detailed analysis of the accuracy and privacy protection of our algorithm.

Our contributions can be summarized as follows:

- We are the first to study privacy protection in clustering using representatives and to propose a privacy-preserving algorithm.
- To the best of our knowledge, we are also the first to study a privacy protection in clustering algorithm using fully homomorphic encryption.
- In terms of privacy, our algorithm leaks no knowledge about each party's data except the clustering result obtained.

*This paper was supported by NSF CNS-0845149 and CCF-0915374

The rest of the paper is organized as follows. In Section 2, we describe the related work in privacy-preserving data mining. In Section 3, we describe the technical preliminaries including the CURE algorithm and the cryptographic tool that we use in this paper. Section 4 describes our privacy-preserving protocol for the CURE algorithm when the data are arbitrarily partitioned between two parties. The analysis of the computation complexity and privacy of our protocol is given in Section 5.

II. RELATED WORK

Privacy preserving data mining using cryptographic protocols has developed for more than 20 years since Lindell et al. provided a privacy-preserving algorithm for ID3 in [3]. Privacy-preserving data mining based on cryptographic protocols can obtain more accurate results than methods based on a randomization algorithm. Since the cryptographic protocol is based on computation complexity, the security of the method is also better than randomization-based algorithms.

The cryptographic protocols that are used in privacy-preserving data mining are mainly based on secure computation. The most famous one is provided by Yao[4], [5] who first introduced secure two-party computation. After that, a secure multi-party computation was proposed, the goal of which is to enable several parties to jointly compute a function without revealing their inputs to each other. Another important work[1] is provided by Rivest, Adleman, and Dertouzos following the invention of RSA encryption[6]. They introduced a concept called “privacy homomorphism”. Since this development, many homomorphic encryption schemes have been proposed. RSA encryption[6] is a multiplicatively homomorphic encryption scheme. ElGamal also provided an elegant multiplicatively homomorphic encryption scheme in [7]. Benaloh provided an additively homomorphic encryption scheme in [8] and Paillier proposed another one [9] in 1999. However, no fully homomorphic encryption was realized until Gentry proposed one [10] using an ideal lattice in 2009. Gentry’s scheme allows one to compute arbitrary functions over ciphertext without the decryption key[11]. After Gentry proposed his fully homomorphic encryption scheme, van Dijk et al. proposed a simpler one using integers instead of the ideal lattice.

Based on these cryptographic protocols, many privacy-preserving algorithms have been proposed in the past decade. Lindell et al. provided a privacy-preserving algorithm[3] for ID3 using Yao’s protocol[5]. Yang et al. proposed a privacy-preserving classification method for customer data using a variant of ElGamal’s encryption [12]. Chen et al. proposed a privacy-preserving backpropagation neural network learning in [13]. All of these privacy-preserving algorithms are based on either additively homomorphic encryption or multiplicatively homomorphic encryption. Since Gentry developed the first fully homomorphic encryption scheme, many researchers have tried to provide privacy-preserving applications based on the FHE. For example, FHE enables a person to submit queries to a search engine (e.g., the user submits an encrypted query, and the search engine computes a succinct encrypted answer without ever looking at the query in plaintext). More broadly, fully homomorphic encryption improves the efficiency of secure multi-party computation. Recently, Chu et al. proposed

a privacy-preserving simrank algorithm using Gentry’s FHE scheme.

To be more concrete, many privacy-preserving clustering algorithms have been proposed in the past decades. Vaidya et al. proposed a privacy-preserving K-means clustering over vertically partitioned data in [14]. However, none of these researchers have been able to develop a privacy-preserving clustering algorithm using fully homomorphic encryption; therefore, we are the first to offer this type of encryption.

III. TECHNICAL PRELIMINARY

A. Arbitrarily Partitioned Data

In two-party distributed privacy-preserving data mining, Party A and Party B have different attributes for different features. More specifically, suppose that training dataset D consists n samples ($D = \{d_1, d_2, d_3, \dots, d_n\}$), and that each sample d_i contains m attributes, which denotes $d_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}$. The arbitrarily partitioned data based on the different attributes of the different features randomly distributed by two parties.

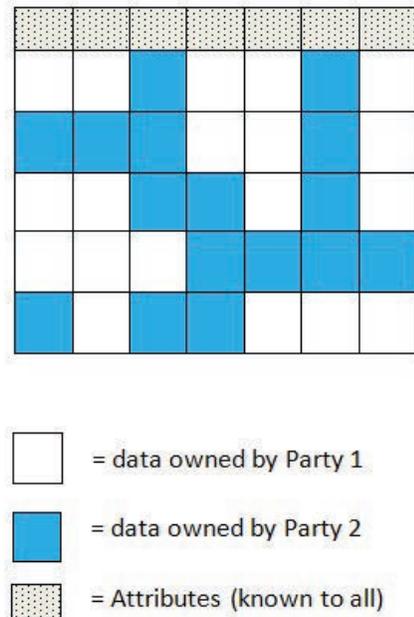


Fig. 1: Data arbitrarily partitioned by two parties

B. Fully Homomorphic Encryption Scheme

Homomorphic encryption is a form of encryption that allows specific types of computations to be carried out on ciphertext and for an encrypted result to be obtained, the decryption of which matches the result of the operations performed in plaintext. For instance, one person could add two encrypted numbers and then another person could decrypt the result without either of them being able to determine the value of the individual numbers.

To generalize this property, consider the constraint $\delta(F(\epsilon(m_1), \epsilon(m_2), \dots, \epsilon(m_n))) = F(m_1, m_2, \dots, m_n)$ for some set of functions F on all m_i in the plaintext space. We say

that an encryption scheme is partially homomorphic if F is a (finite or infinite) proper subset of computable functions. We can also say that an encryption scheme is fully homomorphic if F is the set of all computable functions.

In [10], Gentry provided the first fully homomorphic encryption scheme using ideal lattices. In [15], van Dijk et al. and Gentry et al. constructed a simple fully homomorphic encryption scheme based on Gentry's scheme[10]. van Dijk et al.'s scheme[15] only uses addition and multiplication over integers instead of ideal lattices, which is much simpler conceptually. In van Dijk et al.'s scheme, they proposed two phases to construct a fully homomorphic encryption. The first phase is constructing a somewhat homomorphic encryption. The second phase is squashing the decryption circuit to make it fully homomorphic. The following is a brief introduce of these two phases.

1) *Somewhat Homomorphic Encryption:* In this phase, there are four steps to construct a somewhat homomorphic encryption. They are key generation, encryption, evaluate and decryption[15].

- **Key Generation.**
In the key generation step, an odd η -bit integer p is randomly chosen from $(2\mathbb{Z} + 1) \cap [2^{\eta-1}, 2^\eta)$ as the private key. The corresponding public key is $pk = \langle x_0, x_1, \dots, x_\tau \rangle$, which $x_i \leftarrow D_{\gamma, \rho}(p)$ and $D_{\gamma, \rho}(p) = x = pq + r$ and $q \leftarrow \mathbb{Z} \cap [0, 2^\gamma/p), r \leftarrow \mathbb{Z} \cap (-2^\rho, 2^\rho)$
- **Encryption.**
Then choose a random subset $s \in \{1, 2, \dots, \tau\}$ and a random integer r in $(-2^{\rho'}, 2^{\rho'})$, and output $c \leftarrow [m + 2r + 2 \sum_{i \in S} (x_i)]_{x_0}$
- **Evaluate**
Given the binary circuit C_ε with t inputs, and t ciphertexts c_i , apply the integer addition and multiplication gates of C_ε to the ciphertexts, performing all the operations over the integers, and return the resulting integer
- **Decryption.**
Output $m' \leftarrow (c \bmod p) \bmod 2$.

2) *Squashing the Decryption Circuit:* In van Dijk et al.'s fully homomorphic scheme, they followed Gentry's approach[10] to make their scheme bootstrappable in this phase.

- **Key Generation.**
An odd η -bit integer p is randomly chosen from $(2\mathbb{Z} + 1) \cap [2^{\eta-1}, 2^\eta)$ as the private key. The corresponding public key is $pk = \langle x_0, x_1, \dots, x_\tau \rangle$, which $x_i \leftarrow D_{\gamma, \rho}(p)$ and $D_{\gamma, \rho}(p) = x = pq + r$ and $q \leftarrow \mathbb{Z} \cap [0, 2^\gamma/p), r \leftarrow \mathbb{Z} \cap (-2^\rho, 2^\rho)$. Set $x_p \leftarrow$
After generate the private key and public key, choose at random a $\theta - bit$ vector with Hamming weight θ , $s = \langle s_1, \dots, s_\theta \rangle$, and let $S = \{i : s_i = 1\}$.
Choose at random integers $u_i \in \mathbb{Z} \cap [0, 2^{k+1}), i = 1, \dots, \theta$ subject to the condition that $\sum_{i \in S} u_i = x_p \pmod{2^{k+1}}$. Set $y_i = u_i/2^k$ and $y = \langle y_1, \dots, y_\theta \rangle$. Hence each y_i is a positive number smaller than 2, with k bits of precision after the binary point. Also, $[\sum_{i \in S} y_i]_2 = (1/p) - \Delta_p$ for some $|\Delta_p| < 2^{-k}$.

TABLE I: Denotations of van Dijk et al's scheme[15]

γ is the bit-length of the integers in the public key
η is the bit-length of the secret key
ρ is the bit-length of the noise
τ is the number of integers in the public key

Output the secret key $sk = s$ and public key $pk = (pk^*, y)$.

- **Encryption and Evaluate.**
Generate a ciphertext c^* as before (i.e., an integer). Then for $i \in 1, \dots, \theta$, set $z_i \leftarrow [c^* \cdot y_i]_2$, keeping only $n = \lceil \log \theta \rceil + 3$ bits of precision after the binary point for each z_i . Output both c^* and $z = \langle z_1, \dots, z_\theta \rangle$.
- **Decryption.**
Output $m' \leftarrow [c^* - [\sum_i s_i z_i]]_2$

C. Notations for CURE algorithm

CURE is an efficient data clustering algorithm for large databases[16]. It is more robust for outliers and can identify clusters with non-spherical shapes and wide variances in size.

The CURE algorithm uses distance to denote the relationship of each pair of points. Then it merges the closest pair into the same cluster. It also uses a constant number of well scattered points to represent one cluster so that it can speed up the clustering process for large database. Here we briefly introduce how the CURE algorithm works. First, it calculates the distance of each pair of points and finds the nearest neighbor of each point. It then puts the nearest neighbor of each point along with that point into a queue in increasing order. When the size of the queue is bigger than k , it merges the top two clusters and updates the distance to the new cluster. In the merge process, it uses a constant number of well scattered points to represent the cluster that is shrunk toward the center of the cluster by a fraction α . Then it iteratively process above progress until the size of the queue equals to k . Then, the algorithm outputs these k clusters. [16] gives more details regarding the algorithm. Figure 2 gives an example of the CURE clustering of seven different clusters.

IV. MODEL DESCRIPTION

A. Definitions and Problem Statement

In this paper, we present a privacy-preserving CURE algorithm over arbitrarily partitioned data in different parties. The goal of our method is to protect the data privacy and to achieve the computation of the clustering simultaneously. For simplicity, we will mainly focus on a two-party scenario.

In a two-party scenario, data are arbitrarily partitioned distributed between the two parties. That means for each instance d_i , Party A holds data x_{ij} and Party B holds data $x_{ij'}$. Without loss of generality, for each $d_i \in \mathbb{D}$, $d_i = x_{ij} \cup x_{ij'}$. Party A and Party B want to determine clustering using the CURE algorithm cooperatively without revealing data to each other. For this problem, we propose a privacy-preserving CURE algorithm using fully homomorphic encryption.

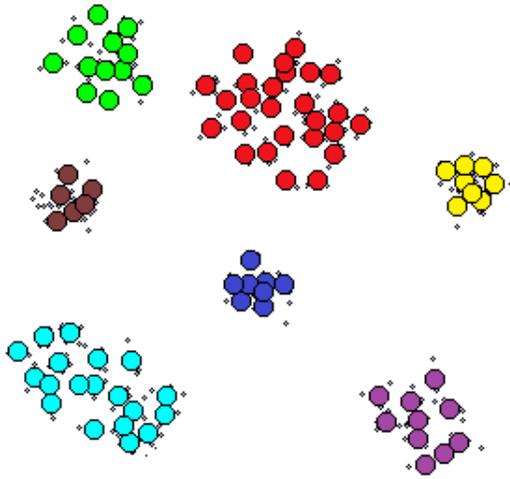


Fig. 2: CURE algorithm with 7 clusters

B. Privacy-preserving CURE algorithm using fully homomorphic encryption

We now describe our privacy-preserving CURE algorithm using fully homomorphic encryption. Party A and Party B share a data set in an arbitrary partition, which we explained in section 3.1. Since the CURE algorithm uses random sampling to handle large data sets, in this step, we can simply use Reservoir sampling to obtain a smaller data set. It is easy to see that this step does not influence the user's privacy. In addition, the CURE algorithm also uses partitioning for increasing the speed. The basic idea is to partition the sample space into p partitions. Each partition contains n/p elements. In the first pass, we partially cluster each partition until the final number of clusters is reduced to n/pq for the constant $q \geq 1$. Then, we run a second clustering pass on n/q partial clusters for all the partitions. For the second pass, we only store the representative points since the merge procedure only requires representative points of previous clusters before computing the new representative points for the merged cluster. Obviously, this step only reduces the execution times. Therefore, the main challenge of the privacy-preserving CURE algorithm is how to protect privacy when Party A and Party B jointly execute the CURE algorithm.

In the original CURE algorithm, a k-d tree is used to store the representative points, and it can sort the multi-dimensional point data efficiently so that it is easy to retrieve. But since we need to preserve the privacy of the data, and this k-d tree which contains a sorting process will influence the efficiency enormously, we simply use a k-d matrix to store these points in stead of a k-d tree. In our two-party model, the k-d matrix is arbitrarily partitioned by Party A and Party B. We propose how to build a heap Q securely using van Dijk et al's scheme.

First, Party A generates a secret key $sk_A = p_A$ and Party B generates his secret key $sk_B = p_B$ based on van Dijk et al.'s scheme. Since the whole computation process will be based on bit manipulation. Party A and Party B need to convert their data to a binary format. Party A generates his public key pk_A and encrypts his data and sends them to Party B. Then Party

B encrypts his data using Party A's public key. For clarity, we assume that parties are both using Euclidean distance as their $Dist$. Therefore, the main challenge of the privacy-preserving CURE cluster algorithm is obtaining a secure comparison of the distance and determining how to find the closet one. Since we use Euclidean distance, we have

$$dist(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots (q_n - p_n)^2} \quad (1)$$

Since we only need to find the two points that have smallest distance, we can use the square of distance for easily computation. For each pair of points, since the attributes of the two points are arbitrarily partitioned by the two parties, the two data sets belong to the same attribute that is either owned by one party or distributed by the two parties. Without generality, we assume that Party A owns a_i^p and Party B owns b_j^p . We can simple get $(a_i^p - b_j^p)^2 = (a_i^p)^2 + (b_j^p)^2 + 2a_i^p \times b_j^p$. $(a_i^p)^2$ and $(b_j^p)^2$ can be calculated by each party itself. For $2a_i^p \times b_j^p$, we can use the full homomorphism of van Dijk et al. scheme to obtain the random share of the distance as shown in algorithm 1 and algorithm 2. For example, in algorithm 2, the random share for Party A is $R_A + dist\{i, j\} + R_B$, and for Party B is R_B . Another random share for Party A is R_A , and for Party B is $R_A + dist\{i', j'\} + R_B$. After obtaining the random share of two distances, we can use secure compare which is proposed in [17](algorithm 3) to compare them. Secure compare[17] uses two binary numbers as the input. First, Party A encrypts each bit and sends them to Party B in order. Party B then calculates $c_i = E_A(x_B^i - x_A^i + 1 + \sum_{j=1}^{i-1} w_j)$ for each i . If $x_A > x_B$, then there must exist one bit that the bits before i are the same and x_A^i is 1 and x_B^i is 0, therefore making $D_A(c_i) = 0$. Since $(x_B^i - x_A^i + 1 + \sum_{j=1}^{i-1} w_j) = x_B^i - x_A^i + 1 + \sum_{j=1}^{i-1} (x_A^j + x_B^j - 2 \cdot x_A^j \cdot x_B^j)$, this can be calculated by Party B using the full homomorphism of van Dijk et al.'s scheme. After getting the $index(i)$ which indicates the instance that has the smallest distance from instance i , we need to compare them again to get smallest one $v := index(u)$ using the algorithm 3 again. Next we can merge u and v . First, we iteratively select c points. We can use variatal algorithm 2 to calculate the distance between p and $\frac{|u|u.mean + |v|v.mean}{|u| + |v|}$. It is easy to get this variatal

algorithm, as we can assume that $\frac{|u|u.mean + |v|v.mean}{|u| + |v|}$ is a point. When calculating it, we can split it into several parts that contain u and v . Then we can use method in algorithm 2 to get the value $R_A + dist(p, \frac{|u|u.mean + |v|v.mean}{|u| + |v|} + R_B)$.

Then Party A and Party B jointly apply algorithm 3 to secure get the $temSet$. After that, we use variant algorithm 2 again to choose the point that is furthest from the previously selected points. That means we let the points shrink toward the mean by a fraction α using a variatal algorithm 2. Then we can insert this new merged cluster into the matrix dataset. After that, we can use the same method (algorithm 2) to update the array $index(i)$. Then we iteratively process above procedure until the size of array Z is equal k . Lastly, we can output these K clusters.

Algorithm 1 Privacy-preserving CURE cluster algorithm

Input Party A and party B with their own data S
Output k clusters
begin
 Party A generate his private key p_A and public key pk_A
 Party B generate his private key p_B and public key pk_B
 initialize an array $Z := index(i)$.
for each of instances i {
 for each of instances $j \neq i$ **do**{
 By applying algorithm 2, party A and party B can jointly compare the distance of i and j and store the smaller one into $index(i)$
 }
end for
 $index(i)$ indicates the instance that has the smallest distance from instance i

While $size(Z) > k$ **do**
 {
 Party A and party B jointly apply algorithm 2 to get $u := extractmin(Z)$
 $v := index(u) // v := u.closest//$
 delete(Z, u)
 $w := u \cup v$
 $tmpSet := \phi$
 for $i := 1$ **to** c **do**{
 $maxDist := 0$
 for each point p in cluster w **do**{
 if $i = 1$
 Party A and B jointly apply algorithm 2 to get
 $minDist := dist(p, \frac{|u|u.mean + |v|v.mean}{|u| + |v|})$
 else
 Party A and B jointly apply algorithm 2 to get
 $minDist := mindist(p, q) : q \in tmpSet$
 Party A and B jointly apply algorithm 3 to secure compare
 if ($minDist \geq maxDist$){
 $maxDist := minDist$
 $maxPoint := p$
 }
 }
 }
 $tmpSet := tmpSet \cup maxPoint$
 }
 for each point p in $tmpSet$ **do**
 Party A and B jointly apply variant algorithm 2 to get
 $w := w \cup \{p + \alpha * (\frac{|u|u.mean + |v|v.mean}{|u| + |v|} - p)\}$
 $insert(M, w)$
 $w.closest := x // x$ is an arbitrary cluster in $Z//$
 for each $x \in Z$ **do**{
 Party A and B jointly apply algorithm 3 to secure compare
 if $dist(w, x) < dist(w, w.closest)$
 $w.closest := x$
 if $x.closest$ is either u or v {
 if $dist(x, x.closest) < dist(x, w)$
 $x.closest := closest_cluster(T, x, dist(x, w))$
 else
 $x.closest := w$
 else if $dist(x, x.closest) > dist(x, w)$ {
 $x.closest := w$
 }
 }
 $insert(Z, w)$
 }
end

Algorithm 2 Privacy-preserving find minimum distance

Input two distance partitioned by party A and B
Output the one has smaller distance
begin
 Party A sends $\{E_A(a_i^p)\}$ to party B for all p_{th} attribute belong to party A and all b_j^p belong to party B.
 Party A also sends $\{E_A(a_j^q)\}$ to party B for all q_{th} attribute belong to party A and all b_i^q belong to party B.
 Party B calculates $\{E_A((b_i^\beta - b_j^\beta)^2 - (2a_i^p \times b_j^p + 2a_j^q \times b_i^q) + R_B)\}$ and sends this to party A
 Party A calculates $\{E_A((a_i^\alpha - a_j^\alpha)^2 + (b_i^\beta - b_j^\beta)^2 - (2a_i^p \times b_j^p + 2a_j^q \times b_i^q) + R_B)\}$ and then decrypts it to get $dist\{i, j\} + R_B$.
 Using same method, party B can get $dist\{i', j'\} + R_A$.
 Party A calculates $R_A + dist\{i, j\} + R_B$
 Party B calculates $R_A + dist\{i', j'\} + R_B$
 By applying algorithm 3, party A and party B can jointly *securecompare*($R_A + dist\{i, j\} + R_B, R_A + dist\{i', j'\} + R_B$) and store the smaller one.

Algorithm 3 Secure Compare two binary number[17]

Input Party A holds $X_A = \{x_A^1, x_A^2 \dots x_A^n\}$. Party B holds $X_B = \{x_B^1, x_B^2 \dots x_B^n\}$
Output $X_A >? X_B$
begin
 Party A sends $\{E_A(x_A^i)\}$ to party B.
 Party B calculates $c_i = E_A(x_B^i - x_A^i + 1 + \sum_{j=1}^{i-1} w_j)$ where $w_j = x_A^j \oplus x_B^j = x_A^j + x_B^j - 2 \cdot x_A^j \cdot x_B^j$
 Party B permutes c_i and send them to party A
 Party A decrypts all c_i using p_A .
if exist any $c_i == 0$
 $X_A > X_B$
else
 $X_A \leq X_B$
end

V. CORRECTNESS AND SECURITY ANALYSIS

We have now described our privacy-preserving CURE algorithm. In this section, we conduct a detailed analysis of the correctness and security of our algorithm.

A. Correctness analysis

After using our privacy-preserving CURE algorithm, Party A and Party B can combine their partial knowledge of the dataset to get the finally clusters. Now we show that this combined result is correct, that is equal to the clusters when using the original algorithm without privacy protection.

Theorem 1. For the clustering process, our algorithm 1 (with the input arbitrarily partitioned by the two parties) produces the clusters that are the same as the corresponding clusters when running the original version of the CURE algorithm with the whole dataset without any privacy protection

Proof. We first show the correctness of algorithm 2. Party A and Party B can jointly compare the two numbers $R_A + dist\{i, j\} + R_B$ and $R_A + dist\{i', j'\} + R_B$ since the correctness of the secure compare two binary number algorithm is guaranteed in [17]. Obviously, the comparison result is same as the comparison result between $dist\{i, j\}$ and $dist\{i', j'\}$, so the correctness of the distance comparison is proven. Now we show that in each iteration, our algorithm can correctly obtain the $R_A + dist\{i, j\} + R_B$. We note that the distance we used in our comparison is the square of the

Euclidian distance. The components of the $dist\{i, j\}$ either belong to Party A or Party B. Since we use van Dijk et al.'s fully homomorphic encryption, we can directly achieve the $E_A(R_A + dist\{i, j\} + R_B)$ through the communication of Party A and Party B. The fully homomorphic property list below has been guaranteed in [15].

- Multiplication homomorphic property $\delta(\epsilon(m_1) \times \epsilon(m_2) \times \epsilon(m_3) \dots \times \epsilon(m_n)) = m_1 \times m_2 \times m_3 \dots \times m_n$
- Addition homomorphic property $\delta(\epsilon(m_1) + \epsilon(m_2) + \epsilon(m_3) \dots + \epsilon(m_n)) = m_1 + m_2 + m_3 \dots + m_n$

Since we let the distance hide behind the two random numbers R_A and R_B , neither of the parties can obtain the intermediate result. This completes the proof of the theorem.

B. Security analysis

In this subsection, we explain why our algorithm is secure in the semi-honest model. In the semi-honest model, we say an algorithm is secure if neither party can learn anything beyond its output from the information obtained throughout the algorithm. In our algorithm, they can only get information from others when they communicate with each other. There are two parts that communicate with each other. In exchanging $E_A(R_A + dist\{i, j\} + R_B)$, since R_B is randomly chosen by Party B, Party A can not learn the value of $dist\{i, j\}$ from $dist\{i, j\} + R_B$. This is also applied to Party B. In secure compare two binary numbers algorithm, the security also has been guaranteed in [17]. Therefore, our algorithm is secure in the semi-honest model.

VI. CONCLUSIONS

In this paper, we have proposed a privacy-preserving algorithm for clustering using representatives when the input data are arbitrarily partitioned between two parties. Our algorithm is correct and provides strong privacy guarantees.

REFERENCES

- [1] Ronald L Rivest, Len Adleman, and Michael L Dertouzos. On data banks and privacy homomorphisms. *Foundations of secure computation*, 4(11):169–180, 1978.
- [2] Geetha Jagannathan and Rebecca N Wright. Privacy-preserving distributed k-means clustering over arbitrarily partitioned data. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 593–599. ACM, 2005.
- [3] Yehuda Lindell and Benny Pinkas. Privacy preserving data mining. In *Advances in Cryptology CRYPTO 2000*, pages 36–54. Springer, 2000.
- [4] Andrew C Yao. Protocols for secure computations. In *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, pages 160–164, 1982.
- [5] Andrew Chi-Chih Yao. How to generate and exchange secrets. In *Foundations of Computer Science, 1986., 27th Annual Symposium on*, pages 162–167. IEEE, 1986.
- [6] Ronald L Rivest, Adi Shamir, and Len Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.
- [7] Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In *Advances in Cryptology*, pages 10–18. Springer, 1985.
- [8] Josh Benaloh and Dwight Tuinstra. Receipt-free secret-ballot elections. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 544–553. ACM, 1994.

- [9] Pascal Paillier. Public-key cryptosystems based on composite degree residuosity classes. In *Advances in cryptology—EUROCRYPT 1999*, pages 223–238. Springer, 1999.
- [10] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 169–178, New York, NY, USA, 2009. ACM.
- [11] Craig Gentry. *A fully homomorphic encryption scheme*. PhD thesis, Stanford University, 2009.
- [12] Martin Hirt and Kazue Sako. Efficient receipt-free voting based on homomorphic encryption. In *Advances in Cryptology—EUROCRYPT 2000*, pages 539–556. Springer, 2000.
- [13] Tingting Chen and Sheng Zhong. Privacy-preserving backpropagation neural network learning. *Neural Networks, IEEE Transactions on*, 20(10):1554–1564, 2009.
- [14] Jaideep Vaidya and Chris Clifton. Privacy-preserving k-means clustering over vertically partitioned data. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 206–215. ACM, 2003.
- [15] Marten Van Dijk, Craig Gentry, Shai Halevi, and Vinod Vaikuntanathan. Fully homomorphic encryption over the integers. In *Advances in Cryptology—EUROCRYPT 2010*, pages 24–43. Springer, 2010.
- [16] Sudipto Guha, Rajeev Rastogi, and Kyuseok Shim. Cure: an efficient clustering algorithm for large databases. In *ACM SIGMOD Record*, volume 27, pages 73–84. ACM, 1998.
- [17] Ivan Damgard, Martin Geisler, and Mikkel Kroigard. Homomorphic encryption and secure comparison. *International Journal of Applied Cryptography*, 1(1):22–31, 2008.

Texture Classification based on Bidimensional Empirical Mode Decomposition and Local Binary Pattern

JianJia Pan*

*Department of Computer Science
Hong Kong Baptist University
Hong Kong

YuanYan Tang*†

†Department of Computer and Information Science
University of Macau
Macau

Abstract—This paper presents a new simple and robust texture analysis feature based on Bidimensional Empirical Mode Decomposition (BEMD) and Local Binary Pattern (LBP). BEMD is a locally adaptive decomposition method and suitable for the analysis of nonlinear or nonstationary signals. Texture images are decomposed to several Bidimensional Intrinsic Mode Functions (BIMFs) by BEMD, which present a new set multi-scale components of images. In our approach, firstly, saddle points are added as supporting points for interpolation to improve original BEMD, and then images are decomposed by the new BEMD to several components (BIMFs). After then, Local Binary Pattern (LBP) in different sizes is used to detect features from different BIMFs. At last, normalization and BIMFs selection method are adopted for features selection. The proposed feature presents invariant while preserving LBP's simplicity. Our method has also been evaluated in CuRet and KTH-TIPS2a texture image databases. It is experimentally demonstrated that the proposed feature achieves higher classification accuracy than other state-of-the-art texture representation methods, especially in small training samples condition.

Keyword: Texture classification, Empirical Mode Decomposition, Local Binary Pattern, Invariant feature

I. INTRODUCTION

Texture analysis is widely recognized as a difficult and challenging computer vision problem. It provides many applications such as remote sensing image, medical image diagnosis, document analysis, and target detection, etc. Recently, using texture methods to face image analysis and motion analysis have presented many applications, which indicate that texture methods can be adopted to many new fields of computer vision problems.

Over the last several decades, there have been many methods proposed for texture classification, such as co-occurrence matrix, Gabor wavelet, Local Binary Pattern [21], maximum response 8 (VZ-MR8) [1], Basic Image Features (BIF) [2] etc. The statistics describing the whole region is then computed from these transformed local descriptors. The Gabor-based filter representation has been shown to be optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency, and is widely used in image analysis [24]. LBP [21] is an operator for image description that is based on signs of different of neighbouring pixels. Varma and Zisserman [1] proposed the statistics VZ-MR8 classification algorithm that uses a rotationally invariant filter bank and

clustering to estimate the full joint probability distribution of filter responses. Basic Image Features presented by Griffin and Lillholm [2] are defined by a partition of the filter-response space of a set of six Gaussian derivative filters and the set of filters describes an image locally up to second order at some scale. Those methods are all state-of-the-art statistics algorithm and present good classification results on many databases.

Among these descriptors, LBP is a widely used local descriptor. It is simplicity and present excellent performance in various texture and face recognition, which has gained increasing attention. However, LBP describes over-local structure in image, and many improvements of original LBP have been proposed [22], [23], [24]. LBP operator is extended to use neighbourhoods for different sizes [21]. The sign and the magnitude of LBP, and the binary code of the intensity of centre pixels were combined together in CLBP [28] to improve texture classification. However, the intensity information is very sensitive to illumination changes; thus, CLBP needs image normalization to remove global intensity effects before feature extraction. Based on local phase and local surface type extracted from Riesz transform, Zhang [35] proposed a rotation invariant LBP feature (M-LBP) for texture classification. By use of Gabor wavelet, the LBP encode the local information and compress the redundancy in the Gabor filtered images in multi-scale and multi-direction and achieve effectiveness in texture representation [4], [3]. For variational applications, many other descriptors based on LBP are proposed [25], [28], [34].

Recently, Empirical Mode Decomposition (EMD), developed by Huang [5], has attracted more and more attentions in past decade and has been used for texture analysis [9]. The EMD method is based on the direct extraction of the energy associated with various intrinsic time scales. Expressed in Intrinsic Mode Functions (IMFs), they are the expansion basis which can be linear or nonlinear as dictated by data. EMD has been used to analyse two-dimensional signals [7], for example, images, which is known as Bidimensional EMD (BEMD).

BEMD presents some better quality than Fourier, wavelet and other decomposition algorithms in extracting intrinsic components of textures because of its data driven property [6], [7]. In this paper, we proposed an efficiency application of saddle points added BEMD [32] combined with LBP in texture

classification, and present the effectiveness of BEMD/BIMFs invariant properties for texture images.

Local Binary Pattern (LBP) is used as texture descriptor to detect the features of texture images' BIMFs. BEMD decomposed the original image to new multi-scale components (Bidimensional Intrinsic Mode Functions). In those new components, LBP histograms can achieve better efficiency than in the original image and present more illumination invariant features to supplement LBP to improve classification accuracy while preserving its simplicity. Experiments show texture image recognition rate based on our method is better than other state-of-the-art texture representation methods.

This paper is an extension of our previous work [32]. In this paper, we further extend the LBP-BEMD feature to variance normalization and BIMFs selection for performance improvement. We also provide a more in-depth analysis and more extensive evaluation.

II. REVIEW OF BEMD

Empirical mode decomposition (EMD) is a data driven processing algorithm, which applies no predetermined filter. EMD is based on the local characteristics scale of data, which is able to perfectly analyse nonlinear and nonstationary signals [5].

Nonstationary signals have statistical properties that vary as a function of time and should be analysed differently than stationary data. Rather than assuming that a signal is a linear combination of predetermined basis functions, in EMD, the data are instead thought of as a superposition of fast oscillations onto slow oscillations. EMD identifies those oscillations that are intrinsically present in the signal and produces a decomposition using these modes as the expansion basis.

EMD decomposes signal into components called Intrinsic Mode Functions (IMFs) satisfying the following two conditions [5]: (a)The numbers of extrema and zero-crossings must be either equal or differ at most by one; (b)At any point, the mean value of the envelope defined by the local maxima and the envelope by the local minima is zero.

Huang [5] also proposed an algorithm called 'sifting' to extract IMFs $J_k(t)$ from the original signal $f(t)$:

$$f(t) = \sum_{k=1}^K J_k(t) + r_K(t) \quad (1)$$

Where $J_k(t)$, $k = 1, \dots, K$ is IMFs and r_K is the residue.

The EMD is originally developed for one-dimensional (1D) data. Nunes.[9] firstly extended it to two dimensional BEMD, and decomposed images to bidimensional IMFs (BIMFs). The Bidimensional EMD (BEMD) process is conceptually the same as the one dimension EMD. The main process of the BEMD can be described as:

Step 1 Identify the local extrema (both maxima and minima) of the image $I(x, y)$;

Step 2 Generate the 2D envelopes by connecting maximum points (respectively, minima points) using surface interpolation. The local mean m is the mean of the two extrema envelopes. Follow Damerval's work [11], Delaunay triangulation and then cubic interpolation on triangles is used in our work;

Step 3 Subtract out the mean from the image to get a proto-BIMF $r = I - m$, judge whether r is a BIMF, if it is, go to Step 4. Otherwise, repeat Step 1 and Step 2 using the proto-BIMF r , until the latest proto-BIMF turns to be a BIMF;

Step 4 Input the proto-BIMF r to the loop from Step 1 to Step 3 to get the next remained BIMFs until it cannot be decomposed further.

After the BEMD, the decomposition of the image can be rewritten as following form:

$$I(i, j) = \sum_{k=1}^K d_k(i, j) + r(i, j) \quad (2)$$

The $d_k(i, j)$ is the BIMFs of the images, and $r(i, j)$ is the residual function.

Although the discussions about EMD/BEMD lack concrete theoretical foundation until now [14], numerous tests demonstrated empirically that EMD is a powerful tool for the analysis of nonlinear and nonstationary data, especially for time-frequency-energy representations [8], [14], [16]. For two dimensional image, there are also some successful applications [15], [19]. In this work, we fine down the saddle points added BEMD combined with LBP features proposed in our previous work [32], and provide a more in-depth analysis and more extensive evaluation.

III. BEMD BASED ON SADDLE POINTS

One practical implementations in BEMD is the local extrema points detection. Which points should be detected as supporting points for the interpolation is an open problem. Mathematical morphology is used to detect local maxima and minima points in Nunes method[9]. Further, the local neighbour location method is also proposed for extrema detection [13]. However, by use of these methods, saddle points may not be detected. Saddle points are local maximum and local minimum points evaluated in different directions, and they also give important supporting features about the local variation of the original function. We added saddle points as supporting points for interpolation, which provided more significant components for texture classification.

In discrete condition, a point $u(x, y)$ in a 2D matrix U is a *saddle point* $u_{saddle}(x, y)$, if the product of the eigenvalues of the Hessian matrix is negatives:

$$u_{saddle}(x, y) = \{u(x, y) | u_{xx}u_{yy} - u_{xy}^2 < 0\} \quad (3)$$

After detecting the saddle points, neighbour location method [13], [31] is used to detect the local maxima or minima points. In ordinary BEMD methods [6], [7], mathematical morphology is used to detect local maxima and minima points, but we found the numbers of extrema points reducing fast. It

means that the component will be too smooth to detect any signification extrema points after one or two times 'sifting'. To improve local extrema points detection, neighbour location method is used to detect extrema points. In this method, a data point $u(x, y)$ is considered as a local maximum (or. local minimum) if its value is strictly larger (or. lower) than the value of u at the nearest neighbours of points (x, y) .

Let the window size for local extrema determination be $(2w + 1) \times (2w + 1)$, then

$$u(x, y) = \begin{cases} u_{max} & \text{if } u(x, y) > u(i, j), \\ u_{min} & \text{if } u(x, y) < u(i, j). \end{cases} \quad (4)$$

Where $\forall(i, j) \in W(x, y)$ and $W(x, y) = \{(i, j) | (x-w) \leq i \leq (x+w), (y-w) \leq j \leq (y+w), i \neq x, j \neq y\}$. From experiment and following the method in [13], we use the 3×3 window ($w = 1$). We find that result is an optimum extrema map for given images. The larger windows can be used in some other applications, but it will lead to a smaller number of extrema points for given texture images.

After then, the detected saddle points are added to maxima or minima points sets. In saddle points set U_{saddle} , a saddle point location (x, y) 's neighbourhood window is $U(k, l) = \{u | (x-T) \leq k \leq (x+T), (y-T) \leq l \leq (y+T)\}$.

$$\begin{aligned} u_{saddle}(x, y) \in U_{max} & \text{ if } N_{max} > N_{min}, \\ u_{saddle}(x, y) \in U_{min} & \text{ if } N_{max} < N_{min}. \end{aligned} \quad (5)$$

where u_{saddle} is saddle points set element (point), U_{max} is maximum points set, U_{min} is minimum points set, N_{max} is the number of maxima points in window $U(k, l)$, and N_{min} is the number of minima points in window $U(k, l)$. It means that if number of maxima points is more than minima points in window $U(k, l)$, saddle point is considered to be maxima point, and vice verse. In experiments, the window size is 5×5 ($T = 2$). The recognition performance is nonsensitive to this saddle point location windows size. Experimental result about relationship between recognition performance and the windows size is shown in Section V-B.

These three type points, saddle points, neighbour local maxima and neighbour local minima points, are detected from image and used as supporting points for BEMD's interpolation. As shown in Section IV-B, by use of saddle points added BEMD, texture images are decomposed into Bidimensional Intrinsic Mode Functions ($BIMF_s$), which represent images' multi-scale components. The saddle points added BEMD detected more details (high local frequencies of oscillation) of images and contributed the performance of texture images classification.

IV. TEXTURE DESCRIPTOR BASED ON BEMD AND LBP

To analyse and classify texture images, we propose using LBP descriptor to extract local features from decomposed BIMFs. And then, the variance normalization and BIMFs selection are employed for performance improvement.

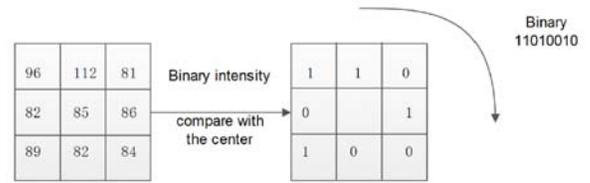


Fig. 1. The LBP operator

A. Local Binary Patterns (LBP)

LBP operator is originally developed for texture description. The operator assigns a label to every pixel of an image by thresholding the 3×3 -neighbourhood of each pixel with the centre pixel value and considering the result as a binary number. Then histogram of the labels can be used as a texture descriptor [21].

The form of the resulting 8-bit LBP code can be defined as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^7 s(u_n - u_c) 2^n \quad (6)$$

where u_c corresponds to the gray value of the centre pixel (x_c, y_c) into gray values of the 8 neighbourhood pixels, and function $s(m)$ is defined as:

$$s(m) = \begin{cases} 1 & \text{if } m \geq 0, \\ 0 & \text{if } m < 0. \end{cases} \quad (7)$$

LBP presents that it will be not affected by any monotonic gray-scale transformation which preserves the pixel intensity order in a local neighbourhood. Each bit of LBP code has the same significance level and that two successive bit values may have a totally different meaning.

To deal with textures at different scales, LBP operator is later extended to use neighbourhoods for different sizes [21]. The local neighbourhood is extended to as a set of sampling points evenly spaced on a circle centred at the pixel to be labelled allows any radius and number of sampling points [22]. If a sampling point is not in the centre of a pixel, it will be rebuilt by bilinear interpolation. The notation (P, R) is defined as the pixel neighbourhood which means P sampling points on a circle of radius of R . Figure.2 shows an example of circular neighbourhoods.

The number of patterns of original LBP grows with respect to the neighbourhood size, to address this problem, Ojala [21] observed that some patterns are more common than others, which is known as uniform patterns ('u2'). The number of transition between zero and one in uniform pattern at most two. For example, the patterns 01111000 and 11001111 are uniform whereas the patterns 11001001 and 01010011 are not. It is measured by:

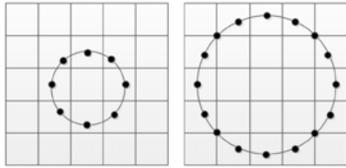


Fig. 2. The circular (8,1) (16,2) neighbourhoods

$$U(LBP_{P,R}) = |s(u_{P-1} - u_c) - s(u_0 - u_c)| + \sum_{n=1}^{P-1} |s(u_n - u_c) - s(u_{n-1} - u_c)| \quad (8)$$

In uniform LBP, there is a separate output label for each uniform pattern and all the non-uniform patterns are assigned to a single label. Thus, the number of different output labels for mapping for patterns of P bits is $(P(P-1) + 3)$ [21]. For instance, the uniform mapping produces 59 output labels for neighbourhoods of 8 sampling points, and 243 labels for neighbourhoods of 16 sampling points. In the following, the mentioned LBP patterns are all uniform patterns.

B. LBP histograms of BIMFs

In this section, the proposed LBP via BIMFs feature frame is introduced. Figure.3 shows an example of texture image and its components BIMFs.

BEMD decomposes an image into its BIMFs basically on local frequency or oscillation information. The first BIMF contains the highest local frequencies of oscillation, the final BIMF contains the lowest local frequencies of oscillation and the residue contains the trend of the data. Corresponding high-frequency components are more robust to illumination changes [30]. BIMFs of image present a set of components of image from high-frequency to low-frequency. At the same time, the BEMD decomposition is an adaptive decomposition method. It is different from wavelet-based multi-scale analysis that characterizes the scale of a signal event using pre-specified basis functions. Moreover, corresponding BIMFs by saddle points added BEMD are able to capture more representative features of the original signal, especially more singular information in high frequency ones.

At the same time, LBP is a nonparametric method, which means that no prior knowledge about the distributions of images is needed.

We use the following procedure to extract texture features:

Firstly, the original image I is decomposed into its BIMF ($BIMF_s(i)$) by use of the saddle points added BEMD:

$$I = \sum_{i=1}^K BIMF_s(i) + r_N \quad (9)$$

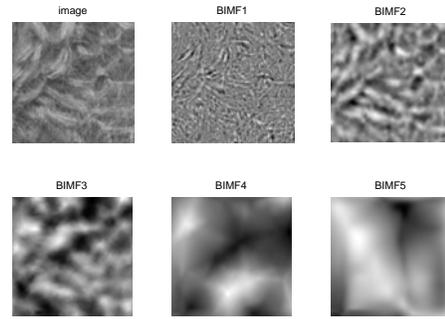


Fig. 3. Texture image and its BIMFs

Secondly, as we can find from texture images' BIMFs (Figure.3), the first and the second BIMF (higher BIMFs) remain the main detail of original image, and the last BIMFs (lower BIMFs) represent information in large scale. In our experiment, histograms of different size LBP ($LBP_{8,1}$ and $LBP_{16,2}$) for different BIMFs are mixed and the best combination is selected experimentally. All LBP patterns used in our algorithm are uniform patterns [21].

Thirdly, the LBP histograms of different BIMFs are normalized. Variance is a measure of how far a set of numbers are spread out from each other. Because we use variational size of LBPs to describe the BIMFs, the distributions of different BIMFs' LBP histograms are incongruous. To normalize LBP histograms, Variance-normalized LBP is defined as:

$$VLBP_{P,R}^{BIMF_s(i)} = LBP_{P,R}^{BIMF_s(i)} / VAR(LBP_{P,R}^{BIMF_s(i)}) \quad (10)$$

where $VAR(LBP_{P,R}^{BIMF_s(i)})$ is the variance of $LBP_{P,R}$ histogram of $BIMF_s(i)$:

$$VAR(x) = \sum_{i=1}^n p_i (x_i - \mu)^2 \quad (11)$$

μ is the expected value $\mu = \sum_{i=1}^n p_i x_i$ and p_i is the probability of x_i . $LBP_{P,R}$ describes the local feature of BIMFs and VAR describes the local variance.

Lastly, in EMD/BEMD literatures, there are findings that residual show trend of the whole signal/image. Figure.4 shows texture image samples and their BIMFs.

The higher BIMFs capture the detail information of original image, and the lower BIMFs capture the coarse contour information. Especially, the illumination and pose variety mainly appears in the residue. Therefore, this indicates that the lower BIMFs are sensitive to the variety. It is well understood that the variety effects can be reduced or eliminated by removing these lower BIMFs and residue. So we can just detecting the first two BIMFs for feature detection, and their LBP histograms are concatenated as the feature vector of image, which is Variance-normalized Saddle points added BEMD LBP:

$$VSBEMDLBP = \{VLBP_{P,R}^{BIMF_s(i)} | i = 1, 2\} \quad (12)$$

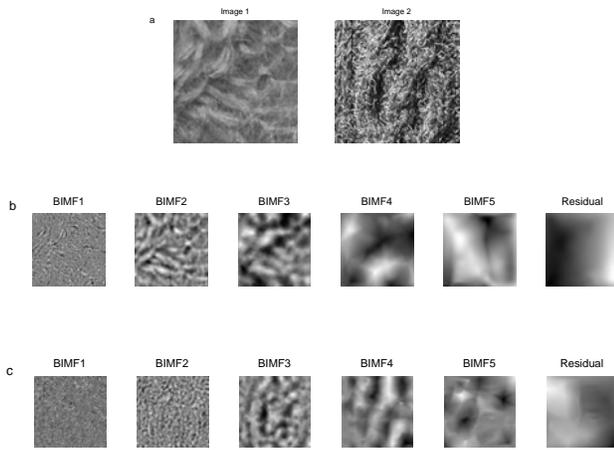


Fig. 4. Images and their BIMFs (a)Images from same class of KTH-TIPS2a (b)BIMFs of Image 1 (c)BIMFs of Image 2

There are many combination choices to combine different size $LBP(P, R)$ with different $BIMF_s(i)$. To find the optimum combination of $LBP(P, R)$ with $BIMF_s(i)$, different combination choices are used and their classification accuracy are compared. Some discussion and experimental result will be presented in Section V-B.

V. EXPERIMENT AND DISCUSSION

To validate the effectiveness of proposed VSEMDLBP feature, we carried out a series of experiment on two large databases compared with other methods: KTH-TIPS2a database [36] and CURET database [37]. Nearest neighbourhood classifier (NN) is used for classification.

A. Databases and dissimilarity measurement

The KTH-TIPS2a database [36] is a database widely used for texture classification and material categorization. KTH-TIPS2a database contains 4 samples of 11 different materials. The sample images at 9 different scales, 12 lighting and pose setups. It contains 11 texture classes with 4572 images. Some examples from different classes are shown in Figure.5. It appears small inter class variations between textures and large intra class variations in the same class. In the top row, all images are of the same texture while in different scales and lighting/pose setups. In the bottom row, the images appear similar and yet they are belonging to different classes.

The images are 200×200 pixels in size (as in ref.[27], we did not include those images which are not of this size, so the experimental data is 10 class with 396 samples pre class, totally 3960), and all of images are transformed into 256 gray levels.

The CURET database [37] contains images of 61 materials and includes many surfaces commonly seen in our environment [24]. Each of the materials in the database has been imaged under different viewing and illumination conditions. The effects of surface normal variations such as specularities, reflections and shadowing are evident. This database also includes some man-made textures, and is highlighted due to abundant imaging

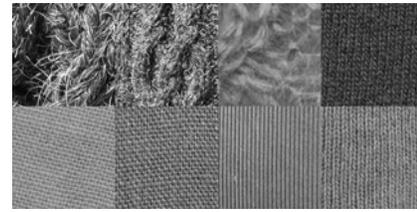


Fig. 5. Some examples from KTH-TIPS2a database

TABLE I. CLASSIFICATION ACCURACY OF PROPOSED METHODS WITH DIFFERENT SIZE WINDOWS IN THE DATABASES

windows size	KTH-TIPS2a	CURET
3×3	96.36%	97.61%
5×5	96.96%	98.00%
7×7	96.30%	97.87%
9×9	96.26%	97.68%
11×11	96.42%	97.33%

conditions. These make it far more challenging and become a benchmark widely used to assess classification performance.

There are 118 images which have been shot from a viewing angle of $< 60^\circ$. Follow ref.[24], in these 118 images, we selected 92 images, from which a sufficiently large region could be cropped (200×200) across all texture classes. And then they are converted all the cropped regions to gray level.

We use χ^2 statistic to measure the dissimilarity of sample and model histograms. Thus, a test sample x_t will be assigned to the class of model x_m that minimizes:

$$D(x_t, x_m) = \sum_{n=1}^N (x_t(n) - x_m(n))^2 / (x_t(n) + x_m(n)) \quad (13)$$

where N is the number of bins, and $x_t(n)$ and $x_m(n)$ are the values of the sample and model histogram at the n^{th} bin, respectively.

B. Parameters selection and feature combination selection in experiment

In section III, we proposed to combine the saddle points to maximum points set or minimum points set based on the numbers of maximum points and minimum points in the saddle point location neighbourhood windows. In this section, we firstly gave some experiments to show the relationship between the size of saddle point location windows and the classification performance. The test data sets are KTH-TIPS2a (40 samples per class for training and 356 samples per class for testing) and CURET (46 samples per class for training and 46 samples per class for testing). The used VSBEMDLBP feature is $VLBP_{8,1}^{BIMF_s(1)} VLBP_{8,1}^{BIMF_s(2)} VLBP_{16,2}^{BIMF_s(1)}$.

As Table.I shows, the classification performance is non-sensitive to the size of saddle point location windows. The value of T don't effect the total number of extrema points. T just effects the distribution of number of added saddle points between maximum set and minimum set. When T changes from 1 to 5 (windows size changes from 3 to 11), the different number of saddle points between maximum set and minimum set accounts for 5-10 percent of total number of saddle points,

TABLE II. CLASSIFICATION ACCURACY OF PROPOSED METHODS WITH COMBINATIONS OF DIFFERENT $LBP_{P,R}$ WITH DIFFERENT $BIMF_s(i)$ IN THE DATABASES

Combination	KTH-TIPS2a	rank	CURET	rank
$V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$	96.96%	2	98.00%	1
$V_{8,1}^1 V_{8,1}^2$	98.27%	1	97.39%	3
$V_{8,1}^2 V_{16,2}^1$	96.38%	3	97.92%	2
$V_{8,1}^1 V_{16,2}^1$	93.19%	9	96.34%	5
$V_{8,1}^1$	93.17%	10	94.38%	8
$V_{16,2}^1 V_{16,2}^2$	92.37%	11	96.91%	4
$V_{8,1}^1 V_{8,1}^2 V_{16,2}^1 V_{16,2}^2$	95.62%	4	92.67%	10
$V_{8,1}^1 V_{8,1}^2 V_{16,2}^2$	94.41%	5	86.77%	14

thus T 's influences on recognition result is small. For other combination features, the same conclusion can be found. In the following experiments, the window size is fixed to 5×5 ($T = 2$).

Further, in Section.IV-B, we proposed the VSBEMDLBP features combining the $LBP_{P,R}$ and $BIMF_s(i)$. The performance of different combination is different. Experiments to 'optimum' the choice of combining different $LBP_{P,R}$ with different $BIMF_s(i)$ is shown in Table.II. The test data sets are KTH-TIPS2a (40 samples per class for training and 356 samples per class for testing) and CURET (46 samples per class for training and 46 samples per class for testing). Because there are many different combinations, we just reported the top five combinations result in the databases.

In order to simplify the expression, $V_{P,R}^i$ is short for the $VLBP_{P,R}^{BIMF_s(i)}$ and $V_{P,R}^1 V_{P,R}^2$ means to concatenate the $VLBP_{P,R}^{BIMF_s(1)}$ and $VLBP_{P,R}^{BIMF_s(2)}$ as the feature vector, etc.

Form TableII, we can find that the best performances are focus on three combinations ($V_{8,1}^1 V_{8,1}^2$, $V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$ and $V_{8,1}^2 V_{16,2}^1$) for different databases. It indicated that the high frequency $BIMF_s(1)$ and $BIMF_s(2)$ present more representative features of the original image, and the high-frequency components are more robust to variant of texture image.

At the same time, other combinations' recognition performance varied in different database, but the top five combinations still can achieve good result. As we discussed before, the lower BIMFs are sensitive to the variety. The top 5 results all don't contain the lower BIMFs ($BIMF_s(4)$ and $BIMF_s(5)$), and even the $BIMF_s(3)$ (in fact, the combination $V_{8,1}^1 V_{8,1}^3 V_{16,2}^1$ achieved rank 7 in CURET database (95.24%)). This indicated that these BIMFs can be removed to reduce variety effects.

In the following experiment, $V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$ feature is used to compare with other methods in CURET database. $V_{8,1}^1 V_{8,1}^2$ feature is used to compare with other methods in KTH-TIPS2a database. And we also reported the top three $V_{8,1}^1 V_{8,1}^2$, $V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$ and $V_{8,1}^2 V_{16,2}^1$ features classification performance in the databases.

C. Classification result on texture database

After selecting the combination of $VLBP_{P,R}^{BIMF_s(i)}$, we evaluated the classification performance using our methods in the two texture databases: KTH-TIPS2a database and CURET database.

The original LBP [21], improved LBP approach (CLBP [28]) and two statistical approaches (VZ-MR8 [1], BIF [2]) are used as comparison. At the same time, to validate the effectiveness of proposed saddle points added BEMD decomposition, we also carried out LBP combined with other transform methods: Gabor filters -LBP [4], [3], M-LBP (Riesz transform) [35], and original BEMD decomposition [11]. The VBEMD-LBP feature is to detect LBP features by our procedure from BIMFs by used of BEMD (no saddle points added BEMD) proposed in [11] (Ref.[11] just proposed BEMD method rather than the features).

1) *Experiment result on KTH-TIPS2a database:* For KTH-TIPS2a database, we firstly repeated the small training samples approach (1-7 samples per class for training, 389 samples per class for testing) with 100 random combinations as training and testing data and the results are reported as average value and shown in Table.III. It means that the training data is independent from physical, materials, illumination, pose, and scale. This small training samples approach only supports a few partial training samples and little knowledge about the data.

Secondly, we repeated the experiment with 100 different random selection of training and testing data (1-50 samples per class for training, 346 samples per class for testing) and reported the proposed and compared approaches results in a range of training set sizes (as in Crosier[2]), which is shown in Figure.6. In Figure.6 we just show the best result of proposed VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2$) and VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2$). The classifier is Nearest Neighbour classifier (NN).

The small training samples result classification accuracies are listed in Table.III. It can be seen that VBEMD-LBP and VSBEMDLBP have better performance than other methods. When the number of training samples pre class is small compared to the testing samples, in this case, 1-7 training samples ps. 389 testing samples, the recognition rates by other methods are dropped, especially for LBP. This is mainly because there are different scales, lighting and pose setups in KTH-TIPS2a and the number of samples per class is large. The proposed methods achieve the highest recognition rates among all the competing methods. Particularly, it is less sensitive to the small sample size problem.

Secondly, from Figure.6, we can note that the performance ranking of the eight representations tested remains the same regardless of the number of images in the training set. This can be seen as confirming the uncommitted nature of the nearest neighbour classifier used with each of the representations. VSBEMDLBP-columns score highest, followed by VBEMD-LBP, BIF and CLBP representations. When the number of training samples are relatively high such as 50 samples per class, the difference between the recognition rates of VSBE-MDLBP and other methods is getting smaller.

The performance of VZ-MR8 is significantly lower than other approaches. VZ-MR8 textons are trained from 40 samples of per class with 122 textons, the totally textons number is 1220. There are 396 samples per class in KTH-TIPS2a, textons from 40 samples per class maybe not representative enough. More training samples and more textons pre class may improve the result, but the training time and storage space will be very large and the cluster result may not be ideal for classification.

TABLE III. CLASSIFICATION ACCURACY OF VSBEMDLBP AND COMPARED APPROACHES IN THE KTH-TIPS2A DATABASE WITH DIFFERENT NUMBER OF TRAINING SAMPLES

No.of training samples	1	2	3	4	5	6	7
LBP[21] (8,1)+(16,2)	40.14%	46.88%	52.36%	57.17%	60.47%	62.59%	64.50%
CLBP[28] (8,1)+(16,2)	45.37%	53.33%	59.19%	63.39%	67.23%	69.23%	71.08%
Gabor-LBP[4] (8,1)	42.04%	49.19%	54.81%	58.91%	62.52%	64.78%	65.88%
M-LBP [35] (8,1)+(16,2)	45.21%	52.22%	59.01%	62.38%	66.16%	68.37%	69.29%
VZ-MR8[1]	38.60%	43.41%	47.41%	50.60%	52.72%	54.52%	56.24%
BIF[2]	46.98%	56.11%	62.33%	66.77%	70.14%	72.53%	73.42%
VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2$)	59.79%	70.28%	76.04%	79.77%	82.11%	84.09%	87.48%
VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$)	55.08%	66.33%	71.78%	75.90%	78.49%	80.92%	82.17%
VBEMD-LBP ($V_{8,1}^2 V_{16,2}^1$)	52.85%	61.93%	67.58%	71.36%	75.22%	77.15%	78.39%
VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2$)	66.96%	76.39%	81.89%	84.99%	86.53%	89.04%	90.23%
VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$)	62.25%	71.80%	76.56%	80.76%	83.10%	85.02%	86.46%
VSBEMDLBP ($V_{8,1}^2 V_{16,2}^1$)	59.53%	68.36%	74.28%	78.34%	80.40%	82.03%	84.08%

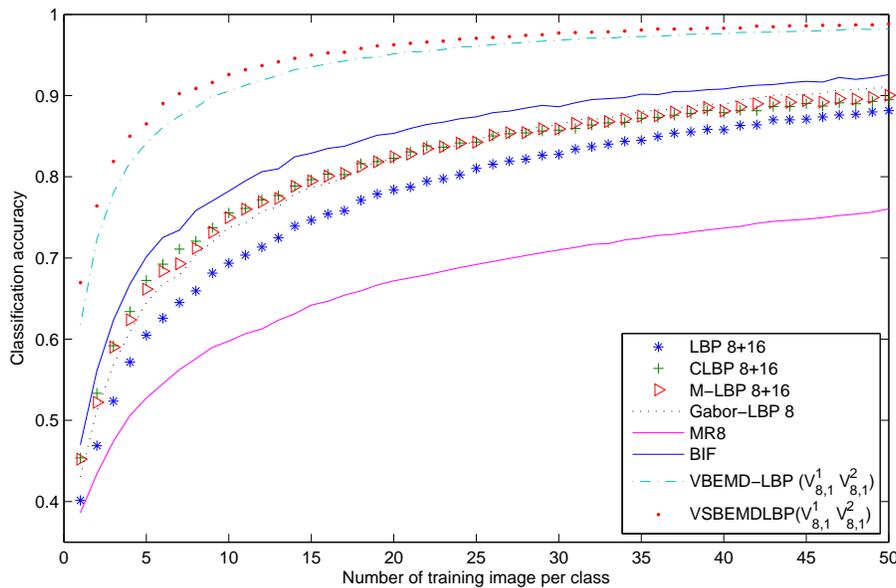


Fig. 6. The mean proportion of correctly classified images over 100 random splits of the KTH-TIPS2a database into training/test data, for a range of training set sizes. The best result for $VSBEMDLBP(V_{8,1}^1 V_{8,1}^2)$ -columns (with 50 training images per class) is 98.83%

Thirdly, there are 396 samples per class in KTH-TIPS2a at different scales, lighting and pose setups, which are high intra-class scatter. Hu [18] and some other researchers have pointed out that the BIMFs' residual is a trend of whole image. For KTH database texture images, the illumination and pose of images can be viewed as trend of images. The high oscillation information ($BIMF(1)$ and $BIMF(2)$) are not only more robust to illumination changes [30] but also more robust to pose and scales changes. In our method, residual and some lower BIMFs are removed in the VBEMD-LBP and VSBEMDLBP features, which means that the variable of samples are reduced by removing the lower BIMFs and residual and achieves a lower intra-class scatter.

2) *Experiment result on CURET database:* For CURET database, there are three training approaches used: 46 training samples C_{46} , 23 training samples C_{23} and 3 training samples C_3 . C_{46} means 46 samples from each class are training data, and other 46 samples are testing data. It is the normal method in previous works [23], [24]. C_{23} means 23 samples from each class be training data, and other 69 samples are testing

data[35]. C_3 means only 3 samples from each class be training data, and other 89 samples are testing data, which is a small training size data. Using the above three settings rather than just one made it possible to better investigate the properties of different operators [24]. The C_{46} data is to simulate the condition that there are enough training samples. C_{23} data is used to simulate the situation of small but comprehensive training set. C_3 data is to simulate the condition that there is only a few partial training samples.

We firstly repeated the three training methods (C_{46} , C_{23} , C_3) with 100 random combinations as training and testing data and the results are reported as average value and shown in Table.IV. Secondly, we repeated the experiment with 100 different random selection of training and testing data (1-46 samples per class for training, 46 samples per class for testing, C_{46}) and reported the combined features results in a range of training set sizes (as in Crosier[2]), which is shown in Figure.7. The small training samples part (1-7 samples per class for training, 46 samples per class for testing, C_{46}) classification accuracies are listed in Table.V. The classifier is

TABLE IV. CLASSIFICATION ACCURACY OF VSBEMDLBP AND COMPARED APPROACHES IN THE CURET DATA

Data setting	C_3	C_{23}	C_{46}
LBP[21] (8,1)+(16,2)	60.14%	88.10%	92.99%
CLBP[28] (8,1)+(16,2)	61.00%	89.91%	94.33%
Gabor-LBP[4] (8,1)	64.82%	90.22%	94.37%
M-LBP [35] (8,1)+(16,2)	67.76%	94.39%	97.59%
VZ-MR8[1]	65.02%	92.48%	96.35%
BIF[2]	73.51%	96.49%	98.56%
VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$)	73.25%	94.76%	97.28%
VBEMD-LBP ($V_{8,1}^2 V_{16,2}^1$)	71.03%	93.70%	96.82%
VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2$)	70.55%	93.14%	96.25%
VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$)	76.29%	95.90%	98.00%
VSBEMDLBP ($V_{8,1}^2 V_{16,2}^1$)	75.18%	95.69%	97.92%
VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2$)	74.63%	94.98%	97.39%

Nearest Neighbour classifier (NN).

Table.IV shows classification accuracy of proposed methods and other methods in CUREt database. By comparing classification rates, we can find that VBEMD-LBP achieve better accuracy that other LBP-based methods and VSBEMDLBP accuracy is higher than other methods except BIF. In the few training samples condition C_3 , accuracies of all methods are dropped, only the proposed features and BIF achieve classification accuracy more than 70% and VSBEMDLBP is even better than BIF. It indicates that the proposed VSBEMDLBP is more robust for real applications where training samples are limited and not comprehensive. On the other hand, Crosier [2] has proposed that the BIF feature achieves 5 – 10% higher classification rate than LBP-based methods in CUREt database, the main superiority of BIF over our method comes from the local feature detector. At the same time, BIF feature is detected by a series of multi-scale filters' responses, which verifies the multi-scale representations will be more efficiency.

We repeated the experiment with 100 different random selection of training and testing data and reported the results in a range of training set sizes, which is shown in Figure.7. We can note that the performance ranking of the eight representations tested remains the same regardless of the number of images in the training set. BIF-columns score highest, followed by VSBEMDLBP, VBEMD-LBP and VZ-MR8 representations. In the small training samples conditions, VSBEMDLBP presents better result than BIF, and their performances are similar after training samples are more than 10. The detail result of the small training samples conditions can be found in Table.V.

In CUREt database, VZ-MR8 textons are trained following the approaches in [1] (10 textons training from 13 samples per class, totally 610 textons). The performance of VZ-MR8 is significantly higher than in another database. VZ-MR8 feature is sensitive to the choice of different number of textons and the textons cluster result in different database.

As Table.III to Table.V show, transform based LBP features present better results compared with LBP and CLBP feature. It shows multi-scale or frequency domain representations extract features in character level, which is proved to be more discriminative feature level. Further, as we discussed before, unlike other priori transform methods (Gabor wavelet, etc.), BIMFs depend on an adaptive decomposition and present a different time-frequency space and more meaningful components.

D. Discussion

In the two databases, the proposed VSBEMDLBP features achieve higher recognition result. Especially, it is less sensitive to the small training sample problem. When the number of samples are relatively high such as 396 samples per class and the number of training samples are relatively low such as 1-7 samples per class in KTH-TIPS2a database, the performance difference between VSBEMDLBP and other methods is higher.

As shown in the Section IV-B, BIMFs of image present adaptive multi-scale components. The higher BIMFs contain the higher local frequencies of oscillation and are more robust to illumination, pose, scale changes. The variety effects can be reduced or eliminated by removing lower BIMFs and residual as in our method. The saddle points added BEMD achieves a better classification result than other transform-based method included the original BEMD. To our best knowledge, the local descriptor based on BIMFs is a new framework to validate the BEMD's powerful compared with other transforms in two dimensional. At the same time, as a decomposition-based method, this framework could be applied to different LBP variants, such as CLBP, Dominant LBP [27], LBP histogram fourier features [23], and other descriptors, for example, BIF.

Although the local descriptor based on BIMFs framework is a powerful method, there are some challenges that should be addressed in the future. The first challenge is choosing the optimum combination of $LBP_{P,R}$ and $BIMF_s(i)$. Though we have validated the performances of the three combinations ($V_{8,1}^1 V_{8,1}^2$, $V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$ and $V_{8,1}^2 V_{16,2}^1$) are better than other combinations in the experiment databases as in section.V-B, the performances of other combinations varied and depended on the texture database. As a future work, more theoretically research is needed.

The next challenge in BEMD's applications is its time complexity. The main time consumption of BEMD is from the many times two dimensional interpolations (the step 2 of Bidimensional EMD (BEMD) process), which is still known to be a time-consuming problem. We compared the time complexity experimentally. The experimental computer is Inter, Core(TM)2, CPU, Q6600 @2.40GHz. The platform is MATLAB R2011a. For a 200×200 pixel image, average BEMD decomposition time was 3.7 seconds, which is more than wavelet (0.06 seconds) and Riesz transform (0.53 seconds). The method with higher time consumption than our method is VZ-MR8 (154 seconds per image for KTH-TIPS2a database and 32 seconds per image for CUREt database), which took a lot of time to cluster textons. BIF took average 3.6 seconds per image to detect feature. LBP and CLBP consumed smaller than 0.05 seconds per image.

Based on the discussion and experimental result of Section IV-B, the best combination of VSBEMDLBP just detected LBP histograms of the first two BIMFs, so in the practice, we can only decompose image to the first two BIMFs and then stop the BEMD processing. Its average decomposition time was 2.5 seconds and faster than BIF. To further reduce computation complex of EMD/BEMD, we have developed a fast decomposition method for one dimensional EMD [33], which only takes 32% time consumption of original EMD and presents more meaningful IMFs. But for two dimensional BEMD, more research is needed to reduce the time while

TABLE V. CLASSIFICATION ACCURACY OF VSBEMDLBP AND COMPARED APPROACHES WITH DIFFERENT NUMBER OF TRAINING SAMPLES (PART) IN C_{46}

No.of training samples	1	2	3	4	5	6	7
LBP[21] (8,1)+(16,2)	41.34%	52.69%	60.31%	65.02%	68.48%	71.40%	73.85%
CLBP[28] (8,1)+(16,2)	41.48%	53.27%	61.15%	66.13%	69.81%	72.91%	75.06%
Gabor-LBP[4] (8,1)	43.98%	55.49%	63.01%	67.71%	71.12%	73.91%	76.17%
M-LBP [35] (8,1)+(16,2)	46.55%	59.99%	67.70%	73.18%	76.90%	79.72%	82.03%
VZ-MR8[1]	43.42%	56.93%	64.35%	70.34%	73.80%	76.66%	78.85%
BIF[2]	50.32%	65.07%	74.44%	78.87%	82.65%	84.99%	86.87%
VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$)	52.43%	65.19%	73.56%	77.75%	80.86%	83.58%	85.38%
VBEMD-LBP ($V_{8,1}^2 V_{16,2}^1$)	52.31%	64.24%	71.16%	76.12%	73.35%	81.56%	83.58%
VBEMD-LBP ($V_{8,1}^1 V_{8,1}^2$)	51.76%	63.46%	70.81%	75.42%	78.38%	80.96%	82.91%
VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2 V_{16,2}^1$)	55.54%	68.46%	76.54%	81.14%	83.83%	86.08%	87.86%
VSBEMDLBP ($V_{8,1}^2 V_{16,2}^1$)	53.96%	67.14%	74.18%	78.90%	82.05%	84.47%	86.14%
VSBEMDLBP ($V_{8,1}^1 V_{8,1}^2$)	54.82%	67.04%	74.24%	78.87%	82.13%	84.43%	86.33%

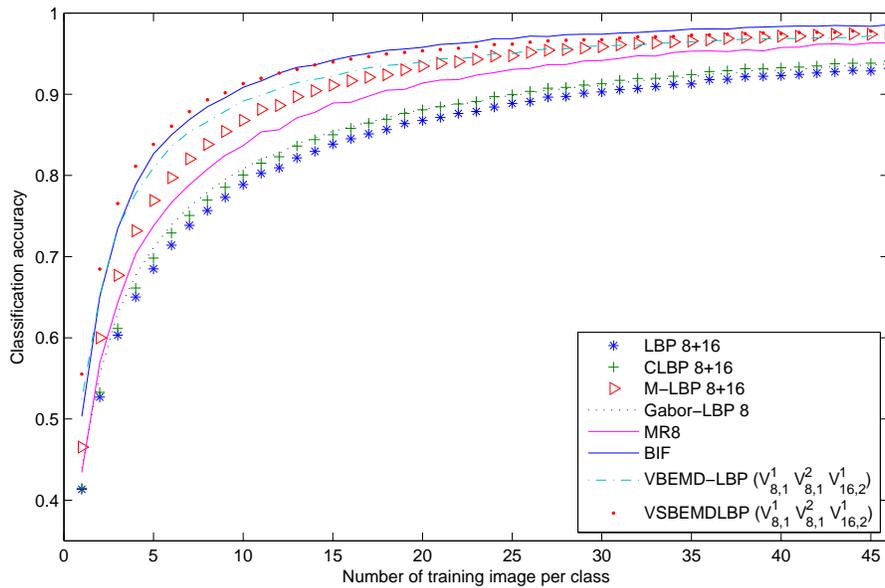


Fig. 7. The mean proportion of correctly classified images over 100 random splits of the CURet database into training/test data, for a range of training set sizes. The best result for BIF-columns (with 46 training images per class) is 98.56%

preserving the decomposition characteristics.

VI. CONCLUSION

Texture analysis is a difficult and challenging computer vision problem. In this paper, a new powerful method (VSBEMDLBP) is proposed for texture classification. An adaptive decomposition method (saddle points added BEMD) is used to supply new components ($BIMF_s$). The saddle points added BEMD detected more details (high local frequencies of oscillation) of images and contributed the performance of texture images classification. At the same time, the higher BIMFs capture the detail information of original image, and the lower BIMFs capture the coarse contour information. Especially, the illumination and pose variety mainly appears in the residue. The higher frequency $BIMF_s$ present more invariant properties for texture classification. In these new adaptive multi-scale components (higher frequency $BIMF_s$), LBP descriptor can achieve better performance than in original images and other transform-based methods. Experiments show the texture image recognition rate based on our method is better than other state-

of-the-art texture representation methods. Especially, it is less sensitive to the small training sample size problem.

ACKNOWLEDGEMENT

The research was supported by University of Macau Research Grants (SRG010-FST11-TYY, MYRG187(Y1-L3)-FST11-TYY, MYRG205(Y1-L4)-FST11-TYY)

REFERENCES

- [1] M. Varma, A. Zisserman, "A statistical approach to texture classification from single image", International Journal of Computer Vision, vol.62(1),pp.61-81, 2005 .
- [2] M. Crosier, LD. Griffin, "Using basic image features for texture classification", International Journal of Computer Vision, vol.88(3), pp.447-460, 2010.
- [3] M. Yang, L. Zhang, "Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary", Lecture Notes in Computer Science, 2010 European Conference on Computer Vision pp.448-461, 2010.

- [4] Y. Su, D. Tao, X.L. Li, X. Gao, "Texture Representation in AAM using Gabor Wavelet and Local Binary Patterns" 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA, pp.3274-3279, 2009
- [5] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", Royal Society.Lond, vol.454, pp.903-1005, 1998.
- [6] N.E. Huang, M.L.C. Wu, S.R. Long, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis", Royal Society A, vol.459, pp.2317-2345, 2003.
- [7] J.C. Nunes, Y. Bouaoune, E. Delechelle, O.Niang, Ph. Bunel, "Image analysis by bidimensional empirical mode decomposition", Image and Vision Computing, vol.21(12), pp.1019-1026, 2003 .
- [8] P. Flandrin, G. Rilling, P. Goncalves, "Empirical mode decomposition as a filter bank", IEEE Signal Processing Letters, vol.11(2), pp.112-114, 2004.
- [9] J.C. Nunes, S. Guyot , E. Dechelle, "Texture analysis based on local analysis of the Bidimensional Empirical Mode Decomposition", Machine Vision and Applications, vol.16(3), pp.177-188, 2005.
- [10] A. Linderherd, "Variable sampling of the empirical mode decomposition of two-dimensional signals", International Journal of Wavelets, Multiresolution and Information Processing., vol.3(3), pp.435-452 , 2005.
- [11] C. Damerval, S. Meignen, V. Perrier, "A fast algorithm for bidimensional EMD", IEEE Signal Processing Letter, vol.12(10), pp.701-704, 2005.
- [12] R. Bhagavatula, S. Marios, M. Acoustics, "Analyzing Facial Images using Empirical Mode Decomposition for Illumination Artifact Removal and Improved Face Recognition", 2007 IEEE International Conference on Speech and Signal Processing, pp.505-508, April 2007.
- [13] M.A.B. Sharif, R.A. Reza, F.K. Jesmin, "Fast and Adaptive Bidimensional Empirical Mode Decomposition Using Order-Statistics Filter Based Envelope Estimation", Journal on Advances in Signal Processing, Article ID 728356, 18 pages, 2008.
- [14] N.E. Huang, Z. Wu, "A review on Hilbert-Huang transform Method and its applications to geophysical studies", Reviews of Geophysics, vol. 46, RG2006, 23 PP., 2008.
- [15] Q.W. Xie, B. Xuan, S.L. Peng, J.P. Li, W.X. Xu, H. Han, "Bandwidth empirical mode decomposition and its application", International Journal of Wavelets, Multiresolution and Information Processing, Vol.6(6), pp.777-798, 2008.
- [16] Z. Wu, N.E. Huang, "On the filtering properties of the Empirical Mode Decomposition", Advances in Adaptive Data Analysis, vol. 2(4), pp.397-414, 2010.
- [17] Z. Wu and N.E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method", Advances in Adaptive Data Analysis, vol.1(1), pp.1-41, 2009.
- [18] X. Hu, W. Xia, S.L. Peng, W.L. Hwang, "Multiple component predictive coding framework of still images", 2011 IEEE International Conference on Multimedia and Expo (ICME), Barcelona, Spain , pp.1-6, July 2011.
- [19] O. Niang, A. Thioune, M.C.E Gueire, E. Delchelle, J. Lemoine, "Partial Differential Equation-Based Approach for Empirical Mode Decomposition: Application on Image Analysis", IEEE Transactions on Image Processing, Vol.21(9), pp.3991-4001, 2012.
- [20] S. Lazebnik, C. Schmid, J. Ponce, "A Sparse Texture Representation Using Local Affine Regions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.27(8), pp.1265-1278, 2005.
- [21] T. Ojala, M. Pietikainen, T.T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with Local Binary Pattern", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24(7), pp.971-987, 2002.
- [22] G. Zhao, M. Pietikainen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29(6), pp.915-928, 2007.
- [23] T. Ahonen, J. Matas, C. He, M. Pietikainen, "Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features", Lecture Notes in Computer Science, Image Analysis, pp.61-70, 2009.
- [24] Y. Guo, G. Zhao, M. Pietikainen, Z. Xu, "Descriptor Learning Based on Fisher Separation Criterion for Texture Classification", Lecture Notes in Computer Science, 2010 Asian Conference on Computer Vision, pp.185-198, 2010.
- [25] Heikkilä, Marko and Pietikäinen, Matti and Heikkilä, Janne, "A texture-based method for detecting moving objects", BMVC, pp.1-10, 2004
- [26] G. Zhao, T. Ahonen, J. Matas, M. Pietikainen, "Rotation-invariant image and video description with local binary pattern features". IEEE Transactions on Image Processing, vol.21(4),pp.1465-1467, 2012.
- [27] S. Liao , M.W.K. Law, A.C.S. Chung, "Dominant Local Binary Patterns for Texture Classification", IEEE Transactions on Image Processing, vol.18(5), pp.1107-1118, 2009.
- [28] Z. Guo, L. Zhang , D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification", IEEE Transactions on Image Processing. vol.19(6), pp.1657-1663, 2010.
- [29] Z. Guo, L. Zhang, D. Zhang, "Rotation invariant texture classification using LBP variance(LBPV) with global matching", Pattern Recognition, vol.43(3), pp.706-719, 2010 .
- [30] Z.B. Zhang , S.L. Ma, D.Y. Wu, "The application of neural network and wavelet in human face illumination compensation", Proc. Advances in Neural Networks, pp.828-835, 2005.
- [31] J.J. Pan, D. Zhang, Y.Y. Tang, "A fractal-based BEMD method for image texture analysis", 2010 IEEE International Conference on Systems, Man, and Cybernetics, Turkey, October, pp.3817-3820, 2010 .
- [32] J.J. Pan, Y.Y. Tang, "Texture analysis based on saddle points-based BEMD and LBP", Lecture Notes in Computer Science, Computer Analysis of Images and Patterns, pp.506-514, 2011.
- [33] J.J. Pan, Y.Y. Tang, "Extremum Mean Empirical Mode Decomposition", 2012 International Congress on Image and Signal Processing, pp.1888-1893 , Chongqing, China, 2012
- [34] F.M. Khellah, "Texture Classification Using Dominant Neighborhood Structure", IEEE Transactions on Image Processing, vol.20(11), pp.3270-3279, 2011.
- [35] L. Zhang, L. Zhang, Z. Guo, D. Zhang, "Monogenic-LBP: A new approach for rotation invariant texture classification", 2010 IEEE International Conference on Image Processing, pp.2677-2680, 2010.
- [36] B. Caputo, E. Hayman, P. Mallikarjuna, "Class-Specific Material Categorisation", 2005 IEEE International Conference on Computer Vision, pp.1597-1604, 2005.
- [37] K.J. Dana, S.K. Nayar, B. van Ginneken, J.J. Koenderink, "Reflectance and texture of real-world surfaces", ACM Transactions on Graphics, Vol.18(1), pp.295-306,, Jan. 1999.

A Bayesian framework for glaucoma progression detection using Heidelberg Retina Tomograph images

Akram Belghith, Christopher Bowd, Madhusudhanan Balasubramanian, Robert N. Weinreb, Linda M. Zangwill
Hamilton Glaucoma Center
University of California San Diego,
La Jolla, California
abelghith@ucsd.edu, cbowd@ucsd.edu, bmadhu@ieee.org, rweinreb@ucsd.edu, lzangwill@ucsd.edu

Abstract—Glaucoma, the second leading cause of blindness in the United States, is an ocular disease characterized by structural changes of the optic nerve head (ONH) and changes in visual function. Therefore, early detection is of high importance to preserve remaining visual function. In this context, the Heidelberg Retina Tomograph (HRT), a confocal scanning laser tomograph, is widely used as a research tool as well as a clinical diagnostic tool for imaging the optic nerve head to detect glaucoma and monitor its progression.

In this paper, a glaucoma progression detection technique is proposed using the HRT images. Contrary to the existing methods that do not integrate the spatial pixel dependency in the change detection map, we propose the use of the Markov Random Field (MRF) to handle a such dependency. In order to estimate the model parameters, a Monte Carlo Markov Chain procedure is used. We then compared the diagnostic performance of the proposed framework to existing methods of glaucoma progression detection.

Keywords—Glaucoma, Markov random field, change detection, Bayesian estimation.

I. INTRODUCTION

Glaucoma refers to a set of eye conditions of great clinical and etiological heterogeneity. It is characterized by the degeneration of optic nerve fibers and often accompanied by an elevated intraocular pressure (IOP). The loss of nerve fibers leads a decrease in the thickness of the retinal nerve fiber layer (RNFL), affects the appearance of the ONH and causes an irreversible damage to the retina. In the course of the disease, the neuroretinal rim gets thinner whereas the optic cup gets bigger [1].

Since the introduction of the ophthalmoscope by Helmholtz in 1851, ophthalmologists have been able to assess the ONH structure associated with glaucoma. Nevertheless, the qualitative clinical assessment of the ONH leads to a considerable inter-observer diagnostic variability [2]. Therefore, development of quantitative measurements for glaucoma detection is important to make the qualitative clinical assessment more objective and reproducible.

For these reasons, sophisticated ocular imaging instruments are providing quantitative parameters of the ONH

in glaucoma such as the Scanning Laser Polarimetry and the Optical Coherence Tomography. In particular, the Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Heidelberg, Germany), a confocal scanning laser tomography, has been commonly used for glaucoma diagnosis since its commercialization 20 years ago [3].

A limited number of studies have been published investigating glaucoma progression detection using HRT images by detecting changes between baseline reference images and follow-up images. In [4], authors assessed the glaucomatous changes using the Topographic Change Analysis (TCA). However, this method requires up to three additional confirmatory follow-up exams to detect changes (progression) [5]. To overcome this requirement, [6] used the Proper Orthogonal Decomposition (POD) for glaucoma detection. However, although this method is successfully applied to HRT images, it does not exploit additional available knowledge, such as spatial dependency among neighboring pixels (*i.e.*, the status of a pixel will depend on the status of its neighborhood). The POD method indirectly utilizes the spatial relationship among pixels by controlling the familywise type I error rate. In [7], a retinal surface reflectance model and a homomorphic filter were used to detect progression from changes in the retinal reflectance over time. Similar to POD, dependency among spatial locations were only indirectly used by controlling familywise type I error rate.

In this paper, we propose a new strategy for glaucoma progression detection. We particularly show that glaucoma detection can be improved if spatial dependency of status of pixels are directly modeled and integrated within the change detection method. In addition, since for each follow-up exam, three follow-up scans are generated by the machine, each follow-up scan is considered as a separate 'information source'. The data-fusion (*i.e.* combining information from three scans per exam) and dependency in the status of pixels in a neighborhood are jointly modeled and addressed using the Markov Random Field (MRF) [8]. Indeed, many studies have tackled the change detection problem by modeling the change detection map as a MRF [9], [10]. The widely used procedure to estimate the different problem param-

ters is the Expectation-Maximization EM algorithm [11]. However, since we used the MRF model with the change detection map as *a priori* for the change detection map, the optimization step is intractable and therefore, we used a Monte Carlo Markov Chain MCMC technique [12] at each EM iteration [13]. The principle of MCMC technique is to generate samples by drawing from the posterior densities [12].

The paper is divided into two sections. In section II, we present our new glaucoma progression detection scheme. In Section III, we present the results of our new algorithm using HRT follow-up exams of participants in the UCSD Diagnostic Innovations in Glaucoma Study. Then, we compare the diagnostic accuracy, robustness and the efficiency of our novel approach to three existing progression detection approaches, topographic change analysis (TCA) [4], proper orthogonal decomposition (POD) [6] and the reflectance based method [7].

II. GLAUCOMA PROGRESSION DETECTION SCHEME

A. Illumination correction

The HRT images can be affected by inhomogeneous background (illumination) and effect may be differing among follow-up exams due to curvature of the retina and differences in the angle of imaging the eye between exams [7]. Although, this problem is not due to glaucoma, it would have an influence on the subsequent statistical analysis and quantitative estimates of glaucoma progression. To this end, a reflectance-based normalization step is performed. Assuming that the optic nerve head is Lambertian, each HRT image I can be simplified and formulated as a product $I = L \times F$ where F is the reflectance and L is the illuminance. Because illumination varies slowly over the field of view, the illuminance L is assumed to be contained in the low frequency component of the image. Therefore, the reflectance F can be approximated by the high frequency component of the image. The reflectance component F describes the surface reflectivity of the retina whereas the illumination component L models the light sources. The reflectance image can then be used as an input to the change detection algorithm [7], [14], [15].

Several methods have been proposed to solve the problem of reflectance and illumination estimation including the homomorphic filtering [16], the scale retinex algorithm [17]–[19] and the isotropic diffusion approach [20]. The reflectance based method of detecting progression uses a homomorphic filter to estimate retinal reflectance [7]. In our new glaucoma progression detection scheme presented in this study, we used the scale retinex approach. The retinex approach has been successfully utilized in many applications, including medical radiography [21], underwater photography [22] and weather images enhancement [23].

B. Change detection

Let us consider the detection of changes in a pair of amplitude images. Change-detection can be formulated

as a binary hypothesis testing problem by denoting the "change" and "no-change" hypotheses as H_1 and H_0 , respectively. We denote by I_0 and I_1 two images acquired over the same scene at times t_0 and t_1 , respectively ($t_0 > t_1$), and coregistered. A difference image R with N pixels was estimated as pixelwise difference between images I_0 (a baseline image) and I_1 (a follow-up image): $R = abs(I_0 - I_1)$. HRT acquires three scans during each exam. Therefore, we obtain three image differences $R(l, i) = \{r(l, i) | l = 1, 2, 3, i = 1, 2, \dots, N\}$. The change detection is handled through the introduction of change class assignments $Q = \{q(i) | i = 1, 2, \dots, N\}$. Accordingly, the posterior probability distribution of change (i.e. $q(i)$) at each pixel location is expressed as:

$$p(q_i = H_j | R, \Theta) = \frac{1}{Z} \exp(-U(H_j | R, \Theta)) \quad (1)$$

where Z is the normalization constant, Θ consists of the model hyperparameters and $U(H_j | R, \Theta)$, $j \in \{1, 2\}$, is the energy function of the MRF model. The energy function is expressed as a linear combination of elementary energies that model both the spatial dependency of pixel classification in each follow-up exam and the information conveyed by all follow up scans (multisource fusion). Note that we have opted for the 8-connectivity neighboring system in our MRF model of pixel classification. The proposed energy function is defined by:

$$U(H_j | R, \Theta) = \sum_{l=1}^3 \gamma_l [\log(p_{j,l}(r(i, l) | q(i) = H_j, \alpha_{j,l}))] + \beta \sum_{i \sim k} \delta(q_i, q_k) \quad (2)$$

where $p_{j,l}(r(i, l) | q(i) = H_j)$ is the probability density function of the i th amplitude difference $r(i)$ belonging to the source l (or HRT scan), where $l = 1, 2, 3$ and conditioned to H_j ; $\alpha_{j,l}$ is the set of the pdf hyperparameters; δ the delta Kroneker function; $k = 1 \dots N$; and γ_l and β are positive parameters. Note that the first part of the energy function models the *a priori* we have on the image differences $R(l)$ conditioned to change and no-change hypotheses which allows us to use the whole information available in every scan (information fusion). The second part models the spatial dependency by the use of the second-order isotropic Potts model with parameter β . Hence, the definition of the energy function favors the generation of homogeneous areas reducing the impact of the speckle noise, which could affect the classification results of the HRT images [24]. The hyperparameters $\beta, \gamma_1, \gamma_2, \gamma_3$ handle the importance of the energy terms. On one hand, β allows us to tune the importance of the spatial dependency constraint. In particular, high values of β correspond to a high spatial constraint. On the other hand, γ_l models the reliability of each HRT follow-up image and it is usually assumed to take on values in $]0, 1]$. This constraint can easily be satisfied by choosing the appropriate *a priori* distribution for γ_l . Note that when $\gamma_l = 0$, all the energy contribution related to the difference-images is removed and hence should be avoided.

Several parametric models can be used to model the distribution of $r(i, l)$ conditioned to H_j . In this work, we opted for the normal distribution. The pdf $p_{j,l}(r(i, l) | q(i) = H_j)$

is then given by:

$$p_{j,l}(r(i,l)|q(i) = H_j, \alpha_{j,l}) = \frac{1}{\sqrt{2\pi}\sigma_{j,l}} \times \exp\left(-\frac{(r(i,l) - \mu_{j,l})^2}{2\sigma_{j,l}^2}\right) \quad (3)$$

where $\mu_{j,l}$ and $\sigma_{j,l}$ stand for the mean and the standard deviation respectively.

The whole set of the model hyperparameters is then given by $\Theta = \{\sigma_{j,l}, \mu_{j,l}, \gamma_l, \beta\}$. In the absence of information prior knowledge, the following priors were used to generate model hyperparameters:

$$p(\sigma_{j,l}^2) = \frac{1}{\sigma_{j,l}} \quad (4)$$

$$p(\mu_{j,l}) = \frac{1}{\sqrt{2\pi}\varphi} \exp\left(-\frac{\mu_{j,l}^2}{2\varphi^2}\right) \quad (5)$$

The non-negativity of the the hyperparameters $\{\gamma_l, \beta\}$ is guaranteed through the use of the exponential densities:

$$p(\gamma_l) = \frac{1}{\kappa} \exp\left(-\frac{\gamma_l}{\kappa}\right) \quad (6)$$

$$p(\beta) = \frac{1}{\eta} \exp\left(-\frac{\beta}{\eta}\right) \quad (7)$$

Note that the values of $\{\eta, \varphi, \kappa\}$ are fixed empirically but do not really influence the results.

To estimate the model parameters and hyperparameters, we propose the use of a Monte Carlo Markov Chain (MCMC) procedure. Specifically, the principle of MCMC method is to generate samples drawn from the posterior densities and then to achieve parameter estimation. We use a Gibbs sampler based on a stationary ergodic Markov chain allowing us to draw samples whose distribution asymptotically follows the *a posteriori* densities.

C. Inference scheme

The Gibbs sampler facilitates efficient sampling from a n-dimensional joint distribution when knowledge about its conditional distributions is available. The aim of the approach is to generate a sample of values from the posterior distribution of the unknown parameters. The inference scheme consists of running the Gibbs sampler for many iterations and at each iteration an estimate for each unknown parameter is produced. At each iteration, the parameter estimate from the last iteration is used to produce new estimate by assuming that the current values for the other parameters are the true values. The main Gibbs sampler steps are described in algorithm 1.

Algorithm 1 Sampling Algorithm

- 1- Initialization of $\Theta^{[0]}$
- 2- For each iteration h repeat:
 - i) Sample $\beta^{[h]}$ from $p(\beta)$
 - ii) Sample $\gamma_l^{[h]}$ from $p(\gamma)$
 - iii) Sample $\sigma_{j,l}^{[h]}$ and $\mu_{j,l}^{[h]}$ from $p(\sigma_{j,l})$ and $p(\mu_{j,l})$
 - iv) Sample $\gamma_l^{[h]}$ from $p(\gamma)$
 - v) Create a configuration of Q basing on R
 - vi) Calculate $p^{[h]}(q_i = H_j | R, \Theta^{[h]})$
 - vii) Repeat i) to vi until Convergence criterion is satisfied

For convergence, we used a burn-in period of $h_{min} = 500$ iterations followed by another 1000 iterations for convergence ($h_{max}=1500$). The change detection map Q is then estimated using the maximum a posteriori MAP estimator: $\hat{Q} = \underset{H_j}{\operatorname{argmax}} \bar{p}_{H_j}$, where $\bar{p}_{H_0} = \frac{1}{h_{max} - h_{min}} \sum_{h=h_{min}+1}^{h_{max}} p^{[h]}(q_i = H_0 | R, \Theta^{[h]})$ and $\bar{p}_{H_1} = 1 - \bar{p}_{H_0}$ and $p^{[h]}(q_i = H_j | R, \Theta)$ is the estimated pdf at the iteration h .

III. EXPERIMENTS

This section aims at validating the proposed framework. Datasets used for validation are presented in the next sub-section. In sub-section III-B, the intensity normalization algorithm reliability is presented. Change detection results on semi-simulated datasets are presented in sub-sections III-C. Finally, the glaucoma progression detection results using datasets are presented in sub-section III-D.

A. Datasets

The proposed framework was experimentally validated with clinical datasets. The three study groups in the clinical datasets have been described previously [6]. In brief, all eligible participants were recruited from the University of California, San Diego Diagnostic Innovations in Glaucoma Study (DIGS) with at least 4 good quality HRT-II exams, at least 5 good quality visual field exams and at least 2 good quality stereo-photographs of the optic disk were included in the study (267 eyes of 202 participants). Two hundred and forty six eyes from 167 glaucoma patients were included as progressing or non-progressing. Thirty six eyes from 33 participants progressed by stereo-photographs and / or showed likely visual field (Progressors) and the rest of the 210 eyes from 148 participants were considered non-progressing (Non-progressors). An additional 21 eyes from 20 participants were normal eyes with no history of IOP > 22 mmHg, normal appearing optic disk by stereo-photography and visual field exams within normal limits (median age of 62.7 years and median HRT-II follow-up of 0.5 years). The UCSD Institutional Review Boards approved the study methodologies and all methods adhered to the Declaration of Helsinki guidelines for research in human subjects and the Health Insurance Portability and Accountability Act (HIPAA).

HRT exams from each study eye were of size 360 x 360 pixels. For each eye, several exams were performed over time. For each exam, the baseline image and the follow-up image are co-registered using built in instrument software and saved together.

Since no change detection map is available for the clinical dataset, we generated two semi-simulated datasets to assess the proposed change detection method. The first semi-simulated dataset (dataset1) was constructed from four normal HRT images. Changes were simulated by permuting 5%, 7%, 10% and 12% of image regions in the four images respectively (Figure 1). The second simulated dataset (dataset2) was constructed from another set of four normal HRT images. Changes were simulated by modifying the intensities of each 15 x 15 pixel-sized regions in the four images randomly with 0.5%, 0.75%, 1% and 1.25% probability of occurrence respectively. The intensities were modified by multiplying randomly the real intensities by 0.5, 0.75, 1.5 or 2 with a 0.25% probability (Figure 2).

B. Intensity normalization algorithm assessment

To assess the proposed illumination correction normalization algorithm and particularly the use of the retinex method for the reflectance estimation, five different methods were used to normalize the image intensities: the proposed method [18] (filter size=20 pixels), the homomorphic filtering method [16] (standard deviation=2 and filter size=20 pixels), the isotropic diffusion method [20] (smoothness constraint=7 pixels), the Discrete Cosine Transform (DCT) method [25] (number of components is 40) and the wavelet-based method (with Daubechies wavelet and three level decomposition) [21]. Fig. 3 presents the intensity normalization of a baseline and follow-up images using the five methods.

All the above methods were evaluated using the simulated datasets, dataset1 and dataset2. For evaluation, we use False Alarm PFA, Missed Detection PMD and Total Error PTE measurements computed in percentage and defined as: $PFA = \frac{FA}{N_F} \times 100\%$, $PMD = \frac{MD}{N_M} \times 100\%$ and $PTE = \frac{MD+FA}{N_M+N_F} \times 100$, where FA stands for the number of unchanged pixels that were incorrectly determined as changed, N_F is the total number of unchanged pixels, MD the number of changed pixels that were incorrectly detected as unchanged, N_M is the total number of changed pixels. Table. I presents false detections, missed detections and total errors for the simulated datasets. As one can see, the retinex reflectance algorithm performs better than the other normalization methods.

C. Change detection results

In order to emphasize the benefit of the proposed change detection algorithm and particularly the use of the Markov model to handle pixel spatial dependency, we compared the proposed method to the following two kernel-based methods and a Bayesian threshold-based method:

Table with 4 columns: Method, False detection, Missed detection, Total Errors. Rows include Retinex, DCT, Isotropic diffusion, Homomorphic filtering, and Wavelet.

TABLE I. FALSE DETECTION, MISSED DETECTION AND TOTAL ERRORS FOR: THE PROPOSED INITIALIZATION METHOD, THE HOMOMORPHIC FILTERING, THE ISOTROPIC DIFFUSION METHOD, THE DCT METHOD AND THE WAVELET METHOD.

Table with 4 columns: Method, False detection, Missed detection, Total Errors. Rows include Proposed method, SVM, SVDD, and Threshold.

TABLE II. FALSE DETECTION, MISSED DETECTION AND TOTAL ERRORS RESULTING FROM: THE PROPOSED CHANGE DETECTION METHOD, THE SVM METHOD, THE SVDD METHOD AND THE THRESHOLD METHOD.

- The Support Vector Data Description SVDD [26] with the Radial Basis Function RBF kernel,
• The Support Vector Machine [27] with the RBF Gaussian kernel,
• A Bayesian threshold-based method [28],

Note that we used the retinex-based intensity normalization for all methods. First, we applied these methods on the semi-simulated datasets. The proposed method tends to perform better than the above methods. This means that the proposed Markov a priori we considered improves the change detection results by utilizing the information about spatial dependency of pixels in our detection scheme.

D. Glaucoma progression detection

We are now faced with the problem of framework validation on clinical datasets. We estimated the sensitivity and the specificity of detecting glaucoma progression as:

sensitivity = TP / (TP + FN); specificity = TN / (TN + FP)

where TP stands for the number of true positive identifications, FN the number of false negative identifications and FP the number of false positive identifications.

In order to emphasize the benefit of the proposed glaucoma progression detection scheme, we have compared the proposed framework with three other published methods: the Topographic Change Analysis (TCA) method [5], the Proper Orthogonal Decomposition POD method [6] and the reflectance based method [7] (Table III). Note that an eye is considered as progressor when, in one exam, the number of changed pixels with an increase of intensity (loss of retinal height) is greater that 5% of the number of pixels within the optic disc [6]. The proposed framework had approximately twice the specificity in the non-progressing eyes

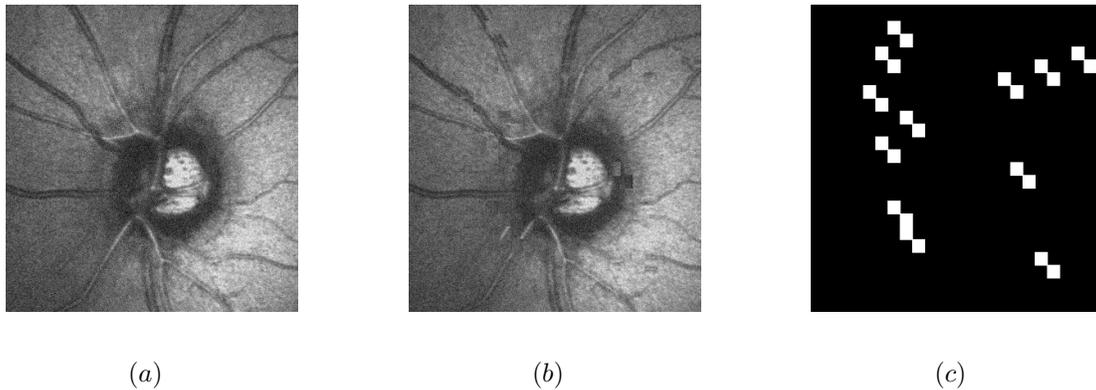


Fig. 1. Semi-simulated image by the permutation of image regions: (a) the original image (b) the simulated image (PSNR=28 dB) and (c) the ground truth.

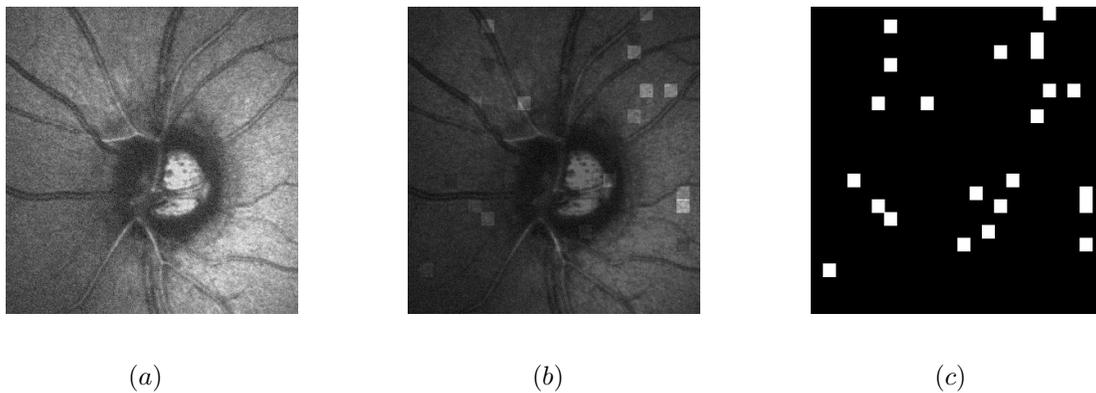


Fig. 2. Semi-simulated image by the modification of image intensities: (a) the original image (b) the simulated image (PSNR=28 dB) and (c) the ground truth.

	Progressor sensitivity	Normal specificity	Non-progressor specificity
Proposed method	87 %	91 %	72 %
POD	78 %	86 %	43 %
TCA	86 %	62 %	21 %
reflectance based method	64 %	100 %	74 %

TABLE III. DIAGNOSTIC ACCURACY OF DIFFERENT METHODS.

(72%) than the POD method (43%). Moreover, it has higher sensitivity in progressor eyes (87%) while maintaining good specificity in normal eyes (91%). Further, in comparison to the reflectance based method, our proposed method provided similar specificity in normal and non-progressing eyes and higher sensitivity in progressing eyes. Increased sensitivity in our proposed scheme is likely because of the fact that we explicitly modeled the spatial dependency of classification among pixels where as dependency is only implicitly accounted by the reflectance based method using familywise type I error rate.

IV. CONCLUSION

In this paper, a Bayesian framework for glaucoma progression detection has been proposed. The task of inferring the glaucomatous changes is tackled with a Monte Carlo Markov Chain algorithm that is used for the first time to our knowledge in the glaucoma diagnosis framework. Modeling and accounting of both spatial dependency of pixels increased the robustness of the proposed change detection scheme compared to the kernel-based method and the threshold method. The validation of the proposed approach using clinical datasets has shown its ability to provide high specificity in non-progressor stable glaucoma eyes and high sensitivity in progressor eyes while maintaining a good specificity in normal eyes.

REFERENCES

- [1] I. Dielemans, PT De Jong, R. Stolk, J.R. Vingerling, D.E. Grobbee, A. Hofman, et al., "Primary open-angle glaucoma, intraocular pressure, and diabetes mellitus in the general elderly population. the rotterdam study," *Ophthalmology*, vol. 103, no. 8, pp. 1271, 1996.
- [2] G.J. Jaffe and J. Caprioli, "Optical coherence tomography to detect and manage retinal disease and glaucoma," *American journal of ophthalmology*, vol. 137, no. 1, pp. 156–169, 2004.

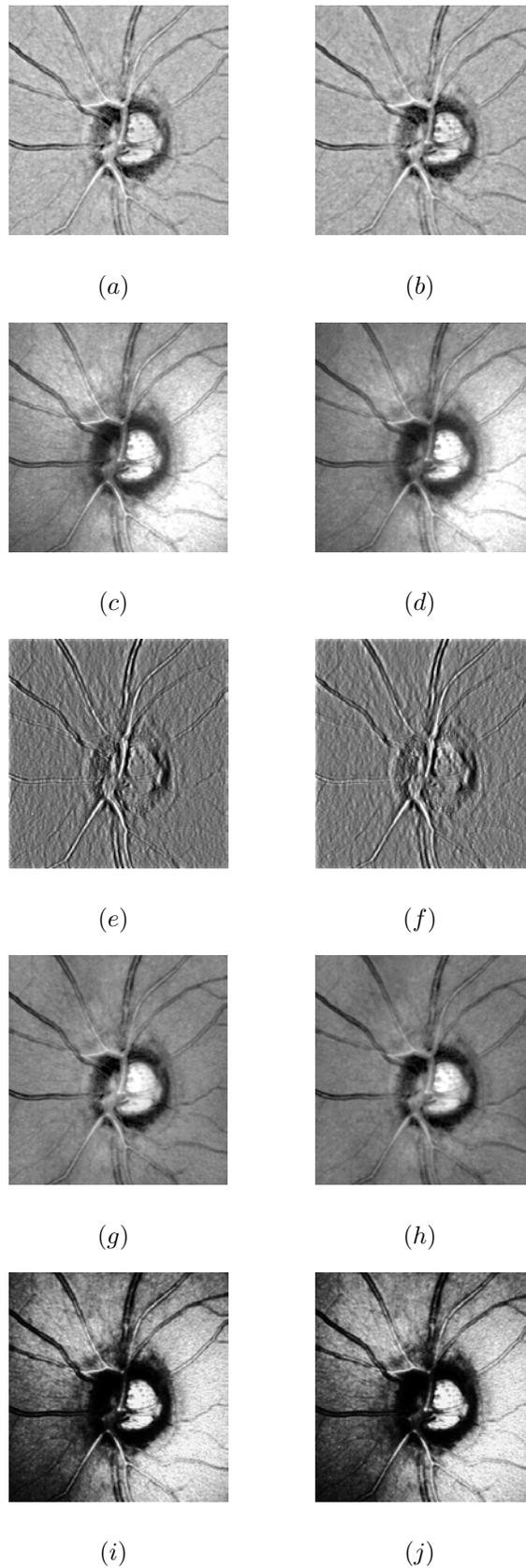


Fig. 3. Intensity normalization results of a baseline and a follow-up images respectively using : (a) (b) the proposed method, (c) (d) the homomorphic filtering, (e) (f) the isotropic diffusion method, (g) (h) the DCT method and (i) (j) the wavelet method.

- [3] A. Mistlberger, J.M. Liebmann, D.S. Greenfield, M.E. Pons, S.T. Hoh, H. Ishikawa, and R. Ritch, "Heidelberg retina tomography and optical coherence tomography in normal, ocular-hypertensive, and glaucomatous eyes," *Ophthalmology*, vol. 106, no. 10, pp. 2027–2032, 1999.
- [4] B.C. Chauhan, J.W. Blanchard, D.C. Hamilton, and R.P. LeBlanc, "Technique for detecting serial topographic changes in the optic disc and peripapillary retina using scanning laser tomography," *Investigative ophthalmology & visual science*, vol. 41, no. 3, pp. 775–782, 2000.
- [5] B.C. Chauhan, D.M. Hutchison, P.H. Artes, J. Caprioli, J.B. Jonas, R.P. LeBlanc, and M.T. Nicolela, "Optic disc progression in glaucoma: comparison of confocal scanning laser tomography to optic disc photographs in a prospective study," *Investigative ophthalmology & visual science*, vol. 50, no. 4, pp. 1682–1691, 2009.
- [6] M. Balasubramanian, D.J. Kriegman, C. Bowd, M. Holst, R.N. Weinreb, P.A. Sample, and L.M. Zangwill, "Localized glaucomatous change detection within the proper orthogonal decomposition framework," *Investigative Ophthalmology & Visual Science*, vol. 53, no. 7, pp. 3615–3628, 2012.
- [7] M Balasubramanian, C Bowd, DJ Kriegman, RN Weinreb, PA Sample, and Zangwill LM, "Detecting glaucomatous changes using only one optic nerve head scan per exam," in *Invest Ophthalmol Vis Sci*, ;51: E-Abstract 2723, 2010.
- [8] S.Z. Li, *Markov random field modeling in computer vision*, Springer-Verlag New York, Inc., 1995.
- [9] Q. Liu, R.J. Scلابassi, C.C. Li, and M. Sun, "An application of map-mrf to change detection in image sequence based on mean field theory," *EURASIP journal on applied signal processing*, vol. 2005, pp. 1956–1968, 2005.
- [10] G. Moser and S.B. Serpico, "Unsupervised change detection from multichannel sar data by markovian data fusion," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 7, pp. 2114–2128, 2009.
- [11] G.J. McLachlan and T. Krishnan, *The EM algorithm and extensions*, vol. 274, Wiley New York, 1997.
- [12] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov chain Monte Carlo in practice*, Chapman & Hall/CRC, 1996.
- [13] G. Celeux, F. Forbes, and N. Peyrard, "Em procedures using mean field-like approximations for markov model-based image segmentation," *Pattern recognition*, vol. 36, no. 1, pp. 131–144, 2003.
- [14] D. Toth, T. Aach, and V. Metzler, "Illumination-invariant change detection," in *Image Analysis and Interpretation, 2000. Proceedings. 4th IEEE Southwest Symposium*. IEEE, 2000, pp. 3–7.
- [15] T. Aach, L. Dumbgen, R. Mester, and D. Toth, "Bayesian illumination-invariant motion detection," in *Image Processing, 2001. Proceedings. 2001 International Conference on*. IEEE, 2001, vol. 3, pp. 640–643.
- [16] D. Toth, T. Aach, and V. Metzler, "Illumination-invariant change detection," in *Image Analysis and Interpretation, 2000. Proceedings. 4th IEEE Southwest Symposium*. IEEE, 2000, pp. 3–7.
- [17] GD Hines, Z. Rahman, DJ Jobson, and GA Woodell, "Single-scale retinex using digital signal processors," in *Global Signal Processing Conference*. Citeseer, 2004, vol. 27.
- [18] W. Wang, B. Li, J. Zheng, S. Xian, and J. Wang, "A fast multi-scale retinex algorithm for color image enhancement," in *Wavelet Analysis and Pattern Recognition, 2008. ICWAPR'08. International Conference on*. IEEE, 2008, vol. 1, pp. 80–85.
- [19] T. Watanabe, Y. Kuwahara, A. Kojima, and T. Kurosawa, "An adaptive multi-scale retinex algorithm realizing high color quality and high-speed processing," *Journal of imaging science and technology*, vol. 49, no. 5, pp. 486–497, 2005.
- [20] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2003, pp. 1055–1055.
- [21] C. Shuyue and Z. Ling, "Chest radiographic image enhancement based on multi-scale retinex technique," in *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*. IEEE, 2009, pp. 1–3.
- [22] N. Brahim, S. Daniel, and D. Gueriot, "Potential of underwater sonar systems for port infrastructure inspection," in *OCEANS 2008*. IEEE, 2008, pp. 1–7.
- [23] KR Joshi and RS Kamathe, "Quantification of retinex in enhancement of weather degraded images," in *Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on*. IEEE, 2008, pp. 1229–1233.
- [24] Joel S Schuman, Tamar Pedut-Kloizman, Ellen Hertzmark, Michael R Hee, Jason R Wilkins, Jeffery G Coker, Carmen A Puliafito, James G Fujimoto, and Eric A Swanson, "Reproducibility of nerve fiber layer thickness measurements using optical coherence tomography," *Ophthalmology*, vol. 103, no. 11, pp. 1889, 1996.
- [25] W. Chen, M.J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 36, no. 2, pp. 458–466, 2006.
- [26] F. Bovolo, G. Camps-Valls, and L. Bruzzone, "A support vector domain method for change detection in multitemporal images," *Pattern Recognition Letters*, vol. 31, no. 10, pp. 1148–1154, 2010.
- [27] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3363–3373, 2006.
- [28] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 874–887, 2005.

Towards a Seamless Future Generation Network for High Speed Wireless Communications*

Kelvin O. O. Anoh¹, Raed A. A. Abd-Alhameed¹, Michael Chukwu², Mohammed Buhari² and Steve M. R. Jones¹

¹Mobile and Satellite Communications Research Centre, University of Bradford (MSCRC), BD1 7DP – United Kingdom

²Centre for Satellite Technology Development, National Space Research Development Agency, Abuja – Nigeria

ABSTRACT—The MIMO technology towards achieving future generation broadband networks design criteria is presented. Typical next generation scenarios are investigated. The MIMO technology is integrated with the OFDM technology for effective space, time and frequency diversity exploitations for high speed outdoor environment. Two different OFDM design kernels (fast Fourier transform (FFT) and wavelet packet transform (WPT)) are used at the baseband for OFDM system travelling at terrestrial high speed for 800MHz and 2.6GHz operating frequencies. Results show that the wavelet kernel for designing OFDM systems can withstand doubly selective channel fading for mobiles speeds up to 280Km/hr at the expense of the traditional OFDM design kernel, the fast Fourier transform.

Keywords—Doppler Effect; Doubly selective fading; frequency-selective fading; OFDM; Wavelet; MIMO

I. INTRODUCTION

The broadband network for next generation communications is a multiple input multiple output (MIMO) network [1-3] defined around 2x2, 4x2 and 4x4 MIMO systems such as the long-term evolution (LTE) and LTE advanced [4]. It uses OFDM for the downlink [5, 6]. The broadband LTE-advanced Release 10, for instance, is now greatly being pursued for building more dependable seamless service and was earlier believed to offer stationary downlink (DL) data size of 1Gb/s and mobile data size of 100Mb/s in wireless communications [7], although it has been extended now to 3Gb/s DL in the subsequent Releases [8]. The air-interface of this technology adopts the OFDM since it can combat inter-symbol interference (ISI) and also converts the frequency selective channel into a flat-fading channel. Also, OFDM is the key air-interface in digital video broadcasting for –terrestrial (DVB-T) and –satellite (DVB-S) wireless communications. Thus, the broadband LTE-advanced uses MIMO-OFDM.

One of the most advocated methods of deploying multi-antenna systems in broadband technology involve the space-time coding [1, 2]. It is the MIMO standard for LTE-Advanced [9]. This is because STBC can exploit the space, time and frequency diversity gains. Most early space-time (ST) coding scheme discuss the case of a frequency-flat fading channel. In practice, the broadband wireless communication channel is frequency-selective and exhibit inter-symbol interference (ISI) that depletes the performance of the system [10]. For broadband wireless systems, the MIMO channels experience frequency-selective fading, which complicates the design of ST codes because of ISI [11].

This is usually addressed by the use of OFDM with MIMO, hence the MIMO-OFDM. The combination of MIMO and OFDM provides an attractive interface solution for next-generation wireless local area networks (WLANS), wireless metropolitan area networks (WMAN) and fourth generation mobile cellular wireless systems [12].

For instance MIMO technology provides two design advantages, namely spatial multiplexing which provides capacity gain, and space-time coding which improves link dependability through diversity gain [11, 12]. These advantages are usually exploited in the design of multi-antenna systems. Applying ST codes only separately over each frequency tone of the OFDM does not exploit frequency diversity. Thus, we choose to construct only space-time frequency (STF) codes that can exploit space, time and frequency diversities.

Broadband network users are necessarily non-stationary. The movement of the user relative to the transmitter cause the signals to fade depending on how fast the channel varies. This is usually discussed as time-selectivity (in addition to the frequency-selectivity) which is the case of a moving receiver. STBC with OFDM can also be studied as space-frequency block codes (SFBC) [13, 14] or space-time frequency block codes (STFBC) [11] over frequency selective channels [15]. The SFBC provides spatial diversity gain only, which sometimes, can be achieved by forward error correction coding with interleaving [12].

The STFBC is preferred over SFBC since the former exploits space, time and frequency gains [10, 11] and the later exploits only space and frequency diversities but not time [10, 13, 14]. The advantage of combining MIMO and OFDM is that MIMO provides capacity and diversity gains while OFDM mitigates the deterring multipath effect. Like this study aims to establish, a high-speed mobile on railway wireless communication technology based on LTE has been discussed in [5] for traditional OFDM architecture. However, it will be shown that the traditional OFDM system is sub-optimal for high speed wireless communications.

In Section II, the OFDM design kernels are presented using the traditional FFT and wavelet transform while the channel system model is discussed in Section III including the Doppler effects. In Section IV, the simulation environment with results are discussed which including the single-input-single-out (SISO) and the MIMO channel, a typical representation of the broadband channel. The conclusion follows in Section V.

*This paper has been presented in part at the CSTD conference week on August 21-22, 2013.

II. ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING

Orthogonal frequency division multiplexing (OFDM) divides wideband into many narrow-bands to increase the data rate. Traditionally kernel for designing OFDM is the FFT. Another kernel that provides signal representation half-way in time-frequency plane, the wavelet transform, is also discussed.

A. FFT-OFDM

OFDM system is used in baseband as modulation method that divides a selected bandwidth into smaller and many narrow bands. In time domain, an N -point FFT OFDM system can be represented as:

$$s[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S_k e^{j2\pi mk/N}, \quad n=0, 1, 2, \dots, N-1 \quad (1)$$

where $s[n]$ is the discrete form of $s(t)$, N is the number of sub-channels, $1/\sqrt{N}$ is a scaling factor with n as the index of the prevalent subcarrier. S_k is the QPSK mapped input symbol of k^{th} sub-channel. The number of FFT points used is equivalent to the number of narrowband sub-channels over which the input symbols are multiplexed. Each of the resulting narrowband sub-channels is modulated by the mapped input bits. Cyclic prefix (CP) at least equal to the length of the channel response, L is pre-appended to each OFDM symbol to combat ISI. To account for the CP, Equation 1a can be modified as:

$$s[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S_k e^{j2\pi mk/N}, \quad -N_g \leq n \leq N-1 \quad (1b)$$

In Equation 1b, N_g denotes the length of CP pre-appended to every OFDM symbol.

B. Wavelet – OFDM

The discrete wavelet transform (DWT) can be used to study multicarrier systems as in [16]. It represents signals in time-frequency domain such that the signal exists neither purely in frequency domain nor purely in time domain. For a mapped input symbol S_k to be transformed by the DWT, the time domain output can be realized from [17];

$$s[n] = \sum_k \sum_{m=0}^{M-1} S_{k,n} \varphi_{m,n}(t) \quad (2)$$

where $S_{k,n}$ represents the n^{th} symbol which modulates the m^{th} -waveform of the k^{th} -constellation. $\varphi_{m,n}(t)$ represents the complex orthogonal DWT basis function similar to the traditional OFDM as:

$$\varphi_{m,n} = \begin{cases} 1 & n = m \\ 0 & \text{elsewhere} \end{cases} \quad (3)$$

where m and n are scales and shifts respectively. If n is the index of each discrete wavelet symbol $s[n]$ of the continuous time symbol $s(t)$, then the wavelet transform is defined as [18]:

$$\psi_{k,a}(t) = e^{i\pi t^2} \quad (4)$$

Now, let a continuous wavelet function be expressed as:

$$\psi_{k,a}(t) = \frac{1}{\sqrt{k}} \psi\left(\frac{t-a}{k}\right) \quad (5)$$

where k and a are the scaling and shifting parameters respectively and $\psi(\cdot)$ is called the mother wavelet. Then, from Equations 5 and 4 the resulting continuous transform can be represented as:

$$s_{CWT}(\tau, k) = \frac{1}{\sqrt{k}} \int_{-\infty}^{\infty} \exp\left[i\pi \frac{(t-\tau)^2}{k^2}\right] s(t) dt \quad (6)$$

Equation 6 has the advantage of time and frequency diversities unlike the FFT transform that has only frequency diversity advantage. In fact, it has been explored that orthogonal wavelet-based OFDM is more robust to ICI and ISI problems than the FFT-based OFDM [19-21]. Meanwhile, the CP pre-appended to the OFDM symbol when using the FFT is omitted in the wavelet-based OFDM. This provides additional 25% spectral efficiency.

III. SYSTEM MODEL

In the following, the system architecture is discussed for the forms of OFDM kernels just discussed. Since the fading for each transmission branch in a multi-antenna system between a pair of transmit and receive antennas are usually independent, the probability that the information is detected correctly is increased [11]. Thus, the probability that the transmit signal will be decoded correctly increases with the diversity order. On that note, let there be a MIMO-OFDM system with N_T antennas and N_R receive antennas defined for N -subcarrier OFDM existing for each transmit antenna. Let there be also n^{th} OFDM subcarrier symbol, $s_k^p(n)$, transmitted over p^{th} transmit antenna at k^{th} OFDM symbol duration. It can be identified that the OFDM symbol of this characteristics– $s_k^p(n)$ – will exploit frequency, space and time diversity gains by the notations following the OFDM symbol, $s(\cdot)$.

A. General System Model

For the FFT-based OFDM, IFFT is applied and the CP is pre-appended to each of the symbols. In the case of wavelet-based OFDM, the wavelet transform is applied to the codeword and then transformed to frequency domain – to fully exploit the frequency diversity gain. The channel is entirely doubly selective (except for times at which the Doppler frequency is explicitly null), although the OFDM technology transforms the frequency-selective channel into a frequency flat fading channel.

In the receiver (after removing the CP for the FFT-OFDM case), the received symbol is OFDM-dettransformed using FFT or forward wavelet transform as the case may be. So for each transmit symbol from one transmit antenna,

$$R_k^q(n) = \sum_{p=1}^{N_T} H_{p,q}(n) S_k^p(n) + Z_k^q(n), \quad \forall n=0, 1, \dots, N \quad (7)$$

where $R_k^q(n)$ is the n^{th} received OFDM symbol with $H_{p,q}(n)$ as the frequency domain channel gain corresponding

to p^{th} transmit antenna and q^{th} receive antenna respective for every subcarrier. The channel gain corresponds to the individual resolvable paths traversed by the OFDM symbols with impulse response vector:

$$\mathbf{h}_{p,q} = [h_{p,q}(0), h_{p,q}(1), h_{p,q}(2), \dots, h_{p,q}(L-1)] \in C^{L \times 1} \quad (8a)$$

where $C^{L \times 1}$ is a complex matrix of $L \times 1$ dimension. The channel gains then follow as:

$$H_{p,q}(n) = \sum_{l=0}^{L-1} h_{p,q}(l) \cdot e^{-j2\pi f\tau_l} \quad (8b)$$

where $e^{-j2\pi f\tau_l}$ is the DFT factor for transforming from time to frequency domain, $Z_k^q(n)$ in Equation 7 is the additive white Gaussian noise (AWGN) while $S_k^p(n)$ is the space-time-frequency (STF) codewords of the N_T transmit antenna, N_R receive antenna of N_Δ OFDM symbol interval. By imaging the STF in three-dimension (3D), STF will be seen as a point on that plane [15]. Like in [22, 23], the diversity gain will increase by the number of subcarriers used such as $N_T N_\Delta N_R$. For each sub-channel, the STF codeword is constructed as follows:

$$S(n) = \begin{bmatrix} S_1^1(n) & \Lambda & S_1^{N_T}(n) \\ M & O & M \\ S_{N_\Delta}^1(n) & \Lambda & S_{N_\Delta}^{N_T}(n) \end{bmatrix} \in C^{N_\Delta \times N_T} \quad (9)$$

In Equation 9, $S(n)$ can be taken as orthogonal space-time block code already discussed in [24, 25]. In terms of STBC, Equation 7 can then be expressed as [26];

$$R(n) = \sqrt{\frac{\rho}{N_T}} H(n)S(n) + Z(n)$$

where ρ is the SNR respective to each transmission branch (antenna), $Z(\cdot)$ is additive noise with circular symmetry and zero mean. To exploit the STBC basic criterion, Equation 9 must be designed according to the design criteria of [25] and given as:

$$s = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \quad (10)$$

Traditionally, full diversity of an orthogonal STBC system is derived from the pairwise error probability between two different codeword matrices (usually transmitted and received), say S and S' bounded as [26]:

$$P(S \rightarrow S') \leq \binom{2r-1}{r} \left(\prod_{i=0}^r \gamma_i \right)^{-1} \left(\frac{\rho}{N_T} \right)^{-r} \quad (11)$$

Equation 11 can also be used to discuss the BER statistics of space-frequency block coding (SFBC); In fact, the BER statistic of STBC and SFBC are similar [27] where the upper bound, r , is the rank of: $(S-S')R(S-S')^H$ where $R = E\{HH^H\}$

is the correlation matrix of the channel matrix and $\gamma_1, \gamma_2, \dots, \gamma_r$ are nonzero eigenvalues of $(S-S')R(S-S')^H$. The superscript in $(\cdot)^H$ is a Hermitian transpose operator and H represents the channel matrix. Signals S and S' are transmitted and received signals respectively. Recently, the BER statistic of the STFBC has been explored in [15, 28].

B. Example of an orthogonal STBC – the Alamouti Code

The traditional Alamouti STBC is followed in which the system will be assumed to be constant over two symbol periods. s_1 and s_2 are transmitted in the first symbol period, and their conjugates (s_1^* and $-s_2^*$) are transmitted in the second symbol period. In Figure 2, this is replaced with the FFT. In Figure 2, s_1 will be transmitted using antenna T1 and s_2 using antenna T2 in the first time period. In the second time period, the negative conjugate of s_2 , i.e. $(-s_2^*)$, and the conjugate of s_1 , i.e. s_1^* , will be transmitted over antennas T1 and T2 respectively.

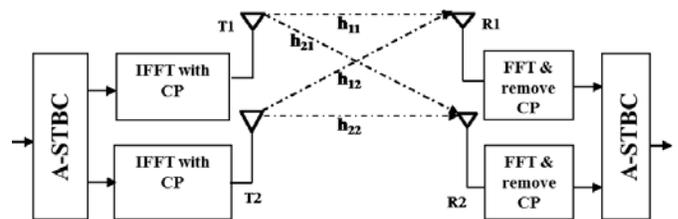


Fig. 2. Simple 2x2 MIMO-OFDM Modulation Architecture

Let $s(n)$ be the discrete equivalent of the original OFDM information $s(t)$, then s_1 and s_2 can be obtained as [29, 30]:

$$s_1 = \frac{1}{\sqrt{2}} s(n) \quad (12a)$$

$$s_2 = \frac{1}{\sqrt{2}} s(n) \quad (12b)$$

The code structure formed from Equation 12 following Equation 9 and 10 is an STF codeword. In that case, S , S_1 and S_2 are frequency domain contents of s , s_1 and s_2 respectively. Equation 12 will convolve with the doubly selective frequency domain channel gains of h_1 and h_2 , as of the channel transfer function in Equation 8b such as:

$$R_1 = H \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} + Z_1 \quad (13a)$$

and,

$$R_2 = H \begin{bmatrix} -S_2^* \\ S_1^* \end{bmatrix} + Z_2 \quad (13b)$$

$$\text{where } H = [H_1 \ H_2] = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \quad (14)$$

By Equation 14, there are up to 2 number of receive antennas. The received signals estimates of the transmitted symbol will then be computed. These estimates are obtained as:

$$D_1 = H_1^H R_1 + R_2^H H_2 \quad (15a)$$

and,

$$D_2 = H_2^H R_1 - R_2^H H_1 \quad (15b)$$

where $(\cdot)^H$ denotes a Hermitian operator. The foregoing description limits $N_T = 2$ and up to $N_R = 2$.

C. Doppler Effect in OFDM Systems - Single Antenna Case

Consider a traditional frequency-selective channel as (in time domain) [31];

$$h(t, \tau) = \sum_{l=0}^{L-1} \alpha(l) e^{j\phi_l} \delta(t - \tau_l) \quad (16)$$

where $\alpha(l)$ is the l^{th} amplitude of the l^{th} path traversed by the OFDM symbol, $\phi_l(t) = 2\pi\{(f_c + f_{d,l})\tau_l - f_{d,l}t\}$ is the phase, f_c is the carrier frequency, τ is the path delay and $\delta(\cdot)$ is the dirac delta function. The received OFDM signal in the receiver will be;

$$r(t, \tau) = \frac{1}{\sqrt{LT}} \sum_{l=0}^{L-1} \alpha(l) e^{j\phi_l} s(t - \tau_l) \quad (17)$$

So, each of the received signals is delayed by τ_l and scaled by $\alpha(l)e^{j\phi_l}$. If frequency domain content of Equation 16 is considered as well as the doubly-selective characteristics of the channel (i.e. t and f):

$$H(t, f) = \sum_{l=0}^{L-1} h(t, \tau_l) e^{-j(2\pi l)/N} \quad (18)$$

Notice that $h(t, \tau_l) \in C^{L \times 1}$, where $C^{L \times 1}$ is a complex matrix of $L \times 1$ dimension. The term ϕ_n influences the phase rotation of received OFDM symbols and it involves $\phi_l = \phi_n + \phi_{off}$, where f_d is the maximum Doppler frequency and ϕ_{off} is the carrier frequency offset. In this work, so set $\phi_{off} = 0$. The transmit signal convolves with Equation 18 such as;

$$R(t, f) = \frac{1}{\sqrt{LT}} \sum_{l=0}^{L-1} H(t, f) S(f) \quad (19)$$

where $S(f)$ is the frequency content of $s(t - \tau_l)$. The Doppler frequency shifts the signal from the carrier frequency which in turn affects the amplitude of the received symbol. For signal compensation, we proceed as follows;

$$Y(t, f) = \frac{H(t, f)^H R(t, f)}{\|H(t, f)\|^2} \quad (20)$$

where $H(t, f)^H$ is the Hermitian of the matrix $H(t, f)^H$. Equation 21 is then transformed into the time domain. If at a time that $|H(t, f)| = 0$, then Equation 21 is modified as:

$$Y(t, f) = \frac{H(t, f)^H R(t, f)}{\|H(t, f)\|^2 + \varepsilon} \quad (21)$$

This is equivalent to:

$$Y(t, f) = \frac{H(t, f)^H \cdot H(t, f) S(t, f)}{\|H(t, f)\|^2 + \varepsilon} + \frac{H(t, f)^H \cdot Z(t, f)}{\|H(t, f)\|^2 + \varepsilon}, \forall 0 \leq \varepsilon \leq 1 \quad (22)$$

where $(\cdot)^H$ is the Hermitian operator, $H(t, f)^H \cdot H(t, f)$ is unitary. The term ε is an error correction factor to reduce the magnification of the noise parameter due to equalization.

D. Doppler Effect in OFDM Systems for Multi-Antenna Case

In earlier studies of the ST codes for MIMO systems, the channel model is usually treated as frequency flat-fading channel. However, each of the transmission branches is independent. This necessitates that the likelihood of decoding the signals correctly increases. MIMO-OFDM systems for realistic broadband channel is non-frequency flat, instead frequency-selective which the OFDM converts to flat-fading channel to eliminate ISI. Meanwhile, for varying mobile speeds the channel also varies in time contributing time-selectivity so that the broadband channel be doubly (time-frequency) selective. This is the case for each transmission channel branch in this study discussed in Section III-A. Thus, for the STF block codes used in this study, each of the transmission branches is necessarily doubly selective so that the Doppler frequency can be accounted for at varying mobile speeds.

IV. SIMULATION RESULTS AND DISCUSSION

The simulation was carried out, first for the single antenna transmission and then the MIMO antennas systems. The operating frequency deployed is the new operating frequency 800MHz and 2.6GHz for LTE-advanced. The rest higher frequencies (900MHz and 1800MHz for GSM) were not investigated. In general, the QPSK modulation scheme was used over OFDM system whose channel delay spread is in the order of the outdoor environment at a sampling frequency of 20KHz.

The mobile speeds are typical for standard road (and railway) travels - 120Km/hr, 150 KM/hr, 200Km/hr, 240Km/hr and 280Km/hr are investigated. The number of OFDM subcarriers used is 64 with 25% cyclic prefix for the case of FFT-OFDM. This is not applicable to wavelet-OFDM by the considerations described in this work, thus offering the wavelet-OFDM further 25% spectrum efficiency at the expense of the FFT-OFDM.

A. SISO OFDM System

The SISO-OFDM systems considered are for the FFT-based and wavelet-based OFDM systems. Each has been simulated over the same design parameters except for the FFT-OFDM using cyclic prefix of 25% only.

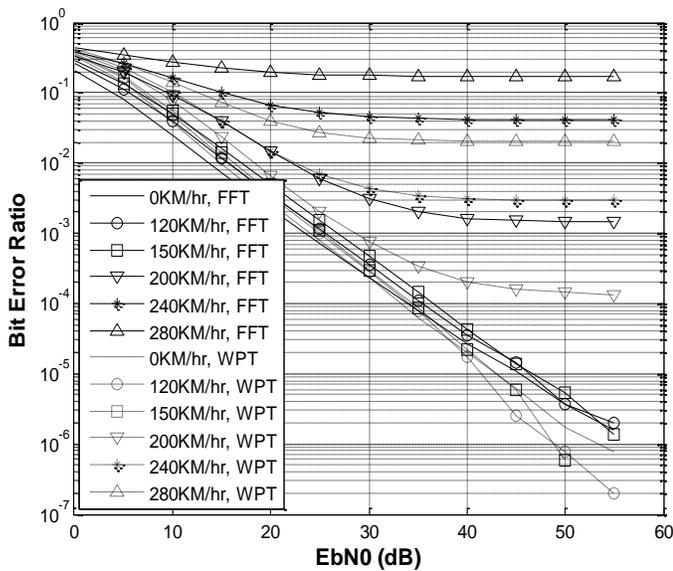


Fig. 3. Comparison of FFT-OFDM and wavelet-OFDM at 800MHz for SISO Systems

Recall that the FFT-OFDM is highly robust over frequency selective channel [32]; $f_d = 0$ in that case. From Figures 3 and 4, it can be seen that the FFT-OFDM contended favourably with the wavelet-OFDM when the channel is dominantly frequency-selective. However, for other speeds shown in the results, wavelet-OFDM well outperformed the FFT-OFDM.

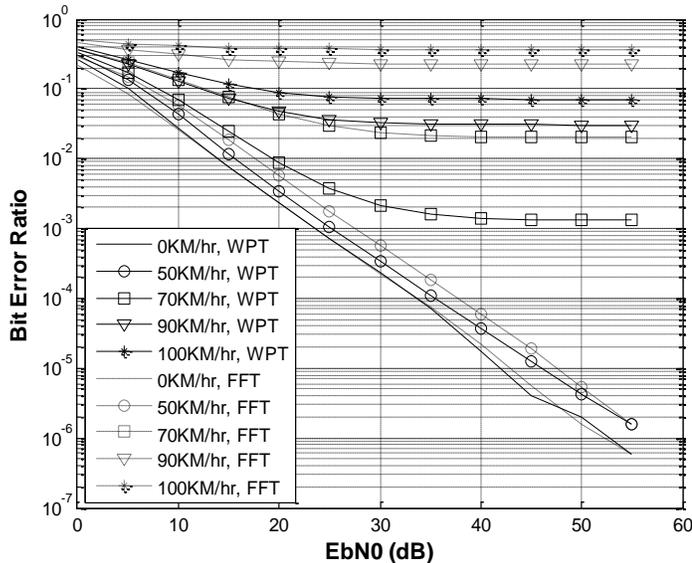


Fig. 4. Comparison of FFT-OFDM and wavelet-OFDM at 2.6GHz for SISO Systems

Results in Figure 4 show the consistence of wavelet-OFDM robustness over doubly selective transmission. Meanwhile, the results reveal that the 2.6GHz has poorer performance than the 800Hz operating carrier frequency since the fading bandwidth of the channel increases with the operating carrier frequency.

B. MIMO-OFDM

In MIMO-OFDM case, Figure 5 depicts the performance of FFT-OFDM and wavelet-OFDM for a MISO System for different mobile speeds.

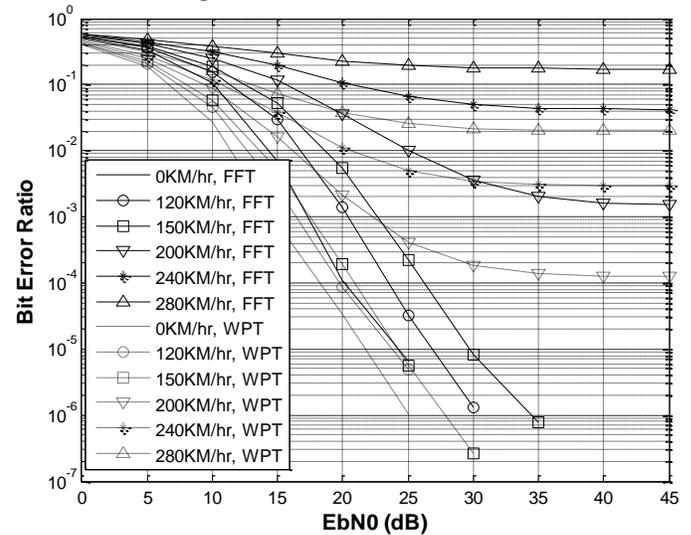


Fig. 5. Comparison of a 2x1 FFT-OFDM and wavelet-OFDM Systems at 800MHz

It is obvious that the wavelet-OFDM outperformed the FFT-OFDM for all speeds. In the next result (Figure 6), it is shown that the result in Figure 5 is improved for a 2x2 MIMO system. We conjecture that this result in Figures 5 and 6 may be consistent for 2x1, 2x2 or even higher order MIMO configurations above 2.6GHz operating frequency, with wavelet-OFDM dominating FFT-OFDM in performance.

In general, the degree at which the OFDM signal shifts from the carrier (or equivalently centre frequency in the baseband) frequency, f_c , by Doppler shift depends on the property of the signal volunteered by the baseband transform. The Doppler Effect leads to the loss in orthogonality of the sub-channels of the OFDM system. Wavelet transforms the signal to possess time and frequency diversities whereas the FFT provides frequency-selectivity (diversity) only. The robustness of the signal against the Doppler shift is then dependent on characterizing signal property nominated by the baseband transform used in the signal multiplexing.

V. CONCLUSION

Future generation wireless communication design of radio air-interface such as the OFDM was presented for two different design kernels – fast Fourier transform and wavelet packet transform. Both OFDM systems were investigated for different mobile speeds which account for varying Doppler frequencies effect in high speed wireless communications typical of real-life scenarios. OFDM converts frequency-selective channel into flat fading channel thus very robust in frequency-selective transmission. However, the FFT-OFDM scheme was observed to be sub-optimal compared to the wavelet transform for OFDM design for high speed mobiles. Results reveal that the wavelet-OFDM has very robust ability in repelling the inter-carrier interference volunteered by fast fading channels due to Doppler Effect than the FFT-OFDM.

Thus, the wavelet transform can be a good candidate for the seamless broadband network radio air-interface in future wireless communications.

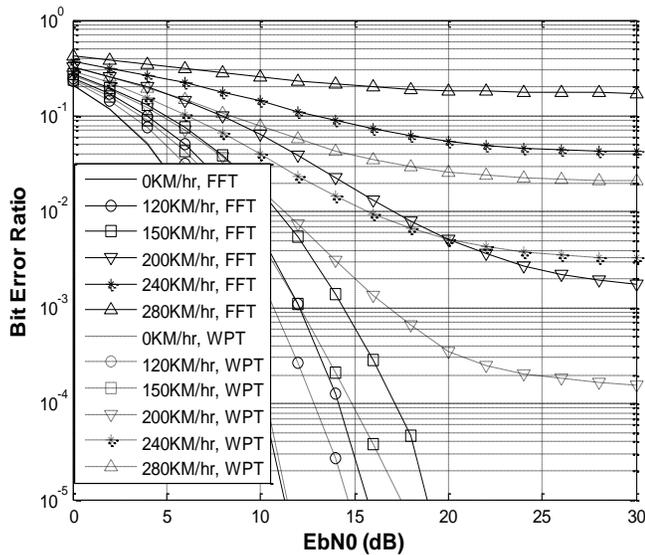


Fig. 6. Comparison of a 2x2 FFT-OFDM and wavelet-OFDM Systems at 800MHz

REFERENCES

[1] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, pp. 271-294, 2004.

[2] H. Yang, "A road to future broadband wireless access: MIMO-OFDM-based air interface," *IEEE Communications Magazine*, vol. 43, pp. 53-60, 2005.

[3] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-advanced-evolving LTE towards IMT-advanced," in *IEEE 68th Vehicular Technology Conference*, 2008. VTC 2008-Fall. 2008, pp. 1-5.

[4] P. E. Mogensen, T. Koivisto, K. I. Pedersen, I. Z. Kovacs, B. Raaf, K. Pajukoski, and M. J. Rinne, "LTE-advanced: the path towards gigabit/s in wireless mobile communications," in *Wireless VITAE 2009. 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*, 2009, pp. 147-151.

[5] G. Tingting and S. Bin, "A high-speed railway mobile communication system based on LTE," in *2010 International Conference On Electronics and Information Engineering (ICEIE)*, 2010, pp. V1-414-V1-417.

[6] A. Omri, R. Hamila, M. O. Hasna, R. Bouallegue, and H. Chamkhia, "Estimation of highly Selective Channels for Downlink LTE MIMO-OFDM System by a Robust Neural Network," *JUSPN*, vol. 2, pp. 31-38, 2011.

[7] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-advanced: next-generation wireless broadband technology [Invited Paper]," *IEEE Wireless Communications*, vol. 17, pp. 10-22, 2010.

[8] J. Wannstrom. (2012). LTE-Advanced. Available: <http://www.3gpp.org/lte-advanced> (Accessed on Sept. 19, 2013)

[9] Cisco, "802.11ac: The Fifth Generation of Wi-Fi Technical White Paper," ed, Aug. 2012.

[10] T. Bao and L. Yongling, "Improved space-time-frequency block code for MIMO-OFDM wireless communications," in *2012 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC)*, 2012, pp. 538-541.

[11] W. Zhang, X. Xiang-Gen, and K. Ben Letaief, "Space-time/frequency coding for MIMO-OFDM in next generation broadband wireless systems," *IEEE Wireless Communications*, vol. 14, pp. 32-43, 2007.

[12] H. Bolcskei, "MIMO-OFDM wireless systems: basics, perspectives, and challenges," *IEEE Wireless Communications*, vol. 13, pp. 31-37, 2006.

[13] H. Bolcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *2000 IEEE Wireless Communications and Networking Conference*, 2000. WCNC. 2000, pp. 1-6.

[14] H. Bolcskei and A. J. Paulraj, "Space-frequency codes for broadband fading channels," in *2001 IEEE International Symposium on Information Theory*, 2001. Proceedings. 2001, p. 219.

[15] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2465-2476, 2002.

[16] O. O. Anoh, N. T. Ali, R. Abd-Alhameed, S. M. Jones, and Y. A. Dama, "On the performance of DWT and WPT modulation for multicarrier systems," in *2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2012, pp. 348-352.

[17] A. Jamin and P. Mähönen, "Wavelet packet modulation for wireless communications," *Wireless Communications and Mobile Computing*, vol. 5, pp. 123-137, 2005.

[18] H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural, "Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms," *JOSA A*, vol. 11, pp. 547-559, 1994.

[19] B. Negash and H. Nikookar, "Wavelet-based multicarrier transmission over multipath wireless channels," *Electronics Letters*, vol. 36, pp. 1787-1788, 2000.

[20] B. Negash and H. Nikookar, "Wavelet based OFDM for wireless channels," in *IEEE VTS 53rd Vehicular Technology Conference*, 2001. VTC 2001 Spring. 2001, pp. 688-691.

[21] Y. Zhang and S. Cheng, "A novel multicarrier signal transmission system over multipath channel of low-voltage power line," *IEEE Transactions on Power Delivery*, vol. 19, pp. 1668-1672, 2004.

[22] F. Fazel and H. Jafarkhani, "Quasi-orthogonal space-frequency and space-time-frequency block codes for MIMO OFDM channels," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 184-192, 2008.

[23] D. H. Nguyen and H. H. Nguyen, "Diversity and Coding Gains of Space-Time-Frequency Coded MIMO-OFDM," in *IEEE, 2008 Vehicular Technology Conference*, 2008. VTC Spring 2008., pp. 518-523.

[24] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Transactions on Communications*, vol. 49, pp. 1-4, 2001.

[25] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451-1458, 1998.

[26] W. Su, Z. Safar, M. Olfat, and K. R. Liu, "Obtaining full-diversity space-frequency codes from space-time codes via mapping," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2905-2916, 2003.

[27] D.-B. Lin, P.-H. Chiang, and H.-J. Li, "Performance analysis of two-branch transmit diversity block-coded OFDM systems in time-varying multipath Rayleigh-fading channels," *IEEE Transactions on Vehicular Technology*, vol. 54, pp. 136-148, 2005.

[28] W. Zhang, X.-G. Xia, and P.-C. Ching, "High-rate full-diversity space-time-frequency codes for broadband MIMO block-fading channels," *IEEE Transactions on Communications*, vol. 55, pp. 25-34, 2007.

[29] K. O. O. Anoh, R. A. Abd-alhameed, J. M. Noras, and S. M. R. Jones, "Wavelet Packet Transform Modulation for Multiple Input Multiple Output Applications," *IJCA*, vol. 63 - Number 7, pp. 46 - 51, 2013.

[30] K. O. O. Anoh, R. A. Abd-Alhameed, and S. M. R. Jones, "An Investigation of FrFT Modulation and FFT-OFDM for MIMO Applications," presented at the 14th Annual Post Graduate Symposium on Convergence of Telecommunications, Networking and Broadcasting, Liverpool, UK, 24 - 25th June, 2013.

[31] G. L. Stüber, *Principles of mobile communication*: Springer, 2011.

[32] T. Wang, J. G. Proakis, E. Masry, and J. R. Zeidler, "Performance degradation of OFDM systems due to Doppler spreading," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 1422-1432, 2006.